Games and Economic Behavior **31,** 106–120 (2000)
doi:10.1006/game.1999.0730, available online at http://www.idealibrary.com on IDEAL®

Implementation of Bargaining Sets via Simple Mechanisms*

David Pérez-Castrillo

Departament d'Economia i d'Història Econòmica & CODE, Universitat Autònoma de Barcelona, 08190 Bellaterra, Barcelona, Spain

and

David Wettstein

Department of Economics, Monaster Center for Economic Research, Ben-Gurion University of the Negev, Beer-Sheva 84105, Israel

Received March 10, 1998

We propose two simple mechanisms that implement two bargaining sets in super-additive environments. The first bargaining set is a close variation of the one proposed by L. Zhou (1994, *Games Econom. Behav.* **6**, 512–526), and the second is the Pareto optimum payoffs of the A. Mas-Colell (1989, *J. Math. Econom.* **18**, 129–139) bargaining set. We adopt a simple framework in which the cooperative outcomes are realized as non-cooperative subgame perfect equilibria in pure strategies of a two-stage game played by an auxiliary set of individuals competing over the cooperative agents. *Journal of Economic Literature* Classification Numbers: C71, C72. © 2000 Academic Press

Key Words: bargaining set; implementation; simple mechanism.

1. INTRODUCTION

Cooperative game theory models the interaction existing in economic and social environments, emphasizing the advantages of cooperation. The different solution concepts put forward in this theory suggest reasonable

*We are grateful to Andreu Mas-Colell, Matthew Jackson, Michael Maschler, and an anonymous referee for very helpful comments. Pérez-Castrillo also acknowledges the financial support by DGES under Projects Nos. PB 96-1192 and PB 97-0181 and the SGR number 98-62. Wettstein gratefully acknowledges support from the Monaster Center for Economic Research. Part of this research was conducted while Wettstein was visiting the Universitat Autònoma de Barcelona, with a grant from the Generalitat de Catalunya.

0899-8256/00 $35.00

ways for the split of the surplus of such cooperation among the various parties.

The most popular cooperative solution concept is the core, which selects those allocations from which no coalition of players can profitably deviate. However, the core ignores subsequent reactions that the deviations can generate. These reactions may render the initial deviation nonprofitable. A subclass of solutions taking into account the future consequences of current actions consists of bargaining sets. The first definition of a bargaining set was suggested by Aumann and Maschler (1964). In recent years, several other bargaining sets have been defined and analyzed by Mas-Colell (1989), Vohra (1991), Vind (1992), and Zhou (1994). An excellent survey of the literature can be found in Maschler (1992). The different definitions of the bargaining set share the common feature of being based on the concepts of objection (deviation) and counter-objection (reaction to the deviation). However, they propose slightly different meanings for these concepts. In contrast to the core, the bargaining sets are non-empty for a large class of games.

A shortcoming common to cooperative solution concepts is the lack of an explicit framework by which they can be reached. The area of research known as implementation theory “rigorously investigates the correspondence between normative goals and institutions designed to achieve (implement) these goals.”[1] Several papers have recently addressed this issue with regard to the core (Pérez-Castrillo (1994), Perry and Reny (1994), Serrano (1995), and Serrano and Vohra (1997)). To the best of our knowledge, only Einy and Wettstein (1999) consider the implementation of bargaining sets. Stearns (1968) and Maschler and Peleg (1976) construct dynamical systems that converge to points in the Aumann and Maschler bargaining set. However, these systems ignore any strategic consideration on the part of the players and can be manipulated by strategic agents.

In this paper, we propose two simple mechanisms that implement two bargaining sets in super-additive environments. The simplicity of these mechanisms is in contrast to the mechanisms suggested in most works. These traditional mechanisms are quite intricate and require a large degree of sophistication on the part of the designer and the participating agents.[2] We adopt a rather simple framework as suggested by Pérez-Castrillo (1994), in which core outcomes are realized as non-cooperative

[1] Quoted from Palfrey (1995).

[2] See, for example, the mechanisms proposed by Moore and Repullo (1988) and Abreu and Sen (1990), which apply to a much larger class of problems and environments.

equilibria of a game played by an auxiliary set of individuals (or institutions) competing over the agents (that are the players in the cooperative game).

The first bargaining set that we implement is a close variation of the one proposed by Zhou (1994). In both sets, the counter-objection must involve some player from both the objection and the remaining set. The only difference between the two sets is that, in contrast to Zhou (1994), we do not impose the condition that a counter-objection should not contain the objection. In the mechanism that we propose, three principals (exogenous players that maximize profits) compete over the agents of the game. Competition occurs in two stages. In the first stage, principals 1 and 2 make simultaneous offers to the agents. Each agent is then provisionally assigned to the principal having made the highest offer to her, where ties are broken in favor of principal 1. If one principal hires the whole set of agents the assignment becomes definitive, salaries are paid, and the game ends. Otherwise, principal 3 has the possibility of submitting a new offer. This new offer has preferential status, in the sense that principal 3 only has to match the maximal provisional offer to attract an agent. However, the final set of workers hired by this principal must contain, if it is not empty, agents from the two groups provisionally formed at stage 1. After this, principals pay salaries to the agents that were assigned to them, receive the value of the coalition composed of these agents, and the game ends. This mechanism implements in subgame perfect equilibrium (SPE) in pure strategies the bargaining set that we propose.

