Bargaining with commitments*

Juan J. Vidal-Puga
Departamento de Estatistica e IO
Universidade de Vigo
36200 Vigo (Pontevedra), Spain
vidalpuga@uvigo.es

June 23, 2003

Abstract

We study a simple bargaining mechanism in which each player puts a prize to his resources before leaving the game. The only expected final equilibrium payoff can be defined by means of selective marginal contributions vectors, and it coincides with the Shapley value for convex games. Moreover, for 3-player games the selective marginal contributions vectors determine the core when it is nonempty.

1 Introduction

Many economic situations can be modelled as a set of agents or players with independent interests who may benefic from cooperation. Once this cooperation is carried out, the question which arises in how the benefits from cooperation should be distributed among the players.

This problem may be approached by taking two different points of view: one of them is axiomatic, or cooperative, and the other is non-cooperative. The axiomatic point of view focuses on finding allocations which satisfy "fair" (or at least "reasonable") properties, such as efficiency (the final outcome must be optimal), symmetry (players with the same characteristics must receive the same), etc. The non-cooperative point of view leads to the study of the allocations which arise in a given non-cooperative environment.

In this paper, we focus on the non-cooperative approach. We study a non-cooperative framework in the context of transferable utility games, where there exists a divisible commodity that all agents value the same in terms of utility.

We consider a society in which individuals have mechanisms which allow them to make absolute bindings at no cost. Schelling (1980) points out that, if more than one individual simultaneous and uncoordinatedly commit themselves,

*Financial support by the Spanish Ministerio de Ciencia y Tecnología through grant BEC2002-04102-C02-01 is gratefully acknowledged.

we may get inefficiency. Business can not take place if their commitments are incompatible. Thus, we consider a protocol or order in which players commit themselves, and study the allocations that arise in subgame perfect equilibria.

For a detailed discussion on the subject of commitments and its implications, the reader is referred to Crawford (1982), where he studies bargaining environments in which each party can make a commitment which is costly to remove (see also Muthoo, 1992, 1996, and Bolt and Houba, 1998). Schelling (1980) points out that a player may enhance his degree of commitment through delegating on a third party. Commitment through delegation may be found in Katz (1991) and Fershtman and Kalai (1997). Delegation with possible renegotiation may be found in Haller and Holden (1997), Bester and Sákovics (2001), and Corts and Neher (2003).

In our framework, players sequentially choose a prize for their resources and commit themselves to that prize. Finally, the last player chooses the resources he wants to buy and clears the market. This protocol generalizes the bargaining game illustrated by Schelling (1980; Appendix B) as follows: two players may divide \$100 as soon as they agree on how to do it. The game terminates at "midnight", when the bell rings. In order to define "agreement", it is supposed that each player keeps his current offer recorded in a way that a referee can check both offers when the bell rings. If the two players have jointly claimed more than \$100, they get nothing. If they have jointly claimed no more than \$100, the gains are divided in accordance. The presence of commitments is illustrated by a "turnstile that permits a player to leave but not to return; his current offer as he goes through the turnstile remains on the books until the bell rings" (Schelling 1980, p. 276).

Several possible extensions of this mechanism for more than two players are given under the generic name of *Demand Commitment Game*. They are discussed in Bennet and van Damme (1991), Selten (1992), Winter (1994), and Dasgupta and Chiu (1998). A common feature in these models is that, if one or more players "go through the turnstile" demanding a feasible amount (i.e. whatever they can assure by themselves is not less than the sum of their commitments), they may form a coalition and leave the game. Thus, some players may leave the game before all the others have a chance to move. In our model, every player (but the last one) commits to a prize. Our mechanism improves on previous ones in two aspects. First, the mechanism is simpler. Second, the range of results is increased. In Winter's and Dasgupta and Chiu's, the Shapley value arises for convex games. But if the game is not convex the equilibrium payoff may be inefficient[1]. We show that in our mechanism the equilibrium outcome is always efficient in a nonrestrictive class of games, and it coincides with the Shapley value for convex games. For 3-player games, Dasgupta and Chiu show that all the possible outcomes constitute the vertices of the core, when the core

[1] In Dasgupta and Chiu's model, efficiency in the non convex case is achieved by means of prizes and penalties from the planner to the players. For penalties large enough, the Shapley value arises for any game, and the planner does not gain nor loose anything in equilibrium. We think this result is unsatisfactory. For example, there may be a utility transfer to the players from outside out of the equilibrium path.

has nonempty interior. We prove that this result also applies in our mechanism. Furthermore, the outcome is also characterized for simple games.

In Section 2, we introduce the notation used throughout the paper. In Section 3 we define a new value for cooperative games, the *selective value*, based on selective marginal contributions vectors. We also study the selective value in some important classes of games. In particular, the selective value coincides with the Shapley value for convex games. In Section 4 we formally describe the non-cooperative mechanism[2]; and we prove that the selective value is the only expected final payoff in subgame perfect equilibria. We have then given an additional non-cooperative motivation to the Shapley value for convex games.

2 The model

We begin with some basic notations. Given a finite set A, by 2^A we denote the cardinal set of A, by $|A|$ the cardinality of A, and by $\mathbb{R}^A$ the set of real $|A|$-tuples whose indices are the elements of A. Given a function $f : S \subset 2^A \to \mathbb{R}$, by $\arg\max\limits_{T\subset S} \{f(T)\}$ we denote the set of subsets $T \subset S$ which maximize $f(T)$.

Let (N, v) be a *TU game* with transferable utility (TU game), where $N = \{1, 2, ..., n\}$ is the set of *players* and v is the *characteristic function*, which assigns a real number $v(S)$ to every *coalition* $S \subset N$, $S \neq \emptyset$ and $v(\emptyset) = 0$. This value $v(S)$ represents the utility that players in S are able to achieve by themselves when playing cooperatively. Following usual practice, we often refer to "the game v" instead of "the TU game (N, v)". We denote by $TU(N)$ the set of all TU games on the set of players N. We denote by TU the set of all TU games.

