Available online at www.sciencedirect.com
ScienceDirect
European Journal of Operational Research 190 (2008) 659–671
EUROPEAN JOURNAL OF OPERATIONAL RESEARCH
www.elsevier.com/locate/ejor

Decision Support

Forming coalitions and the Shapley NTU value ☆

Juan J. Vidal-Puga *

Departamento de Estatistica e IO and Research Group in Economic Analysis, Universidade de Vigo, 36200 Vigo (Pontevedra), Spain

Received 20 October 2006; accepted 2 July 2007
Available online 14 July 2007

Abstract

A simple protocol for coalition formation is presented. First, an order of the players is randomly chosen. Then, a coalition grows by sequentially incorporating new members in this order. The protocol is studied in the context of non-transferable utility (NTU) games in characteristic function form. If (weighted) utility transfers are feasible when everybody cooperates, then the expected subgame perfect equilibrium payoff allocation anticipated before any implemented game is the Shapley NTU value.
© 2007 Elsevier B.V. All rights reserved.

Keywords: Shapley NTU value; Subgame perfect equilibrium; Sequential formation of coalitions

1. Introduction

Endogenous formation of coalitions has been widely studied in the literature. A common approach is to assume that any number of players can simultaneously join a coalition. For example, Hart and Kurz (1983), Chatterjee et al. (1993), Bloch (1996), Okada (1996), and Ray and Vohra (1999) consider situations where coalitions form and abandon the game.

A different approach is to assume that only bilateral mergers occur, and the newly created coalitions keep merging among themselves until a stable coalition structure is created. This is the approach followed by Gul (1989) and Macho-Stadler et al. (2006).

Following the latter approach, this paper studies situations in which a coalition is formed by the sequential inclusion of new members. In contrast with previous models, the collusion is not parallel. Instead, a size-increasing coalition arises that swallows up other individuals like a snowball. An individual agent may choose to join this coalition or continue alone. International bodies like the European Union or NATO provide relevant examples of this coalition formation protocol. In the case of the European Union, in the period between

☆ Financial support by the Spanish Ministerio de Ciencia y Tecnologia and FEDER through grants BEC2002-04102-C02-01 and SEJ2005-07637-C02-01/ECON and by the Xunta de Galicia through grants PGIDIT03PXIC30002PN and PGIDIT06PXIC300184PN is gratefully acknowledged.
* Tel.: +34 986 802014; fax: +34 986 812401.
E-mail address: vidalpuga@uvigo.es

0377-2217/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.ejor.2007.07.006

the customs agreement between Belgium, Luxembourg and the Netherlands came into force (1948) and the current negotiations with Turkey and the Balkan candidates, one re-foundation (1951) and five enlargements (1973, 1981, 1986, 1995 and 2004) have taken place. Even though more than one country officially joined the Union at the same time, negotiations took place individually with each candidate and independently from negotiation with the other candidates. Thus, from a practical perspective, it can be considered that the countries joined the union sequentially. Moreover, accession of new members has caused changes to the laws governing the coalition. In the last scheduled enlargement, for example, which took place in 2004, the old voting system used for a Union of 15, became obsolete and was adapted for a Union of 25 members. Such a change in the laws requires the unanimity of all members, and a lack of unanimity may cause the enlargement to be aborted. In 2001, people in Ireland voted in a referendum against the Nice Treaty, thereby placing in jeopardy the enlargement process. In a second referendum held in 2002, the Irish people voted in favor.

Another example of a coalition formed by a sequential entry of new members is given by situations in which a big company grows by absorbing smaller companies (e.g. Microsoft's policy in the second half of the 1990s). Since such enlargements are often focused on new business sectors, a change in policy in the new enlarged company is frequent.

In this paper, this process is modelled in the set of NTU games by a simple negotiation mechanism. The bargaining model is formulated as a finite-length extensive form game with perfect information. The main idea of the mechanism is the creation and enlargement of a coalition of players. The members of this coalition agree on a rule for sharing their resources. Players outside the coalition can apply to enter the coalition by agreeing with the previously established internal rule. However, in the admission negotiation, candidates may also propose changes to the internal rule. Nonetheless, unanimity will be required for the coalition to change its rules.

In NTU games in which the feasible set associated with the grand coalition has to be a half space,[1] this protocol enables players to obtain at least their marginal contribution in each order for the induced λ-transfer TU game as defined by Shapley (1969). Hence, the Shapley NTU value (Shapley, 1969; see also Aumann, 1985a) arises in equilibrium as an expected payoff allocation, given that the order in which players join the coalition is randomly chosen with the same probability. It is however possible to add more stages (as in Bag and Winter, 1999; Mutuswami and Winter, 2002) to get the value precisely and not in expectation (see Section 5).

Hart and Moore (1990, footnote 11) described a different sequential protocol yielding the Shapley value for TU games in expected terms. This protocol is stated in terms of contracts and guarantees each player his marginal contribution as well. Stole and Zwiebel (1996) studied another sequential protocol yielding the Shapley value for an underlying TU game involving wage bargaining in a firm. In Stole and Zwiebel's model, negotiations are always bilateral between the firm and individual workers. Disagreement between the firm and a worker implies that the previous agreements are renegotiated with this last worker out of the game. Pérez-Castrillo and Wettstein (2001) presented a different protocol yielding the Shapley value in zero-monotonic games. In their model, the players bid for the right to be the proposer. If a proposal is rejected, then the proposer leaves the game and the rest of the players begin the negotiation process without him.

The model presented in this paper differs from these ones in two important aspects: First of all, provisional agreement on a rule takes into account the possibility that not all the remaining players will join the coalition. Hence renegotiations do not happen. Secondly, players are never excluded: they may have the chance to join the coalition in the future if a candidate, in his own admission negotiation, proposes to do so.

In a different framework, Maskin (2003) presented a bargaining protocol in which more than one coalition may form at a time. The next player in the order receives offers from these coalitions to join them. Each player can either join one of the existing coalitions, or form his own coalition and make offers to the next players in the order. As opposed to Maskin's, in the model presented in this paper at most one proper coalition may form, and the proposals are made by the candidates.

Alternative sequential protocols may be found in Chae and Yang (1994) and Suh and Wen (2003, 2006) for pure bargaining problems.

[1] Notice that these NTU games are not equivalent to transfer utility games, because the half space condition is not imposed on proper subcoalitions.

Hart and Mas-Colell (1996) designed a non-cooperative mechanism such that the consistent Shapley value (Maschler and Owen, 1989, 1992) arises, as a subset of the limit of stationary subgame perfect Nash equilibria. As Hart (2004, footnote 9) points out, it would be of interest to obtain similar results for other extensions of the Shapley value (Shapley, 1953) and the Nash solution (Nash, 1950) to NTU games, like the Harsanyi value (Harsanyi, 1963) or the Shapley NTU value.

In Hart and Mas-Colell's model, players must propose payoff allocations. However, it is not uncommon in real negotiations that proposals refer not to final payoff allocations, but to general policies, rules or solutions. For example, in the context of bargaining problems, van Damme (1986), Naeve-Steinweg (1999), and Trockelm (2002) propose mechanisms in which players' proposals are solution concepts (like the Nash solution or the Kalai–Smorodinsky solution). Similarly, in Naeve (1999), players report the set of feasible payoff allocations.