We also implement in SPE the set of efficient allocations in the Mas-Colell (1989) bargaining set through a mechanism similar to, although somewhat more complicated than, the previous mechanism. The main difference between the two mechanisms is that in the second, we assume that principals strictly prefer employing more rather than less agents. This can be interpreted as being part of the design of the game: the designer asks the principals to maximize profits, but also, other things being equal, to hire as many agents as possible.

The rules of the two games are straightforward and rely to a large degree on the intuition underlying the definition of bargaining sets. Using exogenous players allows us to avoid both the use of "modulo (integer) games," and the problem of feasibility of the out-of-equilibrium outcomes.

This paper is organized as follows. In Section 2 we present the basic cooperative game. In Section 3 we implement a close variant of the Zhou (1994) bargaining set, while in Section 4 we implement the Pareto optimal part of the Mas-Colell (1989) bargaining set. Finally, we conclude in Section 5.

2. THE COOPERATIVE GAME

An n-person cooperative game with transferable utility is a pair (N, ν) where $N = \{1, 2, \dots, n\}$ is the set of agents and $v\colon 2^N \to R$ is the characteristic function which satisfies $\nu(\phi) = 0$. The members of 2^N represent all the possible coalitions of agents. The characteristic function yields the maximal payoff any coalition can obtain when its members cooperate. We assume that the game is super-additive, i.e., $\nu(S \cup T) \geq \nu(S) + \nu(T)$ if $S \cap T = \phi$. Moreover, in order to simplify the proofs and to avoid excessive notation, we proceed under the assumption that the preceding inequality is strict if $S \cup T = N$ and $S \notin \{\phi, N\}$. Once we have presented our results under this strong super-additivity assumption, we will indicate how the mechanisms can be modified to obtain the same results under the usual super-additivity assumption.

For any vector $x = (x_1, x_2, \dots, x_n) \in R^n$ and any subset $S \subseteq N$, we write $x(S) = \sum_{i \in S} x_i$. A payoff vector for the game (N, ν) is any vector $x \in R^n$ that satisfies $x(N) \leq \nu(N)$. We denote by $S^c = N/S$ the complement of the subset S in N.

A solution concept for cooperative games associates a set (possibly empty) of payoff vectors with each game. The description of the solution does not involve any procedures through which the agents can achieve an outcome in the solution. The solution concept merely prescribes a set of requirements that payoff vectors must satisfy in order to constitute part of the solution.

The basic components of the definition of a bargaining set are the notions of objection and counter-objection. The different bargaining sets proposed in the literature diverge on the precise definitions of these two concepts. We start by providing the basic set of definitions we are going to use.

Given a payoff vector $x \in R^n$, a pair (S, y) with $S \subseteq N$ and $y \in R^n$ is said to be an *objection* to x if $y(S) \leq \nu(S)$ and $y_i > x_i$ for each i in S. Given an objection (S, y) to the payoff x, a pair (T, z) with $T \subset N$ and $z \in R^n$ is said to be a *counter-objection* to (S, y) if $z(T) = \nu(T)$ and the following two conditions are satisfied:

(BS1) $T \cap S \neq \Phi$ and $T \cap S^C \neq \Phi$.

(BS2) $z_i \geq y_i$ for each $i \in T \cap S$ and $z_i \geq x_i$ for each $i \in T \cap S^c$.

The *core* is the set of all payoff vectors against which there is no objection. It is well known that the core may be empty. Furthermore, the definition of the core ignores subsequent actions. It may well be that the objection to a payoff vector not in the core may set in motion a process that would hurt the objecting individuals. Such a process is the rationale behind the counter-objection. Notice that the counter-objection may generate further deviations, which are ignored in our analysis. See Dutta, Ray, Sengupta, and Vohra (1989) for the construction of a bargaining set that takes into account objections of any order.

To define the bargaining set (BS), we introduce the definition of a justified objection. An objection (S, y) to the payoff x is said to be *justified* if there is no counter-objection to (S, y). The *bargaining set* of the cooperative game (N, ν), denoted by $BS(N, \nu)$, is the set of all payoff configurations $x \in R^n$ against which there is no justified objection.