We say that v is *convex* if $v(T) - v(T\setminus\{i\}) \leq v(S) - v(S\setminus\{i\})$ when $i \in T \subset S$, *zero-monotonic* if $v(S) + v(\{i\}) \leq v(S \cup \{i\})$ when $i \notin S$, and *strictly zero-monotonic* if $v(S) + v(\{i\}) < v(S \cup \{i\})$ when $i \notin S$. Notice that if the game is convex then it is zero-monotonic. We say that v is *monotonic* if $v(T) \leq v(S)$ whenever $T \subset S$.

The *core* $C(v)$ of the game v is the set of vectors $x \in \mathbb{R}^N$ such that $\sum\limits_{i\in N} x_i = v(N)$ and $\sum\limits_{i\in S} x_i \geq v(S)$ for all $S \subset N$. The core of a game may be empty. However, if the game is convex, its core is nonempty.

Let Π be the set of all orders on N. Given $\pi \in \Pi$ and $i \in N$, we define the set of *predecessors* of i under π as the set of players who come before i under the order π. Namely,

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

We also denote $\overline{P_i^\pi} := P_i^\pi \cup \{i\}$.

We define the *marginal contribution* of player i to the game v under the order π as

$$d_i^\pi(v) := v(N\setminus P_i^\pi) - v\left(N\setminus \overline{P_i^\pi}\right).$$

[2] In order to avoid ambiguities, we use the term *non-cooperative mechanism*, or simply *mechanism*, when referring to a non-cooperative game.

This notation differs from the usual[3]. However, this one is more suitable for our purposes. For simplicity, we write d_i^π instead of $d_i^\pi(v)$ and d^π instead of $d^\pi(v)$.

The marginal contributions vectors can be explained as follows. All the players are together in a room. They sequentially leave the room, and by doing so each leaving player receives his marginal contribution to the players still in the room; i.e. the difference between what players can get by themselves before and after he leaves the room.

We define the *Weber set* $W(v)$ of v as the convex hull whose vertices are the vectors d^π's. If v is convex, $W(v) = C(v)$.

A *value* on $G \subset TU(N)$ is a map $f : G \longrightarrow \mathbb{R}^N$. The *Shapley value* (Shapley, 1953) for the game v is defined as the average of the marginal contributions vectors. Namely

$$\varphi(v) := \frac{1}{|\Pi|} \sum_{\pi \in \Pi} d^\pi.$$

We say that a value f on G satisfies *efficiency* if for any $v \in G$, $\sum\limits_{i \in N} f_i(v) = v(N)$. A value f on G satisfies *core selection* if $f(v) \in C(v)$ for all $v \in G$ such that $C(v) \neq \emptyset$.

We say that v is a *simple game* if $v(S) \in \{0, 1\}$ for all $S \subset N$ and $v(N) = 1$. A player i in a simple game is a *veto player* if $i \notin S$ implies $v(S) = 0$. The core of a monotonic simple game v with set of veto players $T \subset N$ is the convex hull of imputations $x \in \mathbb{R}^N$ which satisfy $x_i \geq 0$ for all $i \in T$, $\sum\limits_{i \in T} x_i = 1$ and $x_i = 0$ for all $i \in N \backslash T$. Thus, the core of v is nonempty if and only if $T \neq \emptyset$.

3 The selective value

We define a selective marginal contribution of a player. As the marginal contributions, imagine all players sequentially leave a room. Again, a player gets the difference between what players can get by themselves before and after he leaves the room. However, in computing what a coalition gets by itself, we assume its members may *select* some of the players already outside the room and use their resources by paying the prize they got on leaving. We think that this interpretation may justify the term *selective marginal contributions*.

Formally, let $\pi \in \Pi$. We can assume without loss of generality that $\pi = (1...n)$.

We define the *selective marginal contribution* of player $i \in N$ in the order N as

[3] Usually, the marginal contribution of player i to the game v under de order π is given by $v\left(\overline{P_i^\pi}\right) - v\left(P_i^\pi\right)$.

$$e_i^\pi(v) := \max_{S \subset P_i^\pi} \left\{ v(N \backslash S) - \sum_{j \in P_i^\pi \backslash S} e_j^\pi(v) \right\} - \max_{S \subset P_i^\pi} \left\{ v(N \backslash (S \cup \{i\})) - \sum_{j \in P_i^\pi \backslash S} e_j^\pi(v) \right\}$$

for $i = 1, 2, ..., n-1$ and

$$e_n^\pi(v) := \max_{S \subset P_n^\pi} \left\{ v(N \backslash S) - \sum_{j \in P_n^\pi \backslash S} e_j^\pi(v) \right\}.$$

For simplicity, we write e_i^π instead of $e_i^\pi(v)$ and e^π instead of $e^\pi(v)$.

Notice that what a coalition $N \backslash P_i^\pi$ can get by itself is not $v(N \backslash P_i^\pi)$, but $\max_{S \subset P_i^\pi} \left\{ v(N \backslash S) - \sum_{j \in P_i^\pi \backslash S} e_j^\pi \right\}$; i.e. the coalition can *select* some of the players already outside the room and use their resources by paying their prize.

Notice that

$$e_1^\pi = v(N) - v(N \backslash \{1\}) = d_1^\pi. \tag{1}$$

Next lemma provides a simplification of these formulas.

Lemma 1 *Given $i \in N$ and $\pi \in \Pi$,*

$$\max_{S \subset P_i^\pi} \left\{ v(N \backslash S) - \sum_{j \in P_i^\pi \backslash S} e_j^\pi \right\} = v(N) - \sum_{j \in P_i^\pi} e_j^\pi.$$

Proof. We assume $\pi = (12...n)$. We proceed by induction on i. For $i = 1$ the result is trivial, since $P_1^\pi = \emptyset$. Assume the result is true for $1, 2, ..., i-1$. Let $S_0 \subset P_i^\pi$, $S_0 \neq \emptyset$. We need to prove that

$$v(N \backslash S_0) - \sum_{j \in P_i^\pi \backslash S_0} e_j^\pi \leq v(N) - \sum_{j \in P_i^\pi} e_j^\pi. \tag{2}$$