For more than two players, the payoff allocation proposed by a solution concept depends on which players are involved. Following this line of thought, in our mechanism the proposals are rules that assign a different payoff allocation to each subcoalition, i.e. a rule is a function γ that assigns to each coalition S a vector $\gamma(S) \in \mathbb{R}^S$ that is feasible for S.

The particular class of games to which we restrict ourselves are games (N, V) such that $V(N)$ is delimited by a hyperplane. This class includes TU games, but also includes some proper NTU games, for example, prize games (Hart, 1994).

Section 2 describes the notation used. The mechanism and the main results are presented in Section 3. Several examples are analyzed in Section 4. Some concluding remarks are presented in Section 5. Finally, the results are proved in Appendix A.

2. Preliminaries

Let $\mathbb{R}$ be the set of real numbers. Similarly, $\mathbb{R}_{++}$ and $\mathbb{R}_{+}$ are the set of positive and non-negative real numbers, respectively. Given any finite set S, we denote by $|S|$ the cardinality of S, and by $\mathbb{R}^S$ the set of all functions from S to $\mathbb{R}$. The sets $\mathbb{R}^S_{++}$ and $\mathbb{R}^S_{+}$ are defined accordingly. We also denote by 2^S the power set of S, i.e. $2^S := \{T : T \subset S\}$. A member x of $\mathbb{R}^S$ is an $|S|$-dimensional vector whose coordinates are indexed by members of S; thus, when $i \in S$, we write x_i for $x(i)$. If $x \in \mathbb{R}^T$ (or $x \in \mathbb{R}^N$) and $T \subset S$ (or $T \subset N$), we write x_T for the restriction of x to T, i.e. the members of $\mathbb{R}^T$ whose ith coordinate is x_i. With some abuse of notation, given $x \in \mathbb{R}^S$ and $a \in \mathbb{R}$, we write $(x, a) \in \mathbb{R}^{S\cup\{i\}}$ for the member of $\mathbb{R}^{S\cup\{i\}}$ whose ith coordinate is a and whose restriction to S is x. Given $x, y \in \mathbb{R}^S$, we write $x \geqslant y$ if $x_i \geqslant y_i$ for all $i \in S$.

Let $N = \{1, 2, \ldots, n\}$ be a finite set of *players*. Non-empty subsets of N are called *coalitions*. A *non-transferable utility (NTU) game* on N is a correspondence V that assigns to each coalition S a subset $V(S) \subset \mathbb{R}^S$ that satisfies the following properties:

(A1) For each $S \subset N$, the set $V(S)$ is non-empty, closed, convex, *comprehensive* (i.e., if $x \in V(S)$ and $y \leqslant x$, then $y \in V(S)$), and *bounded from above* (i.e., for each $x \in \mathbb{R}^S$, the set $\{y \in V(S): y \geqslant x\}$ is compact).

(A2) *Normalization*: For each $i \in N$, the maximum of $\{x : x \in V(\{i\})\}$, which we denote by ω_i, is non-negative.

(A3) *Zero-monotonicity*: For each $S \subset N$, $x \in V(S)$ and $i \notin S$, we have $(x, \omega_i) \in V(S \cup \{i\})$. In particular, this implies that $(\omega_i)_{i\in S} \in \mathbb{R}^S$ belongs to $V(S)$.

(A4) The boundary of $V(N)$, which we denote by $\partial V(N)$, is non-level in the positive orthant (i.e., at any point of $\partial V(N) \cap \mathbb{R}^N_{+}$ there exists an outward vector with positive coordinates.)

(A5) For each $S \subset N$, if $x \in \partial V(S)$ with $x_i < 0$ for $i \in T \subset S$, then $\partial V(S)$ at x is parallel to the subspace $\mathbb{R}^T$.

Properties (A1), (A2), (A3), and (A4) are standard properties. The normalization given in (A2) does not affect our results. Property (A4) has previously been used by Hart and Mas-Colell (1996, in hypothesis (A2), p. 359) and Serrano (1997, in assumption A4, p. 61). The hypothesis in Hart and Mas-Colell (1996) is stronger, since it requires non-levelness in every coalition $S \subset N$. Property (A5), which is such that all relevant action occurs in the positive orthant, generalizes the property given in Serrano's assumption A4.

A *transferable utility (TU) game* on N is a function $v : 2^S \to \mathbb{R}$ that assigns to each coalition S a real number $v(S)$ and $v(\emptyset) = 0$. A TU game v on N may also be expressed as the following NTU game on N:

$$V'(S) = \left\{ x \in \mathbb{R}^S : \sum_{i \in S} x_i \leqslant v(S) \right\} \tag{1}$$

for all $S \subset N$.

Let Π be the set of all the orders of players in N. Given $\pi \in \Pi$ and $i \in N$, we define P_i^π as the set of players who come before player i in the order π, namely

$$P_i^\pi := \{ j \in N : \pi(j) < \pi(i) \}.$$

For the sake of notational convenience, we denote $P_{n+1}^\pi : N$.

Let v be a TU game on N and let $\pi \in \Pi$. Given $i \in N$, we define the *marginal contribution* of player i under the order π in the game v as

$$v(P_i^\pi \cup \{i\}) - v(P_i^\pi) \in \mathbb{R}.$$

The *Shapley value* (Shapley, 1953) of a TU game v on N is the vector $Sh(N, v) \in \mathbb{R}^N$ whose ith coordinate is given by

$$Sh_i(N, v) := \frac{1}{|\Pi|} \sum\nolimits_{\pi \in \Pi} \left[v\left(P_i^\pi \cup \{i\}\right) - v\left(P_i^\pi\right) \right] \in \mathbb{R}.$$

Let $\lambda \in \mathbb{R}^N_{++}$ and let $S \subset N$. We define

$$v^\lambda(S) := \sup \left\{ \sum_{i \in S} \lambda_i x_i : x \in V(S) \right\}.$$

Under our hypothesis, this supremum is a maximum:

Lemma 2.1. *For each $S \subset N$, there exists $x \in V(S)$ such that $\sum_{j \in S} \lambda_j x_j = v^\lambda(S)$.*

Proof. It follows from properties (A1) and (A5). □

A vector $x \in V(N)$ is a *Shapley NTU value* (Shapley, 1969) of V if there exists a vector $\lambda \in \mathbb{R}^N_{++}$ such that $\lambda_i x_i = Sh_i(N, v^\lambda)$ for all $i \in N$. Even though the Shapley NTU value may not be unique, Shapley (1969) pointed out that "it is sufficient [for uniqueness to hold] that the Pareto surface coincide with a hyperplane within the individually rational zone". We will refer to this property later (see Theorem 3.1) as $V(N)$ delimited by a hyperplane. The vector $\lambda \in \mathbb{R}^N_{++}$ is, of course, outward to the hyperplane.

For players negotiating to form a coalition, their payoff allocation will only depend on the identity of the coalition members. Thus, we define a *rule* as a function γ which assigns a vector $\gamma(S) \in V(S)$ to each coalition S. Formally, a rule is a "payoff configuration", as defined for example in Hart and Mas-Colell (1996). However, a rule should not be interpreted as a payoff for every coalition, but as an index that indicates payoff allocations when a particular coalition is formed. We denote the set of all rules as Γ.

3. The non-cooperative mechanism

First, an order of the players is randomly chosen. Assume the order is $\pi = (12\ldots n)$. Player 1 then presents a rule $\gamma \in \Gamma$. No restrictions (apart from feasibility) are imposed on γ. Player 2 may either *agree* with γ and join the coalition, or *disagree* with γ and propose a new rule $\tilde{\gamma}$ to player 1. If player 1 accepts the rule (by *voting 'yes'*), the coalition $\{1,2\}$ forms under the new rule $\tilde{\gamma}$, and the turn passes to player 3. If player 1 rejects the rule (by *voting 'no'*), player 2 is excluded from the coalition and the turn passes to player 3.