The BS we define is very close to the definition proposed by Zhou (1994). In his paper, the condition (BS1) also requires that $S/T \neq \phi$. For any given objection, there are more counter-objections satisfying our condition (BS1) than counter-objections satisfying Zhou's definition. Hence, *objecting* is more difficult in our setting and therefore the BS defined here is larger. Since Zhou's BS is non-empty, ours is non-empty as well. Note also that every payoff in the BS is efficient, i.e., $x(N) = \nu(N)$ for all $x \in BS(N, \nu)$. However, it is not necessarily the case that the payoffs in the BS are individually rational, while Zhou's BS does satisfy this property. For example, consider the game $N = \{1, 2, 3\}$, $\nu(1) = 4$, $\nu(2) = \nu(3) = 0$, $\nu(1, 2) = \nu(1, 3) = 16$, $\nu(2, 3) = 17$, $\nu(1, 2, 3) = 24$. The payoff vector $(2, 11, 11)$ is in our BS. However, it is not individually rational for agent 1, hence it is not in Zhou's BS.

We now define the bargaining set proposed by Mas-Colell (1989). It uses different notions of objection and counter-objection. Given a payoff vector $x \in R^n$, a pair (S, y) with $S \subseteq N$ and $y \in R^n$ is said to be a Mas-Colell objection (M-objection) to x if $y(S) = \nu(S)$ and $y_i \geq x_i$ for each $i \in S$ with at least one strict inequality. Given an M-objection (S, y) to x, a pair (T, z) with $T \subseteq N$ and $z \in R^n$ is said to be an M-counter-objection to (S, y) if $z(T) = \nu(T)$ and $z_i \geq y_i$ for all i in $T \cap S$, and $z_i \geq x_i$ for all i in $T \cap S^c$, with at least one strict inequality for some $i \in T$. An M-objection (S, y) to x is justified if there is no M-counter-objection to it.

The *Mas-Colell bargaining set* of the cooperative game (N, υ), denoted hereafter by MBS(N, ν), is the set of all payoff vectors against which there is no justified M-objection. The MBS(N, ν) is non-empty for the class of cooperative games we consider. Furthermore it always contains efficient payoff vectors (see Einy *et al*. (1999)). We denote by POMBS the set of Pareto optimal payoff vectors in MBS.

3. IMPLEMENTATION OF THE BARGAINING SET

In this section we describe a non-cooperative mechanism whose equilibrium outcomes coincide with the BS(N, ν) previously described. The mechanism is intended to solve the problem of a designer who does not know the characteristic function and yet wants to attain outcomes in the BS. The mechanism is to be played by players who are different from the agents described in Section 2. These new players will be called *principals*. There are three principals, 1, 2, and 3, a typical principal will be denoted by j. The principals will compete over the agents via wage offers made to the agents. The payoff of a principal hiring a subset S of the agents is the difference between $v(S)$, that is, the value of the coalition of the agents, and the sum of wages paid out. We interpret this to mean that the designer hires some managers to play the game and offers them the value of the coalitions they hire as payment, but they have to pay the wages they offer to the agents hired.

The mechanism H is played as follows.

Stage 1. Each principal j, for $j = 1, 2$, submits an offer $x^j \in R^n$, where x_i^j is the amount principal j would pay to agent i if he were employed by her. The offers are made simultaneously. Each agent is provisionally assigned to the principal who submits the highest wage for that agent, which is $x_i^Q = \text{Max}\{x_i^1, x_i^2\}$. In case of identical maximal offers, agent i is assigned to principal 1. At the end of stage 1 we have two subsets of agents denoted by S^1 and S^2, where S^j is the set of agents assigned to principal j, for $j = 1, 2$.

If there exists a principal α for which $S^\alpha = N$, the game ends with all agents employed by principal α, the payoff to agent i is x_i^α, the payoff to principal α is $v(N) - x^\alpha(N)$, and the other principals obtain zero profits. If both S^1 and S^2 are non-empty, then the game moves to stage 2, with all principals fully informed about the outcome of stage 1.

Stage 2. Principal 3 can submit a new offer $x^3 \in R^n$. Consider the set $K = \{i \in N \mid x_i^3 \geq x_i^Q\}$. If K contains elements from both S^1 and S^2, then K is assigned to principal 3. Otherwise, principal 3 is not assigned any agents. Denote by T^3 the set of agents assigned to principal 3 (either $T^3 = K$, or $T^3 = \phi$). The principals 1 and 2 hire the sets $T^1 = S^1/T^3$ and $T^2 = S^2/T^3$. The payoff to agent i is x_i^3 if $i \in T^3$ and x_i^Q otherwise. The payoffs to the principals are $v(T^j) - x^j(T^j)$, for $j = 1, 2, 3$.

The main purpose of this section is to show that the mechanism H implements the BS in pure strategies. That is, the set of subgame perfect equilibria in pure strategies (SPE) of the non-cooperative mechanism H coincides with the BS. Previous to the theorem, we present the following two properties of the BS that will be useful in the proof of the implementation result.

LEMMA 1. *(S, y) is a justified objection to $x \in R^n$ if and only if for every $T \subseteq N$ satisfying condition* $(BS1)$:

$$\nu(T) < y(T \cap S) + x(T \cap S^c). \tag{1}$$

Proof. It is easy to see from the definition of a counter-objection that there exists a counter-objection to an objection (S, y) if and only if there exists a set $T \subseteq N$ satisfying condition (BS1) such that $\nu(T) \geq y(T \cap S) + x(T \cap S^c)$. Q.E.D.

