

Bargaining and Bargaining Sets¹

Roberto Serrano² and Rajiv Vohra³

Department of Economics, Brown University, Providence, Rhode Island 02912

Received February 16, 2001; published online January 31, 2002

We argue that, in exchange economies, various versions of the bargaining set cannot be implemented in Nash equilibrium. We then propose an extensive form mechanism that exactly implements the Aumann–Davis–Maschler bargaining set in subgame perfect equilibrium. The mechanism corresponds closely to the description of the bargaining set and respects feasibility in and out of equilibrium. *Journal of Economic Literature* Classification Numbers: C71, C72, C78, D51. © 2002 Elsevier Science (USA)

Key Words: bargaining set; implementation; approximate implementation; Nash equilibrium; subgame perfect equilibrium.

1. INTRODUCTION

The bargaining set can be viewed as a solution concept which attempts to remedy a shortcoming of the core by requiring objections to be immune to counterobjections. Recall that an allocation x does not belong to the core of an exchange economy if there exists a coalition which with its own resources can make all its members better off compared to x . However, it is possible that such an objection has a counterobjection in the sense that another coalition can improve upon x and ensure that all agents common to the two coalitions prefer to join the new coalition. In this case, the original objection is not justified. The bargaining set consists of all allocations to which there exists no justified objection. Since allocations immune to objections are also immune to justified objections, the core is contained in the bargaining set.

Various versions of the bargaining set differ in the way they impose restrictions on the objecting and counterobjecting coalitions, as well as in

¹ We are grateful to David Perez-Castrillo, an Associate Editor, and two anonymous referees for helpful comments. This research was supported by Salomon research awards from Brown University. Serrano also acknowledges an Alfred P. Sloan Foundation research fellowship.

² To whom correspondence should be addressed. E-mail: Roberto_Serrano@brown.edu.

³ E-mail: Rajiv.Vohra@brown.edu.

the requirements of strict versus weak improvements.⁴ The first version of the bargaining set was introduced in Aumann and Maschler (1964) and in Davis and Maschler (1963, 1967).⁵ Mas-Colell (1989) introduced another version of the bargaining set and proved a remarkable equivalence theorem for atomless exchange economies.

The present paper is concerned with implementing these bargaining sets (or, more precisely, the associated correspondences) in the context of an exchange economy. We begin by showing that the bargaining set correspondences do not satisfy Maskin monotonicity and cannot, therefore, be implemented in Nash equilibrium. Next, we turn our attention to implementation in subgame perfect equilibrium through extensive form mechanisms. While it is known from Moore and Repullo (1988) and Abreu and Sen (1990) that this can be done, we seek to do so through mechanisms that fit closely the very description of the corresponding bargaining set. As such, our paper is a contribution to the Nash program for coalitional games. However, we shall insist on this program as being part of the general theory of implementation. In particular, we shall design mechanisms that an uninformed designer (about the consumers' preferences) could use to implement the bargaining sets.

Our work connects with two branches of the literature. First, it relates to the implementation of solutions to exchange economies and general coalitional games in characteristic function form (see, for example, Gul (1989), Hart and Mas-Colell (1996), Hurwicz (1979), Hurwicz, *et al.* (1995), Lagunoff (1994), Perry and Reny (1994), Schmeidler (1980), Serrano (1993, 1995, 1997), Serrano and Vohra (1997), Thomson and Tadenuma (1995)). And second, it links to the literature that proposes non-cooperative procedures to explore stable outcomes in coalitional bargaining situations (see, for example, Bloch (1996), Chatterjee *et al.* (1993), Ray and Vohra (1999)). Our paper contributes to both, by first providing a "bargaining-based" implementation of bargaining sets, and second, by uncovering an interesting set of procedures that deserve further analysis. These are negotiation rules in which proposals, very far from being "cheap talk," must be carefully thought out, as even if rejected at first, they may end up being the final outcome. A plausible setup where this feature may be present is one

⁴The consistent bargaining set, introduced by Dutta *et al.* (1989), tests also for the "credibility" of the counterobjections. That is, unlike the bargaining sets, which stop these chains of objections at the second level, the consistent bargaining set allows for chains of arbitrary length. The result is a set that lies between the core and the bargaining set.