In general, when the turn comes round to player i, he is faced with a coalition $S \subset P_i^\pi$ with a specific rule γ, and a set of players $E = P_i^\pi \setminus S$ who have chosen to stay out of the coalition. Players in S, E and $N \setminus P_i^\pi$ are called *active players*, *passive players* and *candidates*, respectively. Player i must then either agree to join the coalition (in which case, player i becomes an active player and the turn passes to candidate $i+1$) or disagree, proposing both a new rule $\tilde{\gamma}$ and a new coalition $\tilde{S} \subset P_i^\pi \cup \{i\}$ which includes himself and all the members of the old coalition (i.e. $S \cup \{i\} \subset \tilde{S}$). The members of $\tilde{S} \setminus \{i\}$ vote sequentially to either accept or reject this proposal. If they all vote 'yes', the new coalition $\tilde{S}$ forms with the new rule (we say then that the proposal has been accepted), and the turn passes to candidate $i+1$. If at least one member of $\tilde{S} \setminus \{i\}$ votes 'no', then player i becomes a passive player and the turn passes to candidate $i+1$.

When no more candidates remain, we have a coalition $S \subset N$ of active players, a set $E = N \backslash S$ of passive players, and a rule γ for the coalition. The final payoff, thus, for each player $i \in S$ is $\gamma_i(S)$ and all the players $i \in E$ receive their individual payoffs ω_i.

Remark 3.1. elspsrem1Note that a player makes a proposal that assigns a payoff allocation to each coalition, even if he does not belong to the coalition. These ''external'' payoff allocations do not play any role, because the implemented coalition will always contain the proposer of the implemented rule. Hence, the results will not be affected if the proposal of any player is a function that assigns payoff allocations only to the coalitions in which he participates.

However, we would like to point out that these rules are not interpreted as abstract functions, but as policies or courses of action that produce different results (depending on the implemented coalition). Thus, any agent can deduce the outcomes that result from a rule for each coalition, including coalitions that do not contain the proposer.

We now describe the mechanism M formally. We first describe the games $M(\pi, i, S, E, \gamma)$ and $\tilde{M}(\pi, i, S, E, \gamma)$. $M(\pi, i, S, E, \gamma)$ is the subgame that begins when, given the order π, the turn comes round to player i and he faces a coalition of active players S with a proposed rule $\gamma \in \Gamma$, as also a set of passive players E, such that $S \cup E = P_i^\pi$ and $S \cap E = \emptyset$. $\tilde{M}(\pi, i, S, E, \gamma)$ is the subgame which arises after player i disagrees in the subgame $M(\pi, i, S, E, \gamma)$.

Let $\pi \in \Pi$ be an order for the players. Without loss of generality, we can assume that $\pi = (12 \ldots n)$. Given $i \in N \cup \{n+1\}$, $\gamma \in \Gamma$ and $S, E \subset P_i^\pi$ such that $S \cup E = P_i^\pi$ and $S \cap E = \emptyset$, we inductively define the mechanisms $M(\pi, i, S, E, \gamma)$ and $\tilde{M}(\pi, i, S, E, \gamma)$ as described immediately below.

In both $M(\pi, n+1, S, E, \gamma)$ and $\tilde{M}(\pi, n+1, S, E, \gamma)$, each player $i \in S$ receives $\gamma_i(S)$ and each player $i \in E$ receives ω_i.
Assume that both $M(\pi, j, S', E', \gamma')$ and $\tilde{M}(\pi, j, S', E', \gamma')$ are defined for all $j > i$, $\gamma' \in \Gamma$ and S', E' such that $S' \cup E' = P_j^\pi$ and $S' \cap E' = \emptyset$.
In $\tilde{M}(\pi, i, S, E, \gamma)$, player i proposes a rule $\tilde{\gamma} \in \Gamma$ and sets $\tilde{S} \supset S$ and $\tilde{E} \subset E$ such that $i \in \tilde{S}$, $\tilde{S} \cup \tilde{E} = P_i^\pi \cup \{i\}$ and $\tilde{S} \cap \tilde{E} = \emptyset$. If all the members of $\tilde{S} \setminus \{i\}$ vote 'yes' (they are asked in some prespecified order), then the mechanism $M(\pi, i+1, \tilde{S}, \tilde{E}, \tilde{\gamma})$ is played. If at least one member of $\tilde{S} \setminus \{i\}$ votes 'no', then the mechanism $M(\pi, i+1, S, E \cup \{i\}, \gamma)$ is played.
In $M(\pi, i, S, E, \gamma)$, player i can either agree or disagree with (S, E, γ). If the latter, then $\tilde{M}(\pi, i, S, E, \gamma)$ is played. If the former, then $M(\pi, i+1, S \cup \{i\}, E, \gamma)$ is played.
The mechanism M consists of randomly choosing an order $\pi' \in \Pi$ with each order equally likely to be chosen and playing the game $M(\pi') := \tilde{M}(\pi', i', \emptyset, \emptyset, \gamma^0)$, where $\pi'(i') = 1$.

Clearly, for any set of pure (mixed) strategies, this mechanism terminates in finite time. Thus, the (expected) payoffs at termination are well-defined. We will also assume that, if a player feels indifferent to a proposal, then he prefers to agree. This assumption is made in order to avoid problems of coordination among players. In Section 4.2 we show that this tie-breaking rule is necessary for our model. Note that we do not need to make any assumption when players sequentially vote 'yes' or 'no' to a proposal in the subgames $\tilde{M}$.

Theorem 3.1. *If $V(N)$ is bounded by a hyperplane, then there exists a unique expected subgame perfect Nash equilibrium (SPNE) payoff in the negotiation mechanism M, that is the Shapley NTU value. Furthermore, the strategy of a player in SPNE in the negotiation mechanism $M(\pi)$ for any π is robust to deviations by coalitions of his predecessors in π.*

The proof of Theorem 3.1 is provided in the Appendix. The main idea of the proof is that, given any order π, each player has a strategy that ensures his marginal contribution in the order π. Thus, in expected terms, the final payoff is the Shapley NTU value.

Remark 3.2. Following a similar argument as in the proof of Theorem 3.1, it is straightforward to prove that the negotiation mechanism M implements the Shapley value in zero-monotonic TU games.

Notice that this result is not a consequence of Theorem 3.1, because TU games do not satisfy all the assumptions that define NTU games.

4. Some examples

4.1. A classical example

In this section we apply the above procedure to an exchange economy that was described in a series of papers in Econometrica in the 1980s that discussed the applicability of the Shapley NTU value. The reader is referred to Roth (1980), Shafer (1980), Harsanyi (1980), Aumann (1985b), Roth (1986), and Aumann (1986). This controversy has been recently revisited in Montero and Okada (2007).