LEMMA 2. *If there exists a justified objection (S, y) to $x \in R^n$, then there exists a justified objection (S, y') to x such that $y'(S) < \nu(S)$.*

Proof. Take a justified objection (S, y). The number of sets $T \subseteq N$ satisfying condition (BS1) is finite and Eq. (1) is satisfied for all of them. Therefore, there exists an $\varepsilon > 0$ such that $\nu(T) < y'(T \cap S) + x(T \cap S^c)$ for every $T \subseteq N$ satisfying condition (BS1), where $y'_i = y_i - \varepsilon$ for all $i \in N$, and $y'_i > x_i$ for all $i \in S$. Consequently, (S, y') is a justified objection for which $y'(S) < \nu(S)$ is satisfied. Q.E.D.

We now present our first main result.

THEOREM 1. *The mechanism H implements in* SPE *the set* BS(N, ν).

Proof. (a) We first prove that any payoff in BS(N, ν) can be reached as an SPE of H.

Let $x \in$ BS(N, ν). Consider the following strategies. At stage 1, principals 1 and 2 submit the offers $x^1 = x^2 = x$. At stage 2 (if this stage is reached) principal 3 makes an offer that guarantees to her the highest possible profit, given the strategies actually played at stage 1. We claim that these strategies constitute a SPE whose outcome is that principal 1 hires the whole set N at salaries x. (Note that if principals 1 and 2 use the previous strategies, stage 2 is not reached.)

We prove the claim by contradiction. Suppose that there is a profitable deviation by a principal h. The deviation must take place at stage 1, since by definition no profitable deviation exists for principal 3. Denote by $y^h \in R^n$ the deviating offer. Since the deviation gives a strictly positive profit to the principal and $x(N) = \nu(N)$, principal h must end up with a proper subset of N after stage 1, denoted by S^h. Moreover, $y_i^h \geq x_i$ for all $i \in S^h$ and $y^h(S^h) < v(S^h)$. Let us take $y_i^{h\prime} = y_i^h + \varepsilon$ for all $i \in S^h$, with ε small enough so that $y^{h\prime}(S^h) \leq \nu(S^h)$. The pair $(S^h, y^{h\prime})$ satisfies $y_i^{h\prime} > x_i$ for all $i \in S^h$ and $y^{h\prime}(S^h) \leq v(S^h)$, hence it is an objection to x. Given that $x \in$ BS(N, ν), there exists a counter-objection (T, z) to $(S^h, y^{h\prime})$; that is, $T \subseteq N$ satisfies (BS1) (i.e., $T \cap S^h \neq \phi$ and $T \cap S^{hc} \neq \phi$) and $z \in R^n$ satisfies $z(T) \leq \nu(T)$, $z_i \geq y_i^{h\prime}$ for each $i \in T \cap S^h$ and $z_i \geq x_i$ for each $i \in T \cap S^{hc}$.

If the deviation takes place, both principals end up hiring a proper subset of N at stage 1, and the game goes to stage 2. Principal 3 then has the possibility of making the offer x^3, with $x_i^3 = z_i$ for each $i \in T \cap S^{hc}$, $x_i^3 = z_i - \varepsilon$ for each $i \in T \cap S^h$, and $x_i^3 = -\infty$ for each $i \in T^c$ (when we write $x_i = -\infty$ we mean that x_i is lower than any other offer made to agent i). The strategy x^3 guarantees strictly positive profits to principal 3, since agents hired by principal 3 are exactly those in T and $x^3(T) < \nu(T)$. However, it is not necessarily the best strategy for principal 3. Now, we claim that any strategy that maximizes principal 3's profits must leave principal h with non-positive profits. Otherwise, principal 3 will obtain higher profits by hiring both of her group of agents and the remaining group of agents of principal h, at the corresponding salaries, and she will obtain at least the same profits as the sum of the profits of both principals, since the game is super-additive. Therefore, principal h cannot obtain strictly positive profits. Hence, there is no profitable deviation at stage 1.

(b) Now we prove that any payoff that is the outcome of an SPE of H is in BS(N, ν). We proceed by several steps. Throughout the proof, we denote by $x^F \in R^n$ the wages actually paid to the agents. That is, either $x^F = x^Q$, if the game ends at stage 1, or $x_i^F = \text{Max}\{x_i^Q, x_i^3\}$ if the game reaches stage 2.