Let $k = \max\{j : j \in S_0\}$. Hence, $S_0 \backslash \{k\} \subset P_k^\pi$. By induction hypothesis

$$\begin{aligned} e_k^\pi &= v(N) - \sum_{j \in P_k^\pi} e_j^\pi - \max_{S \subset P_k^\pi} \left\{ v(N \backslash (S \cup \{k\})) - \sum_{j \in P_k^\pi \backslash S} e_j^\pi \right\} \\ &\leq v(N) - \sum_{j \in P_k^\pi} e_j^\pi - v(N \backslash S_0) + \sum_{j \in P_k^\pi \backslash (S_0 \backslash \{k\})} e_j^\pi \end{aligned}$$

from where (2) is easily deduced. ■

We can then define the e^{π}'s as

$$e_i^{\pi} = v(N) - \sum_{j \in P_i^{\pi}} e_j^{\pi} - \max_{S \subset P_i^{\pi}} \left\{ v(N \backslash (S \cup \{i\})) - \sum_{j \in P_i^{\pi} \backslash S} e_j^{\pi} \right\}$$

for $i = 1, ..., n-1$ and

$$e_n^{\pi} = v(N) - \sum_{j \in P_i^{\pi}} e_j^{\pi}.$$

Analogously to the Weber set, we define $W^{\sigma}(v)$ as the convex hull whose vertices are the vectors e^{π}'s.

Given a TU game v, we define the *selective value* $\sigma(v)$ as the vector of average selective marginal contributions. Namely

$$\sigma(v) := \frac{1}{|\Pi|} \sum_{\pi \in \Pi} e^{\pi}.$$

Next proposition characterizes the selective value in convex games, monotonic simple games and zero-monotonic 3-player games with nonempty core.

Proposition 2 *a) If v is convex, then the selective value coincides with the Shapley value.*

b) Let (N, v) be a monotonic simple game, and let T be the set of veto players. Then, the selective value is given by

1. *If $T \neq \emptyset$*

$$\sigma_i(v) = \begin{cases} \frac{1}{|T|} & \text{if } i \in T \\ 0 & \text{if } i \notin T \end{cases}$$

2. *If $T = \emptyset$*

$$\sigma_i(v) = \frac{1}{n}$$

for all $i \in N$.

c) Let v be a zero-monotonic game with $n = 3$ and nonempty core $C(v)$. Then, the selective marginal contributions vectors are the vertices of the core. In particular,

$$W^{\sigma}(v) = C(v).$$

The proof of Proposition 2 is located in the Appendix. Next corollary is immediate:

Corollary 3 *The selective value satisfies core selection for convex games, monotonic simple games and zero-monotonic 3 player games.*

Proposition 2a) allows us to extend the results of Winter (1994) and Dasgupta and Chiu (1998) (cf. Theorem 5 below). Winter (1994, p. 271) suggests that the Shapley value is most adequate in convex games. For non-convex games, the selective value does not coincide with the Shapley value. However, it shows a bigger "stability" (in the sense of "core selection") in both simple games (Proposition 2b)) and 3 player games (Proposition 2c)). Proposition 2c), together with the theorems in next section, also extends Theorem 5 in Dasgupta and Chiu (1998).

Remark 4 *Proposition 2c) does not hold for $n > 3$. For example, consider the symmetric game v with $n = 4$ given by $v(N) = 100$, $v(S) = 50$ if $|S| = 2$ or $|S| = 3$, and $v(S) = 0$ otherwise. The core of this game has a single imputation $(25, 25, 25, 25)$. However, $e^{(1234)} = (50, 0, 0, 50) \notin C(v)$.*

4 The bargaining mechanism with commitments

We define here the bargaining mechanism with commitments. The mechanism has n rounds. In the first round, a player is randomly chosen, being each player equally likely to be chosen. Say, player 1 is chosen. Player 1 must then make a commitment $c_1 \in \mathbb{R}$. Another player (say, player 2) is again randomly chosen among the members of $N\backslash\{1\}$. Player 2, aware of player 1's choice, must make a new commitment $c_2 \in \mathbb{R}$, and so on. When the turn reaches player n, he faces a vector $c \in \mathbb{R}^{N\backslash\{n\}}$ of commitments. He must then propose a coalition $E \subset N\backslash\{n\}$ and he gets the resources of $N\backslash E$ by paying c_i to every $i \in N\backslash(E \cup \{n\})$. The final payoff is then c_i for every player $i \in N\backslash(E \cup \{n\})$, $v(N\backslash E) - \sum_{i \in N\backslash(E \cup \{n\})} c_i$ for player n, and $v(\{i\})$ for every $i \in E$. We say then that players in E are *excluded*.

Next theorem shows that the selective value arises in the bargaining mechanism with commitments as the only expected subgame perfect equilibrium payoff.

Theorem 5 *For strictly zero-monotonic games, there exists a unique expected subgame perfect equilibrium payoff; and it is the selective value.*

To prove this result, we need two lemmas. In order to simplify notation, given a game (N, v), a coalition $S \subset N$, a player $i \in N\backslash S$ and a vector $c \in \mathbb{R}^S$, we define

$$A^S(c) := \max_{T \subset S} \left\{ v(N\backslash T) - \sum_{j \in S\backslash T} c_j \right\}$$

$$B_i^S(c) := \max_{T \subset S} \left\{ v(N\backslash(T \cup \{i\})) - \sum_{j \in S\backslash T} c_j \right\}.$$

Notice that, given $i \in N$ and $\pi \in \Pi$

$$e_i^\pi = \begin{cases} A^{P_i^\pi}\left(\left(e_j^\pi\right)_{j \in P_i^\pi}\right) - B_i^{P_i^\pi}\left(\left(e_j^\pi\right)_{j \in P_i^\pi}\right) & \text{if } \pi(i) \neq n \\ A^{P_i^\pi}\left(\left(e_j^\pi\right)_{j \in P_i^\pi}\right) & \text{if } \pi(i) = n. \end{cases}$$

We also denote by $M(S, i, c)$ the subgame which begins when, after players in $S \subset N \backslash \{n\}$ have stated their commitments $c \in \mathbb{R}^S$, it is player i's turn. If $S = \emptyset$, we write $M(\emptyset, i)$.