Example 4.1 (*cf.* *Shafer*, 1980). Consider a pure exchange economy with three players $\{1,2,3\}$ and two commodities $\{x,y\}$. Initial endowments are given by:

$$z_1^0 = (1-\varepsilon, 0),$$
$$z_2^0 = (0, 1-\varepsilon),$$
$$z_3^0 = (\varepsilon, \varepsilon)$$

and utility functions are given by

$$u_1(x,y) = \min\{x,y\},$$
$$u_2(x,y) = \min\{x,y\},$$
$$u_3(x,y) = \frac{x+y}{2}.$$

Commodities x and y may be considered as 'left gloves' and 'right gloves', respectively. Players 1 and 2 only get utility from pairs of gloves. However, player 3 only uses the leather of the gloves. Let i denote an element in $\{1,2\}$. The non-transferable utility (NTU) game is thus given by

$$V(\{i\}) = \{t \in \mathbb{R}^{\{i\}} : t \leqslant 0\},$$
$$V(\{3\}) = \{t \in \mathbb{R}^{\{3\}} : t \leqslant \varepsilon\},$$
$$V(\{1,2\}) = \{(t_1,t_2) \in \mathbb{R}^{\{1,2\}} : t_1 + t_2 \leqslant 1-\varepsilon, t_1 \leqslant 1-\varepsilon, t_2 \leqslant 1-\varepsilon\},$$
$$V(\{i,3\}) = \left\{(t_i,t_3) \in \mathbb{R}^{\{i,3\}} : t_i + t_3 \leqslant \frac{1+\varepsilon}{2}, t_i \leqslant \varepsilon, t_3 \leqslant \frac{1+\varepsilon}{2}\right\}$$
$$V(\{1,2,3\}) = \{(t_1,t_2,t_3) \in \mathbb{R}^{\{1,2,3\}} : t_1 + t_2 + t_3 \leqslant 1, t_1 \leqslant 1, t_2 \leqslant 1, t_3 \leqslant 1\}.$$

The Shapley NTU value proposes a payoff of $(\frac{5(1-\varepsilon)}{12}, \frac{5(1-\varepsilon)}{12}, \frac{5\varepsilon+1}{6})$. For a discussion of this result, the reader is referred to Shafer (1980), Roth (1980) and Aumann (1985b).

The TU game associated with this example is given by $\lambda = (1,1,1)$ and $v^\lambda(\{i\}) = 0$, $v^\lambda(\{3\}) = \varepsilon, v^\lambda(\{1,2\}) = 1-\varepsilon$, $v^\lambda(\{i,3\}) = (1+\varepsilon)/2$, and $v(N) = 1$ (see Fig. 1). In the order $\pi = (312)$, a possible SPNE in the bargaining mechanism would be obtained as follows: players 3 and 1 propose a rule γ that satisfies $\gamma(N) = ((1-\varepsilon)/2, (1-\varepsilon)/2, \varepsilon)$ – i.e. the vector of marginal contributions in the order π – and $\gamma(\{1,3\}) = ((1-\varepsilon)/2, -, \varepsilon)$. Player 2 cannot hope to suggest a more profitable outcome for himself. In fact, players 1 and 3 are indifferent to whether or not player 2 joins them. Player 2 accepts the offer and the final payoff is $\gamma(N)$.

Assume now we are in the NTU game of the example. Players 3 and 1 cannot propose a rule satisfying $\gamma(\{1,3\}) = ((1-\varepsilon)/2, -, \varepsilon)$, because this payoff allocation is not feasible for them. This is equivalent to their wanting to make a non-credible threat to player 2 should he not want to join them.

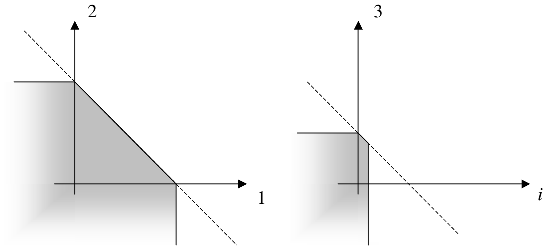

Fig. 1. Feasible outcomes for $\{1,2\}$ and $\{i,3\}$.

However, they can still propose $\gamma(N) = ((1-\varepsilon)/2, (1-\varepsilon)/2, \varepsilon)$ and $\gamma(\{1,3\}) = (0,(1+\varepsilon)/2)$. This means that

- they propose a coalition N in which player 1 receives commodities (x_1, y_1) with $x_1 = y_1 = (1-\varepsilon)/2$; player 2 receives commodities (x_2, y_2) with $x_2 = y_2 = (1-\varepsilon)/2$, and player 3 keeps his initial endowment; and
- should player 2 not join them, they can threaten to form a coalition in $\{1,3\}$ in which player 3 would receive all their commodities (i.e. (x_3, y_3) with $x_3 = 1$ and $y_3 = \varepsilon$), and in which player 1 would receive nothing.

In this case, the threat is credible, because this allocation is feasible for $\{1,3\}$. Again, player 2 cannot suggest a more profitable outcome for himself. Any feasible proposal giving him more than $(1-\varepsilon)/2$ would be rejected by player 1 or 3. This means that, in equilibrium, player 2 would immediately agree to join the coalition. Note that, in this case, players 1 and 3are not indifferent to player 2 joining them.

4.2. The tie-breaking rule

Assume the tie-breaking rule does not hold. Hence, the Shapley NTU value is still an SPNE outcome. However, other SPNE payoff allocations may arise, as the next example shows.

Example 4.2. Let $N = \{1,2,3,4\}$ and let v be defined by $v(S) = |S| + 1$ if $\{1,2,3\} \subset S$ and $v(S) = |S|$ otherwise. Let $\pi = (1234)$. In this example, the vector of marginal contributions in the order π is $d^\pi = (1,1,2,1)$. We consider the following strategies for players in the order π: Players $\{1,2\}$ propose a rule γ satisfying $\gamma(\{1,2\}) = (2,0)$, $\gamma(\{1,2,4\}) = (0,2,1)$ and $\gamma(N) = (1,2,1,1)$. Player 4, when faced with a coalition $S = \{1,2\}$ and a set of passive players $E = \{3\}$ with a rule γ such that $\gamma_4(\{1,2,4\}) = 1$, would apply the following tie-breaking rule: if player 3 is excluded after proposing $\tilde{\gamma}$ with $\tilde{\gamma}_1(N) \geqslant \tilde{\gamma}_2(N)$, then player 4 will agree to join the coalition.[2] If player 3 is excluded after proposing $\tilde{\gamma}$ with $\tilde{\gamma}_1(N) < \tilde{\gamma}_2(N)$, then player 4 will disagree and propose an unacceptable offer, for example $(1,1,1,2)$.

These strategies can be supported as part of a SPNE. Player 3 agrees because he does not expect to propose a positive payoff for himself. If he disagreed and proposed $\tilde{\gamma}$ with $\tilde{\gamma}_1(N) \geqslant \tilde{\gamma}_2(N)$, then player 2 would get 2 by voting 'no'. This means that $\tilde{\gamma}$ will not be accepted unless $\tilde{\gamma}_1(N) \geqslant \tilde{\gamma}_2(N) \geqslant 2$, which would leave player 3 with a negative payoff. If player 3 disagreed and proposed $\tilde{\gamma}$ with $\tilde{\gamma}_1(N) < \tilde{\gamma}_2(N)$, then player 1 would get 1 by voting 'no'. Again, this means that $\tilde{\gamma}$ will not be accepted unless $\tilde{\gamma}_2(N) > \tilde{\gamma}_1(N) \geqslant 2$.

Hence, the final payoff allocation is $(1,2,1,1)$ irrespective of player 4's action.