⁵This set is sometimes referred to as the Aumann–Maschler bargaining set. However, as Maschler (1992, Footnote 15) points out, it is actually due to Davis and Maschler. We will refer to it in this paper as the Aumann–Davis–Maschler (ADM) bargaining set. We thank R. Aumann and M. Maschler for clarifying this point.

in which it is very costly to come up with new proposals (for legal or other frictional reasons).

We introduce an extensive form mechanism that exactly implements the ADM bargaining set correspondence in subgame perfect equilibrium. While our mechanism is closely related to the description of the ADM bargaining set it implements, there are subtleties in the construction of the non-cooperative game which indicate, not surprisingly, that precise details of the mechanism do matter. To the extent that our result provides non-cooperative foundations for the ADM bargaining set, this suggests that these foundations do depend crucially on such details. In any event, we hope that they will provide a better understanding of the bargaining set.⁶

To be faithful to the spirit of implementation theory, our mechanism has the property that outcomes are feasible allocations both in and out of equilibrium. As it turns out, this is not a trivial problem. Recall the description of the bargaining sets outlined above. Suppose that (S, y) is an objection by coalition S to an allocation x . If a counterobjection is raised, by coalition T with an allocation z , it is made to the vector of bundles $(y_{T \cap S}, x_{T \setminus S})$. However, no terminal node of the mechanism should specify this collection of bundles, as in general it will violate feasibility. The only previous papers we are aware of that ask the same question as ours are Einy and Wettstein (1999) and Perez-Castrillo and Wettstein (2000). The paper of Einy and Wettstein (1999) has the undesirable feature that their mechanisms do not respect the feasibility constraints of the economy for some off-equilibrium message profiles. Perez-Castrillo and Wettstein (2000), following Perez-Castrillo (1994), take the approach of adding extra players in the non-cooperative game (the “principals”) to the agents of the economy. The “principals” become residual claimants of the enterprise and may have losses out of equilibrium.

We solve the problem by constructing a multi-stage bargaining game, where the main novelties are two. (1) Depending on the identity of the rejector to a proposal, a different outcome is specified, and (2) after a proposal is put on the table and all the responders agree, the proposer has one extra chance to either sign on to the proposal or reject it. While feature (1) is key to solving the feasibility difficulty outlined above, feature (2) takes care of the related problem of frivolous proposals or responses (frivolous when judged by the logic of the bargaining set).

⁶In Section 5 we report that a modification of this mechanism can be used to implement the Mas-Colell bargaining set. The precise mechanism and the corresponding theorem can be found in Serrano and Vohra (2001).

2. PRELIMINARIES AND DEFINITIONS

Suppose there are l commodities and a set of consumers $N = \{1, \dots, n\}$. A pure exchange economy E is defined as $E = \{(X_i, u_i, \omega_i)_{i \in N}\}$, where $X_i \subseteq \mathbf{R}^l$, $u_i : X_i \mapsto \mathbf{R}$, and $\omega_i \in X_i$ refer to consumer i 's consumption set, utility function, and endowment, respectively. We shall assume that for every $i \in N$, $\omega_i > 0$. We will use the convention $\gg, >, \geq$ to order vectors.

Let \mathcal{N} denote the set of all non-empty subsets (coalitions) in N . For $S \in \mathcal{N}$, we use $-S$ to denote the complement of S . Given a collection of vectors or sets, one for each consumer, we will use subscripts to refer to their restrictions to a particular coalition. For example, $X_N = \prod_{i \in N} X_i$, $X_S = \prod_{i \in S} X_i$, and given $(x_i) \in X_N$, $x_S = (x_i)_{i \in S}$ and $x_{-S} = (x_i)_{i \notin S}$. We will denote by $u_S(x_S)$ the profile of utilities $(u_i(x_i))_{i \in S}$. For the grand coalition we will use $u(x)$ to denote $u_N(x_N)$.