(b1) In any SPE, one principal must hire all the agents. We prove the statement by contradiction. Suppose it is not the case. Since profits must be non-negative at equilibrium (otherwise, a principal with negative profits could deviate making offers $x'_{\rm i} = -\infty$ for $i = 1, \dots, n$, not hiring any agents), it is the case that $x^F(T^j) \le \nu(T^j)$ (note that if there are more than one principal hiring agents, stage 2 has necessarily been reached). Moreover, the super-additivity assumptions imply that $\nu(N) > \nu(T^1) + \nu(T^2) + \nu(T^3)$, since at least two sets are non-empty. Then, we claim that there is a profitable deviation by principal 3: $y^3 = x^F$. Following this deviation, principal 3 hires the whole set N, making profits:

$$\nu(N) - x^F(N) > \nu(T^1) - x^F(T^1) + \nu(T^2) - x^F(T^2) + \nu(T^3) - x^F(T^3) \ge \nu(T^3) - x^F(T^3).$$

(b2) In any SPE, all principals make zero profits. Suppose that it is not the case, so the principal that, at equilibrium, hires all the agents makes positive profits, that is, $\nu(N) > x^F(N)$. Either principal 1 or 2 (or both) obtain zero profits at equilibrium. Suppose it is principal 2. Then, she could deviate from her initial strategy making an offer $y_i^2 = x_i^F + \varepsilon$, for all $i \in N$, with $\varepsilon > 0$ but small enough so that $\nu(N) > x^F(N) + \varepsilon N$. This ensures that principal 2 hires all the agents and makes strictly positive profits.

(b3) In any SPE, $x^F = x^Q$ and $x^F(N) = v(N)$. To prove that $x^F = x^Q$ note, first, that $x^F \geq x^Q$ by definition of x^F and, second, any offer $x_i^3 > x_i^Q$ is dominated by $y_i^3 = x_i^Q$. On the other hand, $x^F(N) = \nu(N)$ is a direct consequence of the property that the principal that, at equilibrium, hires all the agents (b1) makes zero profits (b2).

(b4) The payoff $x^F \in \text{BS}(N, \nu)$. Suppose it is not the case. Then, by Lemma 2, there is a justified objection (S, y) to x^F such that $\nu(S) > y(S)$. Consider the following deviation by principal 1: $y_i^1 = y_i$ for all $i \in S$ and $y_i^1 = -\infty$ for all $i \in S^c$. We claim that this new strategy is a profitable deviation from x^1, thus contradicting the fact x^F is a vector of final wages of an SPE. First, since (S, y) is an objection to x^F, $y_i^1 > x_i^F$ for all $i \in S$, and the set S is hired by principal 1 at stage 1. Second, principal 1 makes strictly positive profits if she keeps S after stage 2, since $\nu(S) > y^1(S)$. Third, given that the deviation (S, y) is justified, there is no counter-objection (T, z) to (S, y); that is, $\nu(T) < y^1(T \cap S) + x^F(T \cap S^c)$ for any T satisfying (BS1). Therefore, every offer in which principal 3 ends up with a non-empty subset of agents yields negative profits to this principal. Hence, principal 1 keeps the set S at stage 2, so y^1 is a profitable deviation.

We have proven that the salaries actually paid to the agents constitute a payoff in $\text{BS}(N, \nu)$. Therefore, the equilibrium outcome constitutes a payoff in $\text{BS}(N, \nu)$. Q.E.D.

The mechanism H is a simple mechanism that allows the designer to reach the set of outcomes that are in the BS of the game. The definition of the BS we use is different from (even if very closely related to) other definitions proposed in the literature. The two features of our BS that are crucial for the implementation via mechanism H (described above) are as follows:

(i) In the definition of a counter-objection, it is sufficient that the agents receive the same payment that they are offered either in the main proposal or in the objection. That is, in the condition (BS2) there must be weak inequalities. Otherwise, in mechanism H, there does not exist a maximum for the program of the principal called to play at stage 2. Moreover, and more importantly, this is the characteristic that makes Lemmas 1 and 2 hold. That is, it ensures that any objector has the option to make strictly positive profits.

(ii) The restrictions that the counter-objector faces must be symmetric in the sense that they apply in the same way to the objector and to the original proposer, so that it is not necessary to identify the objector in order to operate the mechanism. Determining the identity of the objector without knowing the characteristic function (which we assume is known only by the agents and the principals and not by the designer) is very problematic and is not possible in mechanism H.

These two features enable us to make the following comment. If we take an alternative definition of the bargaining set that accords with our definition of an objection, but for which condition (BS1) is different and consider an alternative condition for (BS1) that is symmetric, in the sense that we can exchange the roles of S and S^c without modifying the condition, a mechanism very similar to H implements this alternative bargaining set.

We do not think a mechanism along the lines of mechanism H can implement a BS that fails to satisfy requirement (ii). Zhou's (1994) BS specifies non-symmetric restrictions. Therefore we do not implement it, although it satisfies (i). This is the reason why we adopted a slightly different definition.

On the other hand, the BS proposed by Mas-Colell (1989) does satisfy (ii), but it fails to satisfy (i). In order to implement the MBS, we first have to modify the second step of the mechanism so that there is a maximum of the program of principal 3 and second, to introduce new incentives for the principals to deviate, even if they will not obtain profits. Moreover, the fact that an M-objection requires strict inequalities only for one of the deviating agents also makes the task more difficult. The next section shows one way to deal with the preceding problems.