Vidal-Puga and Bergantiños (2003) model this tie-breaking rule by punishing the players involved in an exclusion with a small penalty $\varepsilon > 0$. Applied to this model, this would mean that any excluded player i would get a utility of almost (but strictly less than) ω_i. The result stated in Theorem 3.1 would also hold (without the tie-breaking rule) if we restricted ourselves to strict zero-monotonic games, i.e. for each $S \subset N$, $x \in V(S)$ and $i \notin S$, the payoff allocation (x, ω_i) belongs to the interior of $V(S \cup \{i\})$.

[2] For simplicity, we assume that player 3 makes an acceptable offer to player 4 (i.e. $\tilde{\gamma}_4(N) = 1$). A more precise description of player 4's strategy is given in Section A.2.

5. Concluding remarks

In this paper a new sequential protocol M of coalition formation is presented. The order in which the players join the coalition is critical for the final payoff allocation in SPNE. When the corresponding NTU game (N, V) satisfies a flat $V(N)$, this allocation is the corresponding vector of marginal contributions. The key factor is the use of threats in the form commitments: Prior to the accession of new members, the coalition in formation should have a internal rule that assigns a different payoff depending on who are its final members.

When the turn reaches a particular player, the protocol determines that he should be able to know his followers' order in advance. This is a key feature, since in SPNE, the accepted rule typically assigns to the grand coalition the corresponding vector of marginal contributions (which depend on the order of the players).

The expected SPNE payoff allocation anticipated before playing any realized game is the Shapley NTU value. Following Bag and Winter (1999) and Mutuswami and Winter (2002), it is possible to add one additional stage to get the value precisely and not in expectation. The detailed protocol M' would be as follows: In the first stage, players play the mechanism M (according to any arbitrary order of the players). At the end of this stage, each player is asked whether he wants to replay M. The mechanism ends if all players answer "NO". Otherwise, the mechanism M is replayed according to a randomly chosen order. The mechanism M' ends after this optional second stage.

Assuming that all the players have a lexicographic preference for the game ending in the first stage, the only final SPNE payoff allocation is the Shapley NTU value. See Bag and Winter (1999, p. 79) for the intuition behind this result.

The application of the Nash program to the Shapley NTU value is still an open question for general NTU games. The expected utility of SPNE allocations over the different orders will usually be inefficient if the feasible set of the grand coalition is strictly convex. It would be of interest to find a non-cooperative mechanism supporting the Shapley NTU value in the most general case. The present paper could be seen as a first step in this direction.

Appendix A

A.1. Proof of Theorem

The proof is structured as follows: first of all, some additional notation is introduced; Secondly, a SPNE that yields the Shapley NTU value is constructed; finally, it is shown that any SPNE yields the Shapley NTU value as the expected final payoff.

A.1.1. Additional notation

Let $(\lambda_i)_{i\in N} \in \mathbb{R}^N_{++}$ and $k \in \mathbb{R}_+$ be such that

$$V(N) = \{x \in \mathbb{R}^N_+ : \sum_{i\in N} \lambda_i x_i = k\} - \mathbb{R}^N_+.$$

Clearly, $v^\lambda(N) = k$.

In order to prove Theorem 3.1, we need some additional notation. Given $x \in \mathbb{R}^S$, we define $x^+ \in \mathbb{R}^S_+$ as the vector whose coordinates are given by $x_i^+ := \max\{0, x_i\}$ for all $i \in S$. By (A5), $x \in V(S)$ implies $x^+ \in V(S)$ for all $S \subset N$.

Let $\pi \in \Pi$. From now on, we assume without loss of generality that $\pi = (12\ldots n)$. In particular, this implies that $P^\pi_{i+1} = P^\pi_i \cup \{i\}$ for all $i \in N$.

Let $\lambda_i d^\pi_i$ be the marginal contribution of player i to the game v^λ in the order π, namely

$$d^\pi_i := \frac{1}{\lambda_i}[v^\lambda(P^\pi_{i+1}) - v^\lambda(P^\pi_i)].$$

Given $x \in \mathbb{R}^{P^\pi_i}$, we define

$$f^\pi_i(x) := \max\left\{y_i : (x, y_i, d^\pi_{N\setminus P^\pi_{i+1}}) \in V(N)\right\}$$

when this maximum exists. In particular, if $x \in V(P^\pi_i)$, this value is well-defined and non-negative.

Note that $f_i^\pi(x)$ represents the maximum payoff that player i can obtain when the players preceding him obtain x and the players following him obtain d^π.

Let $x \in \mathbb{R}^{P_i^\pi}$ be such that $f_i^\pi(x)$ is well-defined. It is straightforward to confirm that

$$f_i^\pi(x) = \frac{1}{\lambda_i}[v^\lambda(P_{i+1}^\pi) - \sum_{j<i} \lambda_j x_j^+]. \tag{2}$$

Given $S \subsetneq N$, we define

$$\kappa(S) := \min\{i \in N \setminus S : P_i^\pi \subset S\}.$$

Thus, $\kappa(S)$ is the excluded player with the lowest index, which means that players in $P_{\kappa(S)}^\pi$ are the first players out of S who come together in the order π. This minimum always exists, because $P_1^\pi = \emptyset$. Moreover, $P_{\kappa(S)}^\pi \subset S$. We also define

$$\Gamma^\pi := \left\{\gamma \in \Gamma : \gamma(S) = (\gamma^+(P_{\kappa(S)}^\pi), \omega_{S \setminus P_{\kappa(S)}^\pi}) \text{ for all} S \subsetneq N\right\}$$

where $\gamma^+(S)$ is such that $\gamma_i^+(S) = \max\{0, \gamma_i(S)\}$ for all $i \in S$.

Thus, Γ^π is the set of (positive) rules which do not share out the resources of the players after the first 'gap' in the coalition (with respect to π). Note that, given $\gamma \in \Gamma^\pi$, we can change $\gamma(N)$ and the resulting rule will still be in Γ^π.

We also define

$$K^\pi := \left\{\gamma \in \Gamma^\pi : \gamma(P_i^\pi) \in \arg\min_{x \in V(P_i^\pi)} \{f_i^\pi(x)\} \text{ for all } i \in N\right\}.$$

This K^π is the set of rules out of Γ^π which give each coalition P_i^π the payoff x that minimizes $f_i^\pi(x)$.

Should player i disagree in equilibrium, his payoff would be $f_i^\pi(x)$ with $x \in \mathbb{R}^{P_i^\pi}$ as the payoff allocation that his predecessors would obtain should they vote 'no' to his proposal. In SPNE, players who precede player i would try to minimize $f_i^\pi(x)$ so that they could extract the maximum surplus from player i. Hence, K^π is a possible set of rules from among which these predecessors could choose their proposals.

It is straightforward to confirm that K^π can also be expressed as

$$K^\pi = \left\{\gamma \in \Gamma^\pi : \sum_{j<i} \lambda_j \gamma_j(P_i^\pi) = v^\lambda(P_i^\pi) \text{ for all } i \in N\right\}.$$

We denote by Θ_i^π the set of feasible (S, E, γ)'s in the subgame $M(i, S, E, \gamma)$, namely

$$\Theta_i^\pi := \{(S, E, \gamma) : S \cup E = P_i^\pi, \ S \cap E = \emptyset \text{ and } \gamma \in \Gamma\}.$$

A.1.2. Existence of SPNE

Given any $(S, E, \gamma) \in \Theta_{n+1}^\pi$, we define

$$b(n+1, S, E, \gamma) := (\gamma(S), \omega_E) \in V(N).$$

Thus, $b(n+1, S, E, \gamma)$ is the final payoff in the fictitious subgame $M(\pi, n+1, S, E, \gamma)$.