We shall assume that the designer knows the endowments of the consumers but not their utility functions. Thus X_i and ω_i will remain fixed and economies will be distinguished simply by the utility functions of the consumers. Let \mathcal{E} denote the class of economies in which for all $i \in N$, the following assumptions are made:

- A1. $\forall i \in N$; $u_i(\cdot)$ is strictly monotonic, in the sense that $u_i(x_i) > u_i(x'_i)$ if $x_i > x'_i$.
- A2. $\forall i \in N$; $u_i(\cdot)$ is continuous.

Each coalition S has a feasible set of consumption plans, $A_S = \{x \in X_S : \sum_{i \in S} x_i \leq \sum_{i \in S} \omega_i\}$.

An allocation $x \in A_N$ is *efficient* if there does not exist $x' \in A_N$ such that $u(x') > u(x)$. An allocation $x \in A_N$ is *individually rational* if $u(x) \geq u(\omega)$.

Given $x \in A_N$, an *objection of i against j* is a pair (S, y_S) satisfying:

- (i) $i \in S, j \notin S, y_S \in A_S$;
 (ii) $u_S(y_S) \gg u_S(x_S)$.

Given $x \in A_N$ and an objection (S, y_S) of i against j , a *counterobjection by j* is a pair (T, z_T) satisfying:

- (i) $j \in T, i \notin T, z_T \in A_T$;
 (ii) $u_k(z_k) \geq u_k(y_k) \forall k \in S \cap T$;
 (iii) $u_k(z_k) \geq u_k(x_k) \forall k \in T \setminus S$.

We say that an objection of i against j is *justified* if there does not exist a counterobjection. The following definition was hinted at in Aumann and Maschler (1964) and introduced in Davis and Maschler (1963).

The (ADM) *bargaining set* $\mathcal{M}_1^i(E)$ of an economy E is the set of efficient and individually rational allocations to which there does not exist a justified objection by any i against any j .

Following the standard notion of objections used in the definition of the core, without reference to particular agents, we have the following notion of objections and counterobjections.

Given $x \in A_N$, an *objection* is a pair (S, y_S) satisfying:

- (i) $y_S \in A_S$;
- (ii) $u_S(y_S) > u_S(x_S)$.

Given $x \in A_N$ and an objection (S, y_S) , a *counterobjection* is a pair (T, z_T) satisfying:

- (i) $z_T \in A_T$;
- (ii) $u_k(z_k) > u_k(y_k) \forall k \in S \cap T$;
- (iii) $u_k(z_k) > u_k(x_k) \forall k \in T \setminus S$.

We say that an objection is *justified* if there does not exist a counterobjection to it. The following definition was introduced in Mas-Colell (1989).⁷

The (Mas-Colell) *bargaining set* $B(E)$ of an economy E is the set of efficient and individually rational allocations against which there does not exist a justified objection.

It is easy to see that given the monotonicity of preferences, the strict inequality, \gg , in the definition of a counterobjection can be substituted for the inequality $>$. It is, however, important that counterobjections are defined with at least one inequality in $B(E)$ but not in $\mathcal{M}_1^i(E)$.

3. NASH IMPLEMENTABILITY OF BARGAINING SETS

In this section we show that, unlike the core correspondence, the ADM bargaining set correspondence and the Mas-Colell bargaining set correspondence cannot be implemented in Nash equilibrium in pure exchange economies. The same is true for the consistent bargaining set of Dutta *et al.* (1989), which is a refinement of the Mas-Colell bargaining set taking account of longer chains of objections to counterobjections. We shall provide two examples of exchange economies to demonstrate that the three bargaining set correspondences do not satisfy Maskin's monotonicity condition (see Maskin (1977, 1985)) and, therefore cannot be implemented in Nash equilibrium.