Lemma 6 *If v is zero-monotonic, $v(\{i\}) \leq A^S(c) - B_i^S(c)$ for all $S \subset N$, $i \in N \backslash S$, and $c \in \mathbb{R}^S$. If v is strictly zero-monotonic, the inequality is strict.*

Proof. Let $S \subset N$, $i \in N \backslash S$, and $c \in \mathbb{R}^S$. Let $T^i \subset N$ such that

$$T^i \in \arg\max_{T \subset S} \left\{ v(N \backslash (T \cup \{i\})) - \sum_{j \in S \backslash T} c_j \right\}.$$

By zero-monotonicity

$$v\left(N \backslash T^i\right) - \sum_{j \in S \backslash T^i} c_j \geq v\left(N \backslash \left(T^i \cup \{i\}\right)\right) + v(\{i\}) - \sum_{j \in S \backslash T^i} c_j$$

which is precisely $B_i^S(c) + v(\{i\})$. Hence, $A^S(c) \geq B_i^S(c) + v(\{i\})$, and thus the result holds. The proof for strict inequality is analogous. ■

Lemma 7 *Let v be a strictly zero-monotonic game. Assume we are in a subgame perfect equilibrium of the subgame $M(S, i, c)$ and $S \neq N \backslash \{i\}$ (i.e. player i has to commit). Then, player i commits to $c_i = A^S(c) - B_i^S(c)$, and he is not excluded.*

Proof. Assume we are in the subgame $M(N \backslash \{\alpha\}, \alpha, c)$ for some $\alpha \in N$, $c \in \mathbb{R}^{N \backslash \{\alpha\}}$ (i.e. player α is due to choose the set of excluded players). Then, player α is due to exclude a coalition $E \subset N \backslash \{\alpha\}$ such that

$$E \in \arg\max_{T \subset N \backslash \{\alpha\}} \left\{ v(N \backslash T) - \sum_{j \in (N \backslash \{\alpha\}) \backslash T} c_j \right\}. \tag{3}$$

Assume now we are in the subgame $M(S, i, c)$ with $S \neq N \backslash \{i\}$. We proceed by a series of claims:

Claim (A): If player i commits to $c_i < A^S(c) - B_i^S(c)$, then he is not excluded.

Claim (B): Player i commits to $c_i \geq A^S(c) - B_i^S(c)$.

Claim (C): If player i commits to $c_i > A^S(c) - B_i^S(c)$, then he is excluded.

Claim (D): Player i commits to $c_i = A^S(c) - B_i^S(c)$, and he is not excluded.

We proceed by backwards induction on $|S|$. Assume $|S| = n-2$, i.e. player i is the last to make a commitment before the last player (say, α) clears the market. Notice that $S = N\backslash\{\alpha, i\}$.

Proof of Claim (A) for $|S| = n-2$. Assume that $c_i < A^S(c) - B_i^S(c)$. We prove that $i \notin E$. Suppose, on the contrary, that $i \in E$. Let $E' \subset S$ be such that

$$E' \in \arg\max_{T \subset S} \left\{ v(N\backslash T) - \sum_{j \in (N\backslash\{\alpha\})\backslash T} c_j \right\}.$$

Then

$$\begin{aligned}
v(N\backslash E') - \sum_{j \in (N\backslash\{\alpha\})\backslash E'} c_j &= \max_{T \subset S} \left\{ v(N\backslash T) - \sum_{j \in (N\backslash\{\alpha\})\backslash T} c_j \right\} = A^S(c) - c_i \\
&> B_i^S(c) = \max_{T \subset N\backslash\{\alpha\}: i \in T} \left\{ v(N\backslash T) - \sum_{j \in S\backslash(T\backslash\{i\})} c_j \right\} \\
&= \max_{T \subset N\backslash\{\alpha\}: i \in T} \left\{ v(N\backslash T) - \sum_{j \in (N\backslash\{\alpha\})\backslash T} c_j \right\} \\
&\geq v(N\backslash E) - \sum_{j \in (N\backslash\{\alpha\})\backslash E} c_j.
\end{aligned}$$

But this contradicts (3). Thus, $i \notin E$, i.e. player i is not excluded, and therefore his final payoff is c_i.

Proof of Claim (B) for $|S| = n-2$. Assume $c_i < A^S(c) - B_i^S(c)$. By Claim (A), player i can improve his final payoff by committing to c_i' such that $c_i < c_i' < A^S(c) - B_i^S(c)$. Hence, $c_i \geq A^S(c) - B_i^S(c)$ in equilibrium.

Proof of Claim (C) for $|S| = n-2$. Assume $c_i > A^S(c) - B_i^S(c)$. We prove that $i \in E$. Suppose, on the contrary, that $i \notin E$. Then $E \subset S$. Let $E' \subset N\backslash\{\alpha\}$ be such that

$$E' \in \arg\max_{T \subset N\backslash\{\alpha\}: i \in T} \left\{ v(N\backslash T) - \sum_{j \in (N\backslash\{\alpha\})\backslash T} c_j \right\}.$$

Then

$$\begin{aligned}
v\left(N \backslash E'\right) - \sum_{j \in (N \backslash \{\alpha\}) \backslash E'} c_j &= \max_{T \subset N \backslash \{\alpha\} : i \in T} \left\{ v\left(N \backslash T\right) - \sum_{j \in (N \backslash \{\alpha\}) \backslash T} c_j \right\} \\
&= \max_{T \subset S} \left\{ v\left(N \backslash \left(T \cup \{i\}\right)\right) - \sum_{j \in (N \backslash \{\alpha\}) \backslash (T \cup \{i\})} c_j \right\} \\
&= B_i^S\left(c\right) > A^S\left(x\right) - c_i \\
&= \max_{T \subset S} \left\{ v\left(N \backslash T\right) - \sum_{j \in (N \backslash \{\alpha\}) \backslash T} c_j \right\} \\
&\geq v\left(N \backslash E\right) - \sum_{j \in (N \backslash \{\alpha\}) \backslash E} c_j .
\end{aligned}$$

Again, this contradicts (3). Thus, $i \in E$ and player i's final payoff is $v\left(\{i\}\right)$.

Proof of Claim (D) for $|S| = n - 2$. Assume player i is excluded. Thus, his final payoff is $v\left(\{i\}\right)$. By Lemma 6, this is strictly less than what he would get by committing to c_i' with $v\left(\{i\}\right) < c_i' < A^\pi\left(c\right) - B_i^\pi\left(c\right)$. Thus, player i is not excluded. By Claim (C), this means that $c_i \leq A^\pi\left(c\right) - B_i^\pi\left(c\right)$. By Claim (B), equality holds.