Consider the following strategies in the subgames $M(\pi, n, S, E, \gamma)$ and $\widetilde{M}(\pi, n, S, E, \gamma)$:

In the subgame $M(\pi, n, S, E, \gamma)$, player n agrees to (S, E, γ) if and only if

$$\gamma_n(S \cup \{n\}) \geqslant f_n^\pi(\gamma(S), \omega_E),$$

which can be restated as

$$b_n(n+1, S \cup \{n\}, E, \gamma) \geqslant f_n^\pi(b_{P_n^\pi}(n+1, S, E \cup \{n\}, \gamma)). \tag{3}$$

In the subgame $\tilde{M}(\pi, n, S, E, \gamma)$, player n proposes $(\tilde{S}, \tilde{E}, \tilde{\gamma})$ such that $\tilde{S} = N$, $\tilde{E} = \emptyset$ and

$$\tilde{\gamma}(N) = (t, f_n^{\pi}(t)) \tag{4}$$

with $t := (\gamma(S), \omega_E)$ and $\tilde{\gamma}(T)$ for all $T \neq N$ given by

$$\tilde{\gamma}(T) = (y, \omega_{T \setminus P^{\pi}_{\kappa(T)}}) \tag{5}$$

and with $y \in V(P^{\pi}_{\kappa(T)})$ such that

$$\sum_{j<\kappa(T)} \lambda_j y_j = v^{\lambda}(P^{\pi}_{\kappa(T)}). \tag{6}$$

Clearly, $\tilde{\gamma} \in K^{\pi}$. In the subgame $\tilde{M}(\pi, n, S, E, \gamma)$, assume player n proposes $(\tilde{S}, \tilde{E}, \tilde{\gamma})$ and $j \in \tilde{S} \setminus \{n\}$. Hence, player j votes 'yes' if and only if

$$b_j(n+1, \tilde{S}, \tilde{E}, \tilde{\gamma}) \geqslant b_j(n+1, S, E \cup \{n\}, \gamma). \tag{7}$$

We have thus defined the strategies of the players in $M(\pi, n, S, E, \gamma)$ for any $(S, E, \gamma) \in \Theta^{\pi}_n$. Assume that for each $j > i$ and each $(S, E, \gamma) \in \Theta^{\pi}_j$, we have defined the strategy profiles in $M(\pi, j, S, E, \gamma)$ and $\tilde{M}(\pi, j, S, E, \gamma)$. Let $b(j, S, E, \gamma) \in V(N)$ be the final payoff allocation when players follow these strategies in $M(\pi, j, S, E, \gamma)$.

We now describe the strategies in $M(\pi, i, S, E, \gamma)$ and $\tilde{M}(\pi, i, S, E, \gamma)$. In $M(\pi, i, S, E, \gamma)$, player i agrees to (S, E, γ) if and only if

$$b_i(i+1, S \cup \{i\}, E, \gamma) \geqslant f_i^{\pi}(b_{P_i^{\pi}}(i+1, S, E \cup \{i\}, \gamma)).$$

The strategies applied in $\tilde{M}(\pi, i, S, E, \gamma)$ are as follows: assume we are in the subgame $\tilde{M}(\pi, i, S, E, \gamma)$. Player i proposes $(P^{\pi}_{i+1}, \emptyset, \tilde{\gamma})$ with $\tilde{\gamma}$ given by

$$\tilde{\gamma}(N) = (t, f_i^{\pi}(t), d^{\pi}_{N \setminus P^{\pi}_{i+1}})$$

with $t = b_{P_i^{\pi}}(i+1, S, E \cup \{i\}, \gamma)$. It is not difficult to confirm that $\tilde{\gamma}(N)$ is well-defined (i.e. $\tilde{\gamma}(N) \in V(N)$). For $S \subsetneq N$, $\tilde{\gamma}(S)$ is given as in (5) and (6). Hence, $\tilde{\gamma} \in K^{\pi}$.

In the subgame $\tilde{M}(\pi, i, S, E, \gamma)$, assume that player i proposes $(\tilde{S}, \tilde{E}, \tilde{\gamma}) \in \Theta^{\pi}_{i+1}$ and $j \in \tilde{S} \setminus \{i\}$. In this case, player j votes 'yes' if and only if

$$b_j(i+1, \tilde{S}, \tilde{E}, \tilde{\gamma}) \geqslant b_j(i+1, S, E \cup \{i\}, \gamma).$$

Under these strategies, therefore, player 1 proposes $(\{1\}, \emptyset, \gamma^{\pi})$ with $\gamma^{\pi}(N) = d^{\pi}$ and the rest of players agree. The coalition is thus formed with all the players and the final outcome is d^{π}.

It is straightforward to prove that these strategies form a SPNE.

A.1.3. Uniqueness of SPNE payoff allocations

We will now prove that every SPNE in $M(\pi)$ has d^{π} as the final outcome. Assume that we are in a SPNE of $M(\pi)$. We will prove that the final payoff allocation in $M(\pi)$ is d^{π}.

Given $i \in N$ and $(S, E, \gamma) \in \Theta^{\pi}_i$, let $b(i, S, E, \gamma) \in V(N)$ be the SPNE payoff allocation in the subgame $M(\pi, i, S, E, \gamma)$.

We now proceed by a series of claims, as follows: Claim A states that passive players do not receive anything; Claim B describes a condition that is sufficient for candidates to agree; Claim C specifies the final payoff allocation when a candidate disagrees and makes a new proposal; Claim E says that each candidate i receives at least d^{π}_i; finally, Claims D and F are technical claims.

Claim A. $b_E(i, S, E, \gamma) = \omega_E$ *for all* $i \in N$, $(S, E, \gamma) \in \Theta^{\pi}_i$.

Claim B. *Assuming we are in* $M(\pi, i, P^{\pi}_i, \emptyset, \gamma)$ *such that* $\gamma_j(N) \geqslant d^{\pi}_j$ *for all* $j \geqslant i$ *and* $\gamma \in K^{\pi}$, *player i agrees.*

Claim C. *Assuming we are in* $\tilde{M}(\pi, i, S, E, \gamma)$, *the final payoff allocation is given by* $(t, f_i^{\pi}(t), d^{\pi}_{N \setminus P^{\pi}_{i+1}})$ *with* $t = b_{P_i^{\pi}}(i+1, S, E \cup \{i\}, \gamma)$.

Claim D. *Assuming we are in* $M(\pi, i, S, E, \gamma)$, *there exists* $T \supset S$, $T \cap E = \emptyset$ *such that* $b_S(i, S, E, \gamma) = \gamma_S(T)$.

Claim E. $b_j(i,S,E,\gamma) \geqslant d_j^\pi$ *for all* $j \geqslant i$, $(S,E,\gamma) \in \Theta_i^\pi$.

Claim F. $\sum_{j\in S}\lambda_j b_j(i,S,E,\gamma) \leqslant v^\lambda(S)$ *for all* $i \in N$, $(S,E,\gamma) \in \Theta_i^\pi$.

Since under Claim E, each player $i \in N$ can ensure a final payoff for himself of at least d_i^π, and d^π, moreover, is an efficient payoff allocation, then we can conclude that the only possible final payoff allocation in SPNE for the subgame $M(\pi)$ is d^π and, furthermore, that the strategy of player i is robust to deviations by coalitions of P_i^π.