⁷We abuse language slightly, as Mas-Colell (1989) did not require efficiency or individual rationality, which made his result the more surprising. Both requirements appear in Vohra (1991).

EXAMPLE 1. Consider the economy E consisting of three identical consumers, two commodities, and non-convex preferences. Each consumer has $u_i(x_{1i}, x_{2i}) = x_{1i}^2 + x_{2i}^2$ and $\omega_i = (1, 1)$. The core of this economy is empty. However, we claim that the allocation x , where

$$x_1 = (1.76, 0) \quad x_2 = (0, 1.76) \quad x_3 = (1.24, 1.24)$$

belongs to the bargaining set. The utilities corresponding to this allocation are

$$u_1 = u_2 = 3.0976 \quad u_3 = 3.0752.$$

Clearly, this allocation is individually rational. It is easy to check that it is efficient. Since any two-player coalition can obtain a utility of 4 for each member, x does not belong to the core. The objecting coalitions are all the three two-player coalitions. Every such objection has a counterobjection. In fact, this claim applies to the definitions of objections and counterobjections in accordance with all the three bargaining sets.

Now consider economy E' , which differs from E only in terms of consumer 3's utility function. In E' , consumer 3's utility function u' is given by

$$u'_3 = \min(x_{31}, x_{32}).$$

In the two-player coalition of players 1 and 3 in E' , all Pareto optimal allocations are along the diagonal. If player 3 gets $(1.24, 1.24)$, player 1 can get a maximum utility of $2 \times 0.76^2 = 1.15$. Thus there does not exist an objection to x from coalition $\{1, 3\}$ or from coalition $\{2, 3\}$; the only objection to x in E' is from coalition $\{1, 2\}$, which implies that x does not belong to any of the bargaining sets in E' . Notice that the change in the utility function of consumer 3 conforms to the hypothesis of Maskin monotonicity; i.e., the relevant lower contour set of consumer 3 is strictly larger in economy E' . This proves that the bargaining set is not monotonic and cannot, therefore, be implemented in Nash equilibrium. Again, this observation applies to all three bargaining sets.

It is natural to ask if a similar example can be constructed in a convex exchange economy. The answer, as we shall see, is yes. However, any such example must have more than three consumers. This is because in three-consumer convex exchange economies the core coincides with the bargaining set and since the core correspondence is monotonic, so is the bargaining set correspondence.⁸

⁸It is easy to see that in a three-player economy, any imputation that belongs to the bargaining set but not to the core must have objections from all two-player coalitions. Since that constitutes a balanced collection, this is impossible in a balanced game.

Instead of explicitly constructing an exchange economy, we will find it convenient to begin with a characteristic function game that is totally balanced.

EXAMPLE 2. Consider the six-player totally balanced TU game (N, v) , where $A = \{1, 2, 3\}$, $B = \{4, 5, 6\}$, $N = A \cup B$, and the characteristic function is the minimal superadditive cover compatible with the following:

$$v(S) = \begin{cases} 0 & \text{if } |S| = 1 \\ 1 & \text{if } |S \cap A| = |S \cap B| = 1 \\ 2 & \text{if } |S \cap A| = 1 \text{ and } |S \cap B| = 2. \end{cases}$$

The only imputation in the core is $(0, 0, 0, 1, 1, 1)$. However, the imputation $w = (1, 1, 1, 0, 0, 0)$ belongs to the ADM, the Mas-Colell, and the consistent bargaining sets (defined in the utility space). The objecting coalitions are all three-player coalitions with one player from A and two players from B and four-player and five-player coalitions with three players from B .