Assume now the claims are true for subgames $M\left(T, j, c'\right)$ with $|T| > |S|$.

Proof of Claim (A). Assume that $c_i < A^S\left(c\right) - B_i^S\left(c\right)$. We have to prove that $i \notin E$. Suppose, on the contrary, that $i \in E$. By induction hypothesis, no player in $N \backslash \left(S \cup \{i\}\right)$ is excluded. Thus, $E \subset S \cup \{i\}$.

Let $E' \subset S$ be such that

$$E' \in \arg\max_{T \subset S} \left\{ v\left(N \backslash T\right) - \sum_{j \in (N \backslash \{\alpha\}) \backslash T} c_j \right\}.$$

Then

$$\begin{aligned}
v\left(N \backslash E'\right) - \sum_{j \in (N \backslash \{\alpha\}) \backslash E'} c_j &= \max_{T \subset S} \left\{ v\left(N \backslash T\right) - \sum_{j \in (N \backslash \{\alpha\}) \backslash T} c_j \right\} \\
&= A^S\left(c\right) - \sum_{j \in (N \backslash \{\alpha\}) \backslash S} c_j \\
&> B_i^S\left(c\right) - \sum_{j \in (N \backslash \{i, \alpha\}) \backslash S} c_j \\
&= \max_{T \subset S \cup \{i\} : i \in T} \left\{ v\left(N \backslash T\right) - \sum_{j \in (N \backslash \{\alpha\}) \backslash T} c_j \right\} \\
&\geq v\left(N \backslash E\right) - \sum_{j \in (N \backslash \{\alpha\}) \backslash E} c_j .
\end{aligned}$$

But this contradicts (3). Thus, $i \notin E$ and player i's final payoff is c_i.

Proof of Claim (B). Analogous to the case $|S| = n-2$.

Proof of Claim (C). Assume $c_i > A^S(c) - B_i^S(c)$. We have to prove that $i \in E$. Suppose, on the contrary, that $i \notin E$. By applying the induction hypothesis, we deduce $E \subset S$. Let $E' \subset N\backslash\{\alpha\}$ be such that

$$E' \in \arg\max_{T \subset S \cup \{i\} : i \in T} \left\{ v(N\backslash T) - \sum_{j \in (N\backslash\{\alpha\})\backslash T} c_j \right\}.$$

Then

$$\begin{aligned}
v(N\backslash E') - \sum_{j \in (N\backslash\{\alpha\})\backslash E'} c_j &= \max_{T \subset S \cup \{i\} : i \in T} \left\{ v(N\backslash T) - \sum_{j \in (N\backslash\{\alpha\})\backslash T} c_j \right\} \\
&= \max_{T \subset S} \left\{ v(N\backslash (T \cup \{i\})) - \sum_{j \in (N\backslash\{\alpha\})\backslash(T \cup \{i\})} c_j \right\} \\
&= B_i^S(c) - \sum_{j \in (N\backslash\{\alpha\})\backslash(S \cup \{i\})} c_j \\
&> A^S(x) - \sum_{j \in (N\backslash\{\alpha\})\backslash S} c_j \\
&= \max_{T \subset S} \left\{ v(N\backslash T) - \sum_{j \in (N\backslash\{\alpha\})\backslash T} c_j \right\} \\
&\geq v(N\backslash E) - \sum_{j \in (N\backslash\{\alpha\})\backslash E} c_j.
\end{aligned}$$

Again, this contradicts (3). Thus, $i \in E$ and player i's final payoff is $v(\{i\})$.

Proof of Claim (D). Analogous to the case $|S| = n-2$. ■

An immediate consequence of Lemma 7 is that the equilibrium payoff of any player depends on the identity of the players who come before and after him but not on the way they are ordered. This feature distinguishes our model from Dasgupta and Chiu's (1991), where the order in which players make their demands is prespecified at the beginning of the mechanism and known by all players.

Next remark is important for the proof of Theorem 10 and Theorem 11 below.

Remark 8 *Zero-monotonicity is only needed in the proof of Claim (D).*

Proof of Theorem 5. By Lemma 7, the only possible payoff in equilibrium is e^π, where $\pi \in \Pi$ is given by the order in which the players commit. We must prove that there exists an equilibrium. We consider the following set of strategies: In the subgame $M(S, i, c)$, player i commits to $c_i = A^S(c) - B_i^S(c)$.

If player i is the last one (i.e. $S = N\setminus\{i\}$), he excludes a coalition $E \subset N\setminus\{i\}$ such that E belongs to

$$\arg\max_{T\subset N\setminus\{i\}} \left\{ v\left(N\setminus T\right) - \sum_{j\in(N\setminus\{i\})\setminus T} c_j \right\}. \tag{4}$$

In case of indifference, player i chooses a coalition E with the minimum cardinality $|E|$.

Clearly, the final payoff when players follow this set of strategies is e^{π}. If player i commits to less than $A^S(c) - B_i^S(c)$, he is not excluded (Claim (A)), but his final payoff decreases. If he commits to more than $A^S(c) - B_i^S(c)$, he is excluded (Claim (C)) and his final payoff is $v(\{i\})$. However, by Lemma 6, this amount is less than e_i^{π}. Thus, it is not optimal for him to deviate. We are then in a subgame perfect equilibrium. ∎

If we consider zero-monotonic games, there exists a subgame perfect equilibrium whose expected payoff outcome coincides with the selective value. A possible equilibrium is the one presented in the proof of Theorem 5. However, there may exist subgame perfect equilibria whose expected final payoff outcome is different from the selective value. Consider the next example.

Example 9 *Let v be the "two-left-one-right-glove" 3-player game given by $v(N) = v(\{1,3\}) = v(\{2,3\}) = 1$ and $v(S) = 0$ otherwise. The selective value of v is the only core allocation $(0,0,1)$. We consider the following strategies: Players 2 and 3 play according the strategies described in the proof of Theorem 5, which implement the selective value. However, if the set given in (4) contains more than one coalition, player 3 will exclude the first coalition in (4) given the preference relation $\emptyset \succ \{1\} \succ \{2\} \succ \{1,2\}$. Moreover, player 1 plays according these strategies except in the subgame $M(\emptyset, 1)$, where he commits to $c_1 = 1$.*

It is not difficult to check that these strategies constitute a subgame perfect equilibrium. When the order in which the players are asked is different from (123), the final payoff is $(0,0,1)$. When the order is (123), the final payoff is $(0,1,0)$. Hence, the selective value is not achieved.