4. IMPLEMENTATION OF THE PARETO OPTIMAL MAS-COLELL BARGAINING SET

We now construct a new mechanism H' whose SPE outcomes coincide with the set of Pareto optimum payoff vectors in the Mas-Colell bargaining set (POMBS). As previously, the players are the principals who compete over the agents. We now assume that there are four principals. We add one further assumption that principals, all other things being equal, strictly prefer employing more rather than less agents. This can be viewed as a tie-breaking rule that introduces lexicographic preferences whereby the principals' first criterion is profit maximization and the second criterion is hiring the largest possible group of agents. This assumption is restrictive but it is necessary to provide the motivation for objections since a justified M-objection always yields zero profits.[3]

[3]Similar lexicographic preferences can also be found in the automata literature. For example, Abreu and Rubinstein (1988) assume that preferences over monetary payoffs and complexity are lexicographic: machine A is strictly preferred to machine B if it yields higher monetary payoff or it yields the same payoff with lower complexity. The treatment of ties has also received attention in the learning literature, where it is shown that different tie breaking rules lead to very different outcomes (see, for example, Monderer and Sela (1996)).

The mechanism consists of two stages. In the first stage, principals 1 and 2 submit wage offers and, in the second stage, principals 3 and 4 possibly bid for the right to make a second proposal, possibly enticing a new set of agents to work for them. The formal description of *the mechanism H'* is as follows.

Stage 1. Principals 1 and 2 submit simultaneous offers $x^1, x^2 \in R^n$. Should one offer weakly dominate another, the weakly dominated offer is discarded. Among the remaining offers, each agent is assigned to the principal who submitted the highest wage for that agent, namely $x_i^Q = \text{Max}\{x_i^1, x_i^2\}$. In case of identical offers, agent i is assigned to the principal whose sum of offers is the smallest; in case of identical sums, he is assigned to principal 1. At the end of this stage there are two subsets of agents denoted by S^1 and S^2, where S^j is the set of agents assigned to principal j.

If there exists a principal α for which $S^\alpha = N$, all agents are employed by principal α, the payoff to agent i is x_i^α, the payoff to principal α is $\nu(N) - x^\alpha(N)$, the other principals obtain zero profits, and the game ends. Otherwise, we go to stage 2.

Stage 2. Principals 3 and 4 simultaneously submit pairs (b^3, x^3) and (b^4, x^4), where $b^3, b^4 \in R_+$ and x^3, $x^4 \in R^n$. Denote by β the principal submitting the largest b^j, where ties are broken in favor of principal 3. If $b^\beta = 0$, principal β is assigned the empty set. In the case where $b^\beta > 0$, principal β pays b^β to agent 1[4] and she is assigned the set $K = \{i \in N \mid x_i^\beta \geq x_i^Q\}$. Denote by T^β the set hired by principal β, either $T^\beta = K$ or $T^\beta = \phi$. Principals 1 and 2 employ the sets $T^j = S^j / T^\beta$, for $j = 1, 2$. The payoff to agent i, for $i = 1, \ldots, n$, is $\text{Max}\{x_i^Q, x_i^\beta\}$, to which agent 1 adds the bid b^β. The payoffs to the principals are $\nu(T^\beta) - x^\beta(T^\beta) - b^\beta$ for principal β, zero for the principal different from β and playing at stage 2, and $\nu(T^j) - x^j(T^j)$, for the principals $j = 1, 2$.

Prior to stating the theorem we introduce one additional definition. Given an allocation $x \in R^n$, let $B(x)$ represent the maximal gain for a principal at stage 2, where x is the vector of maximal offers to the agents at stage 1. That is,

$$B(x) = \text{Sup}\{\nu(T) - z(T) \mid T \subseteq N \text{ and } z_i \geq x_i \text{ for all } i \in T\}.$$

Notice that $B(x) \geq 0$, given that $\nu(\phi) - x(\phi) = 0$. Also note that if the supremum is achieved for coalition T, $z_i = x_i$ for all $i \in T$, so $B(x) = \text{Sup}\{\nu(T) - x(T) \mid T \subseteq N\}$.

[4] The bidding money can go to any agent, or to a third party. The only restriction is that it should not go to any principal, since this could induce strategic behavior.

We now show that the SPE of mechanism H' described above coincide with the set of payoff vectors in POMBS.

THEOREM 2. *The mechanism* H' *implements in SPE the set* *POMBS*(N, v).

Proof. (a) We first prove that any payoff in POMBS(N, v) can be reached as an SPE of H'. Let $x \in$ POMBS. We prove that the following set of strategies constitutes an SPE. At stage 1, $x^1 = x^2 = x$. At stage 2, if it is reached, principals 3 and 4 submit a number $b^3 = b^4 = B(x^Q)$, where x_i^Q is the maximum offer actually made to agent i at stage 1. If $B(x^Q) = 0$, then the wage offers are irrelevant since principals bidding zero are assigned the empty set by definition. Moreover, if $B(x^Q) > 0$, each principal j, for $j = 3, 4$, selects the vector x^j in such a way that $x^j \leq x^Q$ and if she is selected as principal β, then $T^\beta = \{i \in N | x_i^\beta \geq x_i^Q\}$ will be one among the largest sets that produce profit $B(x^Q)$.