Now consider the game (N, v') where v' is the minimal superadditive cover compatible with

$$v'(S) = \begin{cases} 0 & \text{if } |S| = 1 \\ 1 & \text{if } |S \cap A| = |S \cap B| = 1 \\ 2 & \text{if } S = \{1, 4, 5\}. \end{cases}$$

This game is also totally balanced. Now the coalitions $\{1, 4, 5\}$, $\{1, 4, 5, 6\}$, $\{1, 2, 4, 5, 6\}$, and $\{1, 3, 4, 5, 6\}$ are the only coalitions with objections to w . It is easy to see that (S, w') is a justified objection in the sense of $B(E)$, where $S = \{1, 4, 5\}$ and $w'_1 = 1.5$, $w'_4 = w'_5 = 0.25$. This is also a valid objection in the sense of the consistent bargaining set and a justified objection by 1 against 6 in the sense of $\mathcal{M}_1^i(E)$. Thus w belongs to each of the bargaining sets in the game (N, v) but not in the game (N, v') .

Following Shapley and Shubik (1969), we can associate with both these games exchange economies, E and E' , in which all agents have identical concave and homogeneous of degree one utility functions with the following additional properties.⁹ The utility functions are u and u' and there is an allocation x such that $u(x) = u'(x) = w = (1, 1, 1, 0, 0, 0)$, while for all allocations $x' \neq x$, $u'(x) \leq u(x)$. This means that lower contour sets in E'

⁹An explicit construction of exchange economies corresponding to the above characteristic functions can be done as follows. Suppose each consumer i has an endowment of one indivisible unit of commodity i and, in addition to these six commodities, there is another commodity (money) which can be used to transfer utility. Let the utility of each player corresponding to ω_S be $v(S)$ in economy E and $v'(S)$ in economy E' . Finally, notice that we can consider divisible goods by extending the utility functions accordingly to the entire consumption set, maintaining the desired properties.

contain those of E , at the allocation x . Since x belongs to the bargaining set of E but not to the bargaining set of E' , Maskin monotonicity is violated. Thus, we have an example of a convex exchange economy in which none of the bargaining sets can be implemented in Nash equilibrium.

Although exact Nash implementation of a bargaining set is not possible, positive results can be obtained by considering subgame perfect implementation through extensive form mechanisms. In the next section we provide such a mechanism to implement the ADM bargaining set. A variant of this mechanism can be used to implement the Mas-Colell bargaining set, as reported in Section 5; see Serrano and Vohra (2001) for details. One advantage of these mechanisms is that they come close to the story underlying the bargaining sets. In general, they can be thought of as multi-stage bargaining games with the following salient features. There is a pre-bargaining stage (the “conversation stage”), where a status quo is agreed upon before further negotiations. This stage also determines the protocol in further negotiations. In the next stage, an agent makes a proposal to a coalition (a potential “objection”), to be sequentially ratified. A veto results in the status quo while unanimous acceptance leads to the next stage in which potential “counterobjections” are possible. A failed counterobjection imposes a penalty on the proposer. In equilibrium, all agents announce the same allocation, which is in the bargaining set, and the first proposal involves this status quo allocation proposed to the grand coalition. These mechanisms can be viewed as adding an additional stage to the extensive form mechanism constructed in Serrano and Vohra (1997) for implementing the core correspondence.

4. IMPLEMENTATION OF THE ADM BARGAINING SET

We shall construct an extensive form mechanism such that for every economy in \mathcal{E} , the subgame perfect outcomes of the game coincide with the ADM bargaining set allocations.

An *extensive game form or mechanism* is defined as a game tree with possibly simultaneous moves, i.e., as an array $\Gamma = (N, K, g)$, where N is the set of players, K is a game tree, and $g : Z \mapsto A_N$ is the outcome function, where Z denotes the set of terminal nodes of the tree K . We will use $g_i(z)$ to denote consumer i 's commodity bundle corresponding to the allocation $g(z)$. The set of nodes of the tree K is denoted T . The initial node is t_0 . Let M_i^t denote the set of choices available to player i at node t and let M_i denote the set of strategies of player i . Given an economy $E = \{(X_i, u_i, \omega_i)_{i \in N}\}$, the mechanism Γ defines an extensive form game (Γ, E) , where the payoff to the players corresponding to the strategy profile m is $u(g(m))$.