We can prove these claims by backward induction on i. We consider a fictitious case for $i = n+1$, where the subgames $M(\pi,n+1,S,E,\gamma)$ are trivial subgames in which players receive $b(n+1,S,E,\gamma) = (\gamma(S),\omega_E)$. Hence, the claims for this case are trivial.

We will now prove the claims for $i \leqslant n$.

Proof of Claim A. Assume we are in the subgame $M(\pi,i,S,E,\gamma)$. If player i agrees (or he disagrees and his new proposal is, furthermore, rejected), under the induction hypothesis, players in E get ω_E. Assume, on the other hand, that player i disagrees and makes an acceptable proposal. It is well-known that, in equilibrium, player i would make a proposal that would leave the responders indifferent to voting 'yes' or 'no'. By induction hypothesis, any responder $j \in E$ receives ω_j if he votes 'no'. We can thus conclude the result. □

Proof of Claim B. The induction hypothesis for Claim B holds for $i+1,\ldots,n$ if player i agrees. Thus, by induction hypothesis applied to Claim B, we know that player i gets a payoff of $\gamma_i(N)$ by agreeing. Assume player i disagrees and proposes a different $\tilde{\gamma}$. If this proposal is rejected, under Claim A, player i receives $\omega_i \leqslant d_i^\pi$ and thus is either strictly worse off (when $\omega_i < d_i^\pi$) or does not apply the tie-breaking rule (when $\omega_i = d_i^\pi$). If the proposal is accepted by players in P_i^π, this means that each will receive at least what they would obtain by rejecting the proposal. Under Claim D, this is $\gamma_{P_i^\pi}(T)$ for some $T \supset P_i^\pi$ with $i \notin T$. Since $\gamma \in \Gamma^\pi$ and $i \notin T$, we know that $\gamma_{P_i^\pi}(T) = \gamma(P_i^\pi)$. Under Claim E, the final payoff for player i is no more than $f_i^\pi(\gamma(P_i^\pi))$. But $\gamma \in K^\pi$, which means that $\sum_{j<i}\lambda_j\gamma_j(P_i^\pi) = v^\lambda(P_i^\pi)$, and so

$$f_i^\pi(\gamma(P_i^\pi)) = \frac{1}{\lambda_i}\left[v^\lambda\left(P_{i+1}^\pi\right) - \sum_{j<i}\lambda_j\gamma_j\left(P_i^\pi\right)\right] = \frac{1}{\lambda_i}\left[v^\lambda\left(P_{i+1}^\pi\right) - v^\lambda\left(P_i^\pi\right)\right] = d_i^\pi.$$

Here again, player i is either strictly worse off or does not apply the tie-breaking rule. □

Proof of Claim C. Let $t = b_{P_i^\pi}(i+1,S,E\cup\{i\},\gamma)$, and consider the following strategy for player i: given $\varepsilon > 0$, player i proposes $(P_{i+1}^\pi,\emptyset,\tilde{\gamma})$ such that $\tilde{\gamma} \in K^\pi$ and $\tilde{\gamma}(N) = (t, f_i^\pi(t), d_{N\setminus P_{i+1}^\pi}^\pi) + x^\varepsilon$, where $\lambda_j x_j^\varepsilon = \varepsilon$ for all $j \neq i$, and $\lambda_i x_i^\varepsilon = -(n-1)\varepsilon$. Under Claim B, we know that this proposal is bound to be accepted should players in P_i^π vote 'yes'. So, by voting 'yes', players in P_i^π obtain more than what they would obtain by voting 'no'. So they vote 'yes' and player i gets a final payoff of $f_i^\pi(t) - \frac{n-1}{\lambda_i}\varepsilon$. Since this is true for any $\varepsilon > 0$, we can conclude that player i can obtain at least $f_i^\pi(t)$ in SPNE. Moreover, player i cannot obtain more by making an acceptable offer, since any player $j \in P_i^\pi$ can ensure a payoff of t_j for himself by rejecting any new offer. □

Proof of Claim D. If player i agrees, under the induction hypothesis, $b_{S\cup\{i\}} = \gamma_{S\cup\{i\}}(T)$ with $T \supset S\cup\{i\}$ (thus $T \supset S$) and $T \cap E = \emptyset$. Hence, $b_S(i,S,E,\gamma) = \gamma_S(T)$. If player i disagrees, under Claim C, $b_S(i,S,E,\gamma) = b_S(i+1,S,E\cup\{i\},\gamma)$. Under the induction hypothesis, $b_S(i+1,S,E\cup\{i\},\gamma) = \gamma_S(T)$ with $T \supset S$ and $T \cap (E\cup\{i\}) = \emptyset$ (thus $T \cap E = \emptyset$). Hence, $b_S(i,S,E,\gamma) = \gamma_S(T)$. □

Proof of Claim E. By induction hypothesis, the result is true for $j > i$. Let $\dot{b} = b(i+1,S,E\cup\{i\},\gamma)$. Under Claim C, player i can, by rejecting the proposal, ensure a payoff for himself of

$$f_i^\pi\left(\dot{b}_{P_i^\pi}\right) = \frac{1}{\lambda_i}\left[v^\lambda\left(P_{i+1}^\pi\right) - \sum_{j<i}\lambda_j\dot{b}_j\right]$$

under Claim A,

$$= \frac{1}{\lambda_i}\left[v^\lambda\left(P^\pi_{i+1}\right) - \sum_{j\in S}\lambda_j \dot{b}_j - \sum_{j\in E}\lambda_j\omega_j\right]$$

under Claim F,

$$\geqslant \frac{1}{\lambda_i}\left[v^\lambda\left(P^\pi_{i+1}\right) - v^\lambda(S) - \sum_{j\in E}\lambda_j\omega_j\right]$$

according to zero-monotonicity, $v^\lambda(S) + \sum_{j\in E}\lambda_j\omega_j \leqslant v^\lambda(P^\pi_i)$, and so

$$\geqslant \frac{1}{\lambda_i}\left[v^\lambda\left(P^\pi_{i+1}\right) - v^\lambda\left(P^\pi_i\right)\right] = d^\pi_i. \qquad \square$$

Proof of Claim F. Assume we are in $M(\pi, i, S, E, \gamma)$. For simplicity sake, we denote $b = b(i, S, E, \gamma)$ and $\dot{b} = b(i+1, S, E \cup \{i\}, \gamma)$. If player i disagrees, under Claim C, players in P^π_i get $b_{P^\pi_i} = \dot{b}_{P^\pi_i}$. Under the induction hypothesis,

$$\sum_{j\in S}\lambda_j b_j = \sum_{j\in S}\lambda_j \dot{b}_j \leqslant v^\lambda(S).$$

Let $\ddot{b} = b(i+1, S\cup\{i\}, E, \gamma)$. If player i agrees, then $b = \ddot{b}$. Moreover, under Claim C, player i would not agree if $\ddot{b}_i < f^\pi_i(\dot{b}_{P^\pi_i})$. Thus,

$$b_i = \ddot{b}_i \geqslant f^\pi_i\left(\dot{b}_{P^\pi_i}\right) = \frac{1}{\lambda_i}\left[v^\lambda(P^\pi_{i+1}) - \sum_{j<i}\lambda_j \dot{b}_j\right]$$

under Claim A,

$$= \frac{1}{\lambda_i}\left[v^\lambda(P^\pi_{i+1}) - \sum_{j\in S}\lambda_j \dot{b}_j - \sum_{j\in E}\lambda_j\omega_j\right]. \tag{8}$$