If the previous strategies are followed, principal 1 will hire all the agents at stage 1 with zero profits and the agents will be paid according to x. To see that the previous strategies constitute a SPE of the mechanism H', let us first analyze stage 2. The strategy of a principal j, for $j = 3, 4$, in stage 2 consists of sending the message $(B(x^Q), x^j)$, where the offer x^j leaves principal j with one of the coalitions guarantying profits $B(x^Q)$. The coordinates of x^j over the coalition would coincide with the wage offers each coalition member holds from the previous stage. To see that these messages comprise a Nash equilibrium, consider a subgame where $B(x^Q) > 0$ (if $B(x^Q) = 0$, then any deviation would generate either no change or negative profits). The outcome of the previous strategies is that principal 3 employs one of the largest possible set of agents while still making zero profits, since the bid $B(x^Q)$ is precisely the difference between the value of the coalition hired by the principal and the sum of the salaries paid to this coalition. A higher bid would generate losses for any principal. A lower bid for principal 3 would result in principal 3 not hiring any agents and thus would make the principal worse off, while it could not change the outcome if made by principal 4.

Let us now analyze stage 1. Suppose there is a deviation y^h which, when announced by principal $h \in \{1, 2\}$ in stage 1, leads to a strictly preferred outcome for principal h. This deviation must leave principal h with positive profits if $h = 1$, or with non-negative profits and a non-empty set of employees if $h = 2$. In either case, following the deviation, both principals employ non-empty sets of agents at the end of stage 1. Hence, following a potentially profitable deviation, the mechanism moves into stage 2. Denote by S^h the set of agents provisionally employed by the deviating principal at the end of stage 1.

We now show that $B(x^Q) > 0$ (note that $x_i^Q = y_i^h$ if $i \in S^h$ and $x_i^Q = x_i$ if $i \in S^{hc}$). If (S^h, y^h) is an M-objection, there exists a counter-objection (T, z) with $z(T) < v(T)$; hence $B(x^Q) > 0$. The deviation y^h will not be an M-objection when $y_i^h = x_i$ for all $i \in S$. However, if this is the case the deviator must be principal 1, since principal 2 could not have gained from such a deviation (this is so because an offer by principal 2 with $y_i^2 = x_i$ for all $i \in S$ and $y_i^2 \leq x_i$ for all $i \in S^c$ with at least one strict inequality is weakly dominated by the offer of principal 1, which results in not hiring any agents). If principal 1 is the deviator, this can be profitable only if $v(S^1) > y^1(S^1)$; hence $B(x^Q) > 0$ holds as well. Therefore, the bid that principals 3 and 4 will submit at stage 2 is positive.

As a result of the Nash equilibrium for the game beginning at stage 2, the deviating principal h in stage 1 cannot be better off. We show that either principal h has strictly negative profits or ends up employing no agents at all. Suppose that at the end of stage 2, h employs a non-empty set T^h of agents with $y^h(T^h) \leq v(T^h)$. Then, due to the super-additivity assumption, it is not possible that the winning bid and wage offer proposed in stage 2 were the best response. The coalition could have been enlarged without decreasing profits. Hence, in the case where principal h employs any agents the profits are negative. Therefore, there cannot be a profitable deviation in stage 1.

(b) We now prove that any payoff that is the outcome of an SPE of H' is in the POMBS(N, v). Note that we can mimic the proofs of steps (b1) and (b2) in Theorem 1 to prove the following two claims. First, in any SPE one principal must hire all the agents. Second, in any SPE all principals make zero profits.

Given the two previous properties, the sum of payments by the principal who hires all the agents in equilibrium is $v(N)$, so we obtain efficiency. If this payment vector does not belong to the POMBS, there exists a justified objection (S, y). A principal $h \in \{1, 2\}$ not employing any agents who deviate at stage 1 offering $y_i^h = y_i$ for $i \in S$ and $y_i^h = -\infty$ for $i \in S^c$ can employ the coalition S (the offer y is not weakly dominated by x and $y(N) < x(N)$, so the agents in S are assigned to principal h). If coalition S after stage 2 is maintained, the deviating principal is better off. Since the objection is justified, any principal bidding $b^j > 0$ in stage 2 would make strictly negative profits. Hence, the outcome of stage 1 remains, and the deviating principal has profited. This is in contradiction to the fact that we were at an SPE. Q.E.D.

We have constructed the mechanisms implementing the bargaining sets under a strong super-additivity assumption. This hypothesis simplifies the presentation since every efficient outcome involves the formation of the grand coalition. We now indicate the modifications needed when efficient outcomes can also be generated by partitions other than the grand coalition. Let us call an efficient structure a partition of the set of agents such that the sum of the value of the coalitions in the partition equals the value of the grand coalition. Denote by D the maximum number of coalitions an efficient structure could possibly contain.