A *subgame perfect* equilibrium of a game (Γ, E) is a strategy profile $\bar{m} \in M_N$ such that the restriction of the strategies to every subgame constitutes a Nash equilibrium in the subgame. Let $\text{SPE}(\Gamma, E)$ denote the set of all allocations corresponding to subgame perfect equilibria of the game (Γ, E) .

A mechanism in extensive form Γ is said to *implement in subgame perfect equilibrium* the bargaining set in all economies over the class \mathcal{E} if $\text{SPE}(\Gamma, E) = \mathcal{M}_1^i(E)$ for all $E \in \mathcal{E}$.

We need some additional notation before defining our mechanism. For every i , pick $\epsilon_i \in \mathbf{R}_+^l$ such that $\omega_i - \epsilon_i \in \mathbf{R}_+^l$. This is possible since $\omega_i > 0$ for all i . Let Π denote the set of all permutations of N , i.e., one-to-one functions from N to N . Given $\pi = (\pi_i)$, where $\pi_i \in \Pi$ for every $i \in N$, define $p(\pi)$ to be the composition of the permutations (π_i) , i.e., $p(\pi) = \pi_1(\pi_2(\cdots(\pi_i(\cdots\pi_n)\cdots))$. The i th element of $p(\pi)$ will be denoted $p(\pi)_i$. Notice that for every $i \in N$, given π_{-i} and $\pi^* \in \Pi$, there exists $\pi'_i \in \Pi$ such that $p(\pi'_i, \pi_{-i}) = \pi^*$. In particular, any $i \in N$ can make a unilateral change in π_i to make himself or herself the first player in the order p . We shall interpret $p(\pi)$ to be an endogenously determined protocol in our extensive form game.

We shall construct a mechanism in extensive form consisting of four stages.

Stage 0

Every player i chooses simultaneously from the choice set $M_i^0 = A_N \times \Pi$. A typical choice of player i will be denoted by $m_i^0 = (x^i, \pi_i)$. Note that x^i refers to player i 's announcement of an allocation; i.e., $x^i = (x_j^i)_{j \in N}$. We will generally use superscripts to denote an agent's announcement of a profile.

Let $m^0 = (m_i^0)$ represent the profile of stage 0 messages and let $1(m^0) = p(\pi)_1$ and $n(m^0) = p(\pi)_n$ denote the first and the last players according to the order $p(\pi)$.

If for any i and j , $x^i \neq x^j$, the outcome is that player $n(m^0)$ receives $\omega_{n(m^0)} - \epsilon_{n(m^0)}$ and all other players receive their initial endowments. If $x^i = x^j = x^*$ for all i and j in N , proceed to stage 1. In this case we will refer to x^* as the status quo.

Stage 1

Player $i = 1(m^0)$ chooses a coalition S containing i , a player $j \notin S$ (if $S \neq N$), and $y \in A_S$. Let $S = \{1_S(m^0), 2_S(m^0), \dots, k_S(m^0)\}$, where $j_S(m^0)$ refers to the j th player in S according to the order $p(\pi)$. Of course, $1_S(m^0) = 1(m^0)$. After player i 's proposal, responses occur as follows: first the other members of S respond sequentially to the proposal (starting with player $2_S(m^0)$ and going up to $k_S(m^0)$) by either accepting it or rejecting it. If any player in S rejects the proposal, the final outcome is x^* . If all

members of S accept y , player j (if $S \neq N$) responds by either accepting or rejecting. If j accepts the proposal (or j does not exist because $S = N$), coalition S is assigned y and all players not in S are assigned their initial endowments. If player j rejects the proposal play moves to stage 2.