If we want to obtain the selective value for general zero-monotonic games, we have to make additional assumptions. For example, Moldovanu and Winter (1994) assume that "each player prefers to be a member of *large* coalitions rather than smaller ones provided that he earns the same payoff in the two agreements". If we make the same assumption in our model, the selective value is implemented. Formally, we consider the following tie-breaking rule:

- If a player i is indifferent between committing to c_i' or c_i and $c_i' < c_i$, he strictly prefers to commit himself to c_i'.
- If the last player is indifferent between excluding E' or E and $E' \subsetneq E$, he strictly prefers to exclude E'.

Theorem 10 *When players follow the tie-breaking rule in zero-monotonic games, the selective value is the unique expected subgame perfect equilibrium payoff.*

Proof. The set of strategies used in the proof of Theorem 5 constitutes a subgame perfect equilibrium for the bargaining mechanism, and furthermore it satisfies the tie-breaking rule.

Now, we prove that, with this tie-breaking rule, Lemma 7 still holds for (not-strictly) zero-monotonic games. More specifically, Claims (A), (B), (C) and (D) hold. Furthermore, the proof of Claims (A), (B) and (C) are analogous. We must prove Claim (D) for the new hypothesis.

Proof of Claim (D) for $|S| = n - 2$. By Claim (A), Claim (C) and the tie-breaking rule, we deduce that $c_i = A^S(c) - B_i^S(c)$. We have to prove that player i is not excluded. Assume he is excluded, i.e. $i \in E$.

If $v(\{i\}) < A^S(c) - B_i^S(c)$, by Claim (A) he can improve his final payoff by committing to c_i' with $v(\{i\}) < c_i' < A^S(c) - B_i^S(c)$. Thus, $v(\{i\}) = A^S(c) - B_i^S(c)$; i.e. $v(\{i\}) = c_i$.

Let $E' = E \backslash \{i\}$. By zero-monotonicity

$$\begin{aligned} v(N \backslash E') - \sum_{j \in (N \backslash \{\alpha\}) \backslash E'} c_j &\geq v(N \backslash E) + v(\{i\}) - \sum_{j \in (N \backslash \{\alpha\}) \backslash E} c_j - c_i \\ &= \max_{T \subset N \backslash \{\alpha\}} \left\{ v(N \backslash T) - \sum_{j \in (N \backslash \{\alpha\}) \backslash T} c_j \right\} \end{aligned}$$

and thus player α is indifferent between excluding E or E'. Since $E \subsetneq E'$, he does not follow the tie-breaking rule. This contradiction proves that player i cannot be excluded.

Proof of Claim (D). Analogous to the case $|S| = n - 2$.

Thus, the only possible payoff in equilibrium is e^π. ∎

Vidal-Puga and Bergantiños (2003) model this tie-breaking rule by "punishing" with a small penalty $\varepsilon > 0$ the players involved in an exclusion. We can do the same in our model. In particular, we assume each excluded player must pay $\varepsilon > 0$. We call this modification the *ε-bargaining mechanism with commitments*. The structure of the mechanism is the same as before. This means that the strategies available for players are the same in both mechanisms. The only difference lies on the following aspect of the payoff function. When the last player presents a coalition E of excluded players, the final payoff is $v(\{i\}) - \varepsilon$ for each $i \in E$.

Theorem 11 *For any $\varepsilon > 0$, the ε-bargaining mechanism with commitments has a unique expected subgame perfect equilibrium payoff for zero-monotonic games, and it is the selective value.*

Proof. By analogous arguments to those presented in the proof of Theorem 10, we only need to prove Claim (D) of Lemma 7. Assume player i commits to $c_i = A^S(c) - B_i^S(c)$ and it is excluded. Then, his final payoff is $v(\{i\}) - \varepsilon$.

By Lemma 6 and Claim (A), this is strictly less than what he would get by committing to c_i' with $v(\{i\}) - \varepsilon < c_i' < A^S(c) - B_i^S(c)$. This contradiction proves the Claim. ■

Remark 12 *As in Vidal-Puga and Bergantiños (2003), the result is also true if the penalty to the excluded players is agent-dependent, i.e., any player i has a penalty $\varepsilon(i) > 0$ for being excluded.*

5 Appendix

Proof of Proposition 2. a) Let $\pi \in \Pi$. Given $i \in N$, we first prove that

$$v\left(N \backslash \overline{P_i^\pi}\right) \geq v\left(N \backslash \left(S \cup \{i\}\right)\right) - \sum_{j \in P_i^\pi \backslash S} e_j^\pi \tag{5}$$

for all $S \subset P_i^\pi$ and

$$e_i^\pi = d_i^\pi. \tag{6}$$

We proceed by induction on $\pi(i)$. For $\pi(i) = 1$, (5) is trivial and (6) coincides with (1). Assume $\pi(i) > 1$ and the results are true for all j such that $\pi(j) < \pi(i)$. Let $S \subset P_i^\pi$. Then, (5) is equivalent to

$$v\left(N \backslash \overline{P_i^\pi}\right) \geq v\left(N \backslash \left(S \cup \{i\}\right)\right) - \sum_{j \in P_i^\pi \backslash S} d_j^\pi. \tag{7}$$

We prove (7) by inverse induction on $|S|$. For $S = P_i^\pi$, it is trivial. Assume (7) is true for coalitions $T \subset P_i^\pi$ such that $|S| < |T| \leq |P_i^\pi|$. Let i^* be the first player in $P_i^\pi \backslash S$, i.e. the only player in $\arg\min_{j \in P_i^\pi \backslash S} \{\pi(j)\}$. Thus

$$P_{i^*}^\pi \subset S. \tag{8}$$

Let $S^* := S \cup \{i^*\} \subset P_i^\pi$. By induction hypothesis

$$\begin{aligned} v\left(N \backslash \overline{P_i^\pi}\right) &\geq v\left(N \backslash \left(S^* \cup \{i\}\right)\right) - \sum_{j \in P_i^\pi \backslash S^*} d_j^\pi \\ &= v\left(N \backslash \left(S \cup \{i, i^*\}\right)\right) + d_{i^*}^\pi - \sum_{j \in P_i^\pi \backslash S} d_j^\pi. \end{aligned} \tag{9}$$