So, under Claims A and E,

$$\begin{aligned}\sum_{j\in S}\lambda_j b_j &= \sum_{j\in N}\lambda_j b_j - \sum_{j\in E}\lambda_j b_j - \lambda_i b_i - \sum_{j>i}\lambda_j b_j \leqslant v^\lambda(N) - \sum_{j\in E}\lambda_j\omega_j - \lambda_i b_i - \sum_{j>i}\lambda_j d^\pi_j\\ &= v^\lambda\left(P^\pi_{i+1}\right) - \sum_{j\in E}\lambda_j\omega_j - \lambda_i b_i\end{aligned}$$

under (8),

$$\leqslant v^\lambda(P^\pi_{i+1}) - \sum_{j\in E}\lambda_j\omega_j - \left[v^\lambda\left(P^\pi_{i+1}\right) - \sum_{j\in S}\lambda_j \dot{b}_j - \sum_{j\in E}\lambda_j\omega_j\right] = \sum_{j\in S}\lambda_j \dot{b}_j$$

under the induction hypothesis,

$$\leqslant v^\lambda(S). \qquad \square$$

A.2. The tie-breaking rule

A more precise description for the strategy of player 4 in Example 4.2 is the following: If player 3 is excluded after proposing $\tilde{\gamma}$ with $\tilde{\gamma}_1(N) \geqslant \tilde{\gamma}_2(N)$ and $\tilde{\gamma}_4(N) \geqslant 1$, or $\tilde{\gamma}_1(\{1, 2, 3\}) \geqslant \tilde{\gamma}_2(\{1, 2, 3\})$ and $\tilde{\gamma}_4(N) < 1$, then player *4* will agree to join the coalition. If player 3 is excluded after proposing $\tilde{\gamma}$ with $\tilde{\gamma}_1(N) < \tilde{\gamma}_2(N)$ and $\tilde{\gamma}_4(N) \geqslant 1$, or $\tilde{\gamma}_1(\{1,2,3\}) < \tilde{\gamma}_2(\{1,2,3\})$ and $\tilde{\gamma}_4(N) < 1$, then player 4 will disagree and propose an unacceptable offer, such like $(1,1,1,2)$.

References

Aumann, R., 1985a. An axiomatization of the non-transferable utility value. Econometrica 53, 599–612.

Aumann, R., 1985b. On the non-transferable utility value: A comment on the Roth–Shafer examples. Econometrica 53, 667–677.

Aumann, R., 1986. Rejoinder. Econometrica 54, 985–989.

Bag, P.K., Winter, E., 1999. Simple subscription mechanisms for excludable public goods. Journal of Economic Theory 87, 72–94.

Bloch, F., 1996. Sequential formation of coalitions in games with externalities and fixed payoff division. Games and Economic Behavior 14, 90–123.

Chae, S., Yang, J.-A., 1994. An *N*-person pure bargaining game. Journal of Economic Theory 62 (1), 86–102.

Chatterjee, K., Dutta, B., Ray, D., Sengupta, K., 1993. A noncooperative theory of coalitional bargaining. Review of Economic Studies 60, 463–477.

Gul, F., 1989. Bargaining foundations of the Shapley value. Econometrica 57, 81–95.

Harsanyi, J.C., 1963. A simplified bargaining model for the *n*-person cooperative game. International Economic Review 4, 194–220.

Harsanyi, J.C., 1980. Comments on Roth's paper, 'Values for games without side payments'. Econometrica 48, 477.

Hart, S., 1994. On prize games. In: Meggido, N. (Ed.), Essays in Game Theory. Springer, Berlin, pp. 111–121.

Hart, S., 2004. A comparison of non-transferable utility values. Theory and Decision 56, 35–46.

Hart, S., Kurz, M., 1983. Endogenous formation of coalitions. Econometrica 51 (4), 1047–1064.

Hart, S., Mas-Colell, A., 1996. Bargaining and value. Econometrica 64, 357–380.

Hart, O., Moore, J., 1990. Property rights and the nature of the firm. Journal of Political Economy 98, 1119–1158.

Macho-Stadler, Inés, Pérez-Castrillo, D., Porteiro, N., 2006. Sequential formation of coalitions through bilateral agreements in a Cournot setting. International Journal of Game Theory 34 (2), 207–228.

Maschler, M., Owen, G., 1989. The consistent Shapley value for hyperplane games. International Journal of Game Theory 18, 389–407.

Maschler, M., Owen, G., 1992. The consistent Shapley value for games without side payments. In: Selten, R. (Ed.), Rational Interaction. Springer, New York, pp. 5–12.

Maskin, E., 2003. Bargaining, coalitions, and externalities. Presidential address to the Econometric Society.

Montero, M., Okada, A., 2007. Riskly versus riskless bargaining procedures: The Aumnann–Roth controversy revisited. Japanese Economic Review 58 (2), 294–302.

Mutuswami, S., Winter, E., 2002. Subscription mechanisms for network formation. Journal of Economic Theory 106, 242–264.

Naeve, J., 1999. Nash implementation of the Nash bargaining solution using intuitive message spaces. Economic Letters 62, 23–28.

Naeve-Steinweg, E., 1999. A note on van Damme's mechanism. Review of Economic Design 4, 179–187.

Nash, J.F., 1950. The bargaining problem. Econometrica 18, 155–162.

Okada, A., 1996. A noncooperative coalitional bargaining game with random proposers. Games and Economic Behavior 16, 97–108.

Pérez-Castrillo, D., Wettstein, D., 2001. Bidding for the surplus: A non-cooperative approach to the Shapley value. Journal of Economic Theory 100, 274–294.

Ray, D., Vohra, R., 1999. A theory of endogenous coalition structures. Games and Economic Behavior 26, 286–336.

Roth, A.E., 1980. Values of games without side payments: Some difficulties with current concepts. Econometrica 48, 457–465.

Roth, A.E., 1986. The non-transferable value: A reply to Aumann. Econometrica 54, 981–984.

Serrano, R., 1997. Reinterpreting the kernel. Journal of Economic Theory 77, 58–80.

Shafer, W., 1980. On the existence and interpretation of value allocations. Econometrica 48, 467–477.

Shapley, L.S., 1953. A value for *n*-person games. In: Kuhn, H.W., Tucker, A.W. (Eds.), Contributions to the Theory of Games II. Princeton University Press, Princeton, NJ, pp. 307–317.

Shapley, L.S., 1969. Utility comparison and the theory of games. La Décision, Agrégation et Dynamique des Ordres de Préférence. Editions du Centre de la Recherche Scientifique, Paris, pp. 251–263.

Stole, L.A., Zwiebel, J., 1996. Intra-firm bargaining under non-binding contracts. Review of Economic Studies 63, 375–410.

Suh, S.-C., Wen, Q., 2003. Multi-agent bilateral bargaining and the Nash bargaining solution. Working paper 03-W06, Vanderbilt University.

Suh, S.-C., Wen, Q., 2006. Multi-agent bilateral bargaining with endogenous protocol. Journal of Mathematical Economy 42 (1), 61–73.

Trockelm, W., 2002. A universal meta bargaining implementation of the Nash solution. Social Choice and Welfare 19, 581–586.

van Damme, E., 1986. The Nash bargaining solution is optimal. Journal of Economic Theory 38, 78–100.

Vidal-Puga, J.J., Bergantiños, G., 2003. An implementation of the Owen value. Games and Economic Behavior 44, 412–427.