We introduce two modifications in the mechanism H. First, we assume that there are at least $D + 1$ principals competing over the agents at stage 1. Second, at stage 2, that is reached if there are at least two principals which are provisionally assigned agents at stage 1, a non-empty set hired by principal $D + 2$ must intersect with all non-empty sets provisionally assigned to principals at stage 1. This mechanism implements the BS. As to mechanism H', if we assume that there are at least $D + 1$ principals competing over agents at stage 1, then the mechanism implements the POMBS.

5. CONCLUSIONS

We have constructed two simple mechanisms that implement two bargaining sets. The first is a variation on the Zhou (1994) bargaining set and the second is the Pareto optimal subset of the Mas-Colell (1989) bargaining set. These mechanisms are more straightforward than those previously suggested by Einy and Wettstein (1999). The simplicity is obtained at the cost of introducing an auxiliary set of individuals (as in Pérez-Castrillo (1994)).

The implementation of the Pareto optimal subset of the Mas-Colell bargaining set turns out to be less elegant than that of the first bargaining set due to the need for the principals to use an explicit tie breaking rule. While one can see this as a weakness of the implementation of the Mas-Colell bargaining set, the need for the rule is actually due to the fact that Mas-Colell justified objections are always efficient, that is, they all share the whole value of the coalition among its members. Hence, a property that seems desirable from a cooperative point of view interferes with the non-cooperative implementation through our class of mechanisms.

Appropriately modified mechanisms could similarly implement other bargaining sets where the treatment of objections and counter-objections is symmetric. It should be noted that the bargaining sets of Zhou (1994) and Aumann−Maschler (1964) do not fall into this class. It might well be that in order to implement these two bargaining sets one would have to resort to more sophisticated constructions.

REFERENCES

Abreu, D., and Rubinstein, A. (1988). “The Structure of Nash Equilibrium in Repeated Games with Finite Automata,” *Econometrica* **56**, 1259–1282.

Abreu, D., and Sen, A. (1990). “Subgame Perfect Implementation: A Necessary and Almost Sufficient Condition,” *J. Econom. Theory* **50**, 285–299.

Aumann, R., and Maschler, M. (1964). “The Bargaining Set for Cooperative Games,” in *Advances in Game Theory* (M. Dresher, L. S. Shapley, and A. W. Tucker, Eds.). Princeton, NJ: Princeton Univ. Press.

Dutta, B., Ray, D., Sengupta, K., and Vohra, R. (1989). “A Consistent Bargaining Set,” *J. Econom. Theory* **49**, 93–112.

Einy, E., Holzman, R., and Monderer, D. (1999). “On the Least Core and Bargaining Sets,” *Games Econom. Behav.* **28**, 181–188.

Einy, E., and Wettstein, D. (1999). “A Non-Cooperative Interpretation of Bargaining Sets,” *Rev. Econom. Design*, to appear.

Mas-Colell, A. (1989). “An Equivalence Theorem for a Bargaining Set,” *J. Math. Econom.* **18**, 129–139.

Maschler, M. (1992). “The Bargaining Set, Kernel, and Nucleolus,” in *Handbook of Game Theory* (R. J. Aumann and S. Hart, Eds.), Vol. 1. Amsterdam: Elsevier Science Publishers.

Maschler, M., and Peleg, B. (1976). “Stable Sets and Stable Points of Set-valued Dynamic Systems with Applications to Game Theory,” *SIAM J. Control Optim.* **14**, 985–995.

Monderer, D., and Sela, A. (1996). “A 2 × 2 Game without the Fictitious Play Property,” *Games Econom. Behavior* **14**, 144–148.

Moore, J., and Repullo, R. (1988). “Subgame Perfect Implementation,” *Econometrica* **56**, 1191–1220.

Palfrey, T. R. (1995). “Implementation Theory,” in *Handbook of Game Theory* (R. J. Aumann and S. Hart, Eds.), Vol. 3. Amsterdam: Elsevier Science Publishers.

Pérez-Castrillo, D. (1994). “Cooperative Outcomes through Non-Cooperative Games,” *Games Econom. Behavior* **7**, 428–440.

Perry, M., and Reny, P. (1994). “A Non-Cooperative View of Coalition Formation and the Core,” *Econometrica* **62**, 795–817.

Serrano, R. (1995). “A Market to Implement the Core,” *J. Econom. Theory* **67**, 285–294.

Serrano, R., and Vohra, R. (1997). “Non-Cooperative Implementation of the Core,” *Soc. Choice Welf.* **14**, 513–525.

Stearns, R. E. (1968). “Convergent Transfer Schemes for *N*-person Games,” *Trans. Amer. Math. Soc.* **134**, 449–459.

Vind, K. (1992). “Two Characterizations of Bargaining Sets,” *J. Math. Econom.* **21**, 89–97.

Vohra, R. (1991). “An Existence Theorem for a Bargaining Set,” *J. Math. Econom.* **20**, 19–34.

Zhou, L. (1994). “A New Bargaining Set of an *N*-Person Game and Endogenous Coalition Formation,” *Games Econom. Behavior* **6**, 512–526.