By (8), $\overline{P_{i^*}^\pi} \subset S \cup \{i, i^*\}$. Thus, $N \backslash \overline{P_{i^*}^\pi} \supset N \backslash \left(S \cup \{i, i^*\}\right)$ and, by convexity,

$$v\left(N \backslash P_{i^*}^\pi\right) - v\left(N \backslash \overline{P_{i^*}^\pi}\right) \geq v\left(N \backslash \left(S \cup \{i\}\right)\right) - v\left(N \backslash \left(S \cup \{i, i^*\}\right)\right). \tag{10}$$

Since $d_i^\pi = v\left(N \backslash P_{i^*}^\pi\right) - v\left(N \backslash \overline{P_{i^*}^\pi}\right)$, we apply (10) to (9) in order to obtain (7).

We prove now (6)

$$\begin{aligned}
e_i^\pi &= v(N) - \sum_{j\in P_i^\pi} e_j^\pi - \max_{S\subset P_i^\pi}\left\{ v\left(N\backslash\left(S\cup\{i\}\right)\right) - \sum_{j\in P_i^\pi\backslash S} e_j^\pi \right\} \\
&= v(N) - \sum_{j\in P_i^\pi}\left[v\left(N\backslash P_j^\pi\right) - v\left(N\backslash\overline{P_j^\pi}\right)\right] - v\left(N\backslash\overline{P_i^\pi}\right) \\
&= v(N) - \left[v(N) - v\left(N\backslash P_i^\pi\right)\right] - v\left(N\backslash\overline{P_i^\pi}\right) \\
&= v\left(N\backslash P_i^\pi\right) - v\left(N\backslash\overline{P_i^\pi}\right) = d_i^\pi.
\end{aligned}$$

This completes the proof of part a).

b) Let $\pi \in \Pi$. We can assume without loss of generality that $\pi = (123...n)$.

Assume first $T \neq \emptyset$. Let $i_0 \in T$ be the first veto player in the order π. We prove that $e_i^\pi = 0$ for all $i < i_0$ by induction on i. For $i = 1$, $e_1^\pi = v(N) - v(N\backslash\{1\}) = 0$ because v is monotonic and 1 is not a veto player. Assume $e_j^\pi = 0$ for $1 \leq j < i$. Since $j < i_0$ and i_0 is the first veto player in π, we deduce that j is not a veto player. By induction hypothesis

$$e_j^\pi = v(N) - \max_{S\subset P_j^\pi} v\left(N\backslash\left(S\cup\{j\}\right)\right)$$

by monotonicity

$$= v(N) - v\left(N\backslash\{j\}\right)$$

which equals 0 because v is monotonic and j is not a veto player.

Now, we calculate $e_{i_0}^\pi$. If $i_0 < n$,

$$e_{i_0}^\pi = v(N) - \max_{S\subset P_{i_0}^\pi} v\left(N\backslash\left(S\cup\{i_0\}\right)\right) = v(N) - v\left(N\backslash\{i_0\}\right) = 1.$$

If $i_0 = n$,

$$e_{i_0}^\pi = v(N) = 1.$$

Let $i > i_0$. We calculate e_i^π. If $i < n$,

$$\begin{aligned}
e_i^\pi &= v(N) - 1 - \max\left\{v\left(N\backslash\{i\}\right) - 1, v\left(N\backslash\{i, i_0\}\right)\right\} \\
&= -\max\left\{v\left(N\backslash\{i\}\right) - 1, 0\right\} = 0.
\end{aligned}$$

If $i = n$,

$$e_i^\pi = v(N) - 1 = 0.$$

Thus, $e_{i_0}^\pi = 1$ and $e_i^\pi = 0$ otherwise. Since each veto player has the same probability of being the first one in an order $\pi \in \Pi$, we conclude the result.

Assume now $T = \emptyset$. We prove $e_i^\pi = 0$ for all $i < n$ by induction on i. For $i = 1$, $e_1^\pi = v(N) - v(N \setminus \{1\}) = 0$ because v is monotonic and 1 is not a veto player. Assume $e_j^\pi = 0$ for $j < i < n$. By induction hypothesis

$$e_i^\pi = v(N) - v(N \setminus \{i\})$$

which equals 0 because v is monotonic and player i is not a veto player.

Finally, $e_n^\pi = v(N) = 1$. Thus, $e_n^\pi = 1$ and $e_i^\pi = 0$ otherwise. Since each player has the same probability of being the last one in an order, we conclude the result of part b).

c) The proof is similar to those of Theorem 5 in Dasgupta and Chiu (1998), although the computations are different. Let v be a zero-monotonic game with $n = 3$ and nonempty core $C(v)$. We first prove[4]

$$\max\{y_i : y \in C(v)\} = v(N) - v(N \setminus i) \tag{11}$$

for all $i \in N$.

Let $x_i = \max\{y_i : y \in C(v)\}$. We can assume without loss of generality that $i = 1$.

For any $y \in C(v)$, $y_2 + y_3 \geq v(23)$. Since $y_1 + y_2 + y_3 = v(N)$, we conclude $y_1 \leq v(N) - v(23) = v(N) - v(N \setminus 1)$. Thus, $x_1 \leq v(N) - v(N \setminus 1)$.

Let $y \in C(v)$ such that $y_1 < v(N) - v(N \setminus 1)$.

If $y_2 = v(2)$ and $y_3 = v(3)$, by zero-monotonicity

$$y_1 = v(N) - [v(2) + v(3)] \geq v(N) - v(23) = v(N) - v(N \setminus 1).$$

Thus, $y_2 > v(2)$ or $y_3 > v(3)$. We assume without loss of generality that $y_2 > v(2)$.

Let

$$0 < \varepsilon < \min\{v(N) - v(N \setminus 1) - y_1, y_2 - v(2)\}.$$

Let $y^\varepsilon = y + (\varepsilon, -\varepsilon, 0)$. It is not difficult to check that $y^\varepsilon \in C(v)$. Moreover, $y_1^\varepsilon > y_1$. Thus, $x_1 = v(N) - v(N \setminus 1)$. This proves (11).

Let $\pi \in \Pi$. We assume without loss of generality that $\pi = (123)$.

We define x^π as the vertex of $C(v)$ associated to π. Namely

$$\begin{aligned}
x_1^\pi &= \max\{y_1 : y \in C(v)\} \\
x_2^\pi &= \max\{y_2 : y \in C(v), y_1 = x_1\} \\
x_3^\pi &= v(N) - x_1^\pi - x_2^\pi.
\end{aligned}$$

We prove that $x^\pi = e^\pi$. By (11)

$$x_1^\pi = v(N) - v(23) = e_1^\pi.$$

[4] In this section, we use $v(N \setminus 1)$ instead of the more cumbersome $v(N \setminus \{1\})$. Similarly, $v(ij) = v(\{i, j\})$ and so on.

Moreover

$$\begin{aligned} e_2^\pi &= v(N) - e_1^\pi - \max\{v(13) - e_1^\pi, v(3)\} \\ &= v(23) - \max\{v(13) - v(N) + v(23), v(3)\}. \end{aligned}$$

We study two cases

Case 1: $v(13) - v(N) + v(23) \geq v(3)$. Then

$$e_2^\pi = v(23) - v(13) + v(N) - v(23) = v(N) - v(13).$$

We show that

$$v(N) - v(13) = \max\{y_2 : y \in C(v), y_1 = v(N) - v(23)\} = x_2^\pi.$$

Let $y \in C(v)$ such that $y_1 = v(N) - v(23)$. Then

$$y_2 = v(N) - y_1 - y_3 = v(23) - y_3$$

since $y \in C(v)$, we have $y_1 + y_3 \geq v(13)$ and thus

$$\leq v(23) + y_1 - v(13) = v(N) - v(13).$$

Therefore, $x_2^\pi \leq v(N) - v(13)$.

Let $y \in C(v)$ such that $y_1 = v(N) - v(23)$ and $y_2 < v(N) - v(13)$. We consider $y^\varepsilon := y + (0, \varepsilon, -\varepsilon)$ with $0 < \varepsilon < v(N) - v(13) - y_2$. So, $y_1^\varepsilon = v(N) - v(23)$. It is straightforward to show that $y^\varepsilon \in C(v)$.

Thus, $x_2^\pi = v(N) - v(13) = e_2^\pi$.

Case 2: $v(13) - v(N) + v(23) < v(3)$. Then

$$e_2^\pi = v(23) - v(3).$$

We show that $v(23) - v(3) = \max\{y_2 : y \in C(v), y_1 = v(N) - v(23)\} = x_2^\pi$.

Let $y \in C(v)$ such that $y_1 = v(N) - v(23)$. Then

$$y_2 = v(N) - y_1 - y_3 = v(23) - y_3$$

since $y \in C(v)$, we have $y_3 \geq v(3)$ and thus

$$\leq v(23) - v(3).$$

Therefore, $x_2^\pi \leq v(N) - v(3)$.

Let $y \in C(v)$ such that $y_1 = v(N) - v(23)$ and $y_2 < v(23) - v(3)$. We consider $y^\varepsilon := y + (0, \varepsilon, -\varepsilon)$ with $0 < \varepsilon < v(23) - v(3) - y_2$. So, $y_1^\varepsilon = v(N) - v(23)$. It is straightforward to show that $y^\varepsilon \in C(v)$.

Thus, we conclude that $x_2^\pi = v(23) - v(3) = e_2^\pi$.

Since $x_1^\pi = e_1^\pi$ and $x_2^\pi = e_2^\pi$, it is trivial to check that $x_3^\pi = e_3^\pi$. Hence, $x^\pi = e^\pi$. ■

6 References

1. Bennet E. and van Damme E. (1991) *Demand commitment bargaining: The case of apex games.* In: Selten (ed) Game equilibrium models III: Strategic bargaining, Springer-Verlag, Berlin.

2. Bester H. and Sákovics J. (2001) *Delegated bargaining and renegotiation.* Journal of Economic Behavior and Organization **45**, 459-473.

3. Bolt W. and Houba H. (1998) *Strategic bargaining in the variable threat game.* Economic Theory **11**, 57-77.

4. Corts K.S. and Neher D.V. (2003) *Credible delegation.* European Economic Review **47**, 395-407.

5. Crawford V. (1982) *A theory of disagreement in bargaining.* Econometrica **50**, 607-638.

6. Dasgupta A. and Chiu Y.S. (1998) *On implementation via demand commitment games.* International Journal of Game Theory **27**, 161-189.

7. Fershtman C. and Kalai E. (1997) *Unobserved delegation.* International Economic Review **38**, 763-774.

8. Haller H. and Holden S. (1997) *Ratification requirement and bargaining power.* International Economic Review **38**, 825-851.

9. Hart S. and Mas-Colell A. (1996) *Bargaining and value.* Econometrica **64**, 357-380.

10. Katz M. (1991) *Game-playing agents: unobservable contracts as precommitments.* Rand Journal of Economics **22**, 307-328.

11. Moldovanu B. and Winter E. (1994) *Core implementation and increasing returns to scale for cooperation.* Journal of Mathematical Economics **23**, 533-548.

12. Muthoo A. (1992) *Revocable commitment and sequential bargaining.* Economic Journal **102**, 378-387.

13. Muthoo A. (1992) *A bargaining game based on the commitment tactic.* Journal of Economic Theory **69**, 134-152.

14. Schelling T. C. (1980) *The strategy of conflict.* Harvard University Press. Cambridge. Massachusetts. First edition in 1960.

15. Selten R. (1992) *A demand commitment model of coalition bargaining.* In: Selten R. (ed) Rational iteraction, Essays in honor of John C. Harsanyi, Springer-Verlag, Berlin.

16. 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: 307-317.

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

18. Winter E. (1994) *The demand commitment bargaining and snowballing cooperation*. Economic Theory **4**, 255-273.