Stage 2

Player j names a coalition T containing at least two players ($j \in T$ and $i \notin T$) and a proposal $z \in A_T$. Then, according to the protocol induced on $T \setminus \{j\}$ by the message profile of stage 0, the other players in T respond sequentially. They may accept or reject the proposal of player j . If any player in $T \setminus \{j\}$ rejects the proposal, player j receives 0. The outcome for the other players depends on the following two cases: (a) If the first rejector belongs to $S \cap T$, the final outcome is $(y, 0_j, (\omega_k)_{k \notin S \cup \{j\}})$. (b) If the first rejector belongs to $T \setminus S$, the final outcome is $(x_{-j}^*, 0_j)$. If all players in $T \setminus \{j\}$ accept the proposal z , the game moves to stage 3.

Stage 3

In this stage, player j makes the final choice whether or not to accept the counterproposal. If he accepts, the final outcome is (z, ω_{-T}) . If he rejects, the final outcome is $(x_j^*, (x_k^*(\epsilon))_{k \neq j})$, where $x_k^*(\epsilon) < x_k^*$.

Remark 1. The outcome of each strategy profile in this mechanism is feasible, in and out of equilibrium. Notice also that stage 0 is designed to ensure that all agents coordinate on a particular status quo, since the lack of such coordination imposes a penalty on the last player, and this can be avoided by a suitable change in the permutation announced by this player. This feature is similar to the use of “integer games” in the implementation literature.

Remark 2. Although as written the mechanism Γ prescribes some outcomes where resources are wasted, this can be avoided when there are at least three agents. Following a rejection by agent k of a counterproposal made by j , it suffices to assign to agent $i \neq j, k$ the bundle x_j^* or z_j (in addition to the bundles awarded to i in the corresponding outcomes of the last stage of Γ). And for outcomes ending in stage 0, ϵ taken from the last player is awarded to some other agent, i .

THEOREM 1. *The extensive form mechanism Γ implements in subgame perfect equilibrium the ADM bargaining set $\mathcal{M}_1^i(E)$ in the class of economies \mathcal{E} .*

Proof. We begin by showing that if $x^* \in \mathcal{M}_1^i(E)$, then $x^* \in \text{SPE}(\Gamma, E)$. Consider a strategy profile \bar{m} defined as follows:

(i) $\bar{m}_i^0 = (x^*, \pi^e)$ for all i , where π^e denotes the identity permutation.

(ii) Consider a subgame following a status quo y^* agreed upon in stage 0. If $y^* \in \mathcal{M}_1^i(\mathbf{E})$, every player i proposes $(S, y, j) = (N, y^*)$ at every node of stage 1 where he has to make a proposal. If $y^* \notin \mathcal{M}_1^i(\mathbf{E})$, every player i makes a proposal that, given the continuation strategies, maximizes his payoff.

(iii) Consider a subgame following a status quo y^* agreed upon in stage 0 and a proposal (S, y, j) made by player i . Player $k \in S$ accepts if and only if $u_k(t_k) \geq u_k(y_k^*)$, where t_k is the unique bundle that k receives in the continuation equilibrium after $S \setminus \{i\}$ accepts (see below).

(iv) Consider a subgame following a status quo y^* agreed upon in stage 0 and a stage 1 proposal (S, y, j) made by player i and accepted by $S \setminus \{i\}$. Player j rejects it if $y^* \in \mathcal{M}_1^i(\mathbf{E})$, (S, y) is an objection to y^* , and he has a counterobjection. Then he proposes one of the best (from his point of view) counterobjections (T, z) such that $T \neq \{j\}$.¹⁰ Otherwise, player j accepts or rejects the stage 1 proposal on the basis of backwards induction. If he rejects, he makes one of the best offers from his point of view out of the set of counterproposals (T, z) that will be accepted and that he prefers to his endowment.

(v) In every subgame in stage 2 following a status quo y^* agreed upon in stage 0, a proposal (S, y, j) , and a counterproposal (T, z) , respondent k responds as follows.

Suppose that all respondents following k , if any, accept the proposal. If $k \in T \cap S$, k accepts if and only if either one of the following conditions holds,

$$u_j(z_j) \geq u_j(y_j^*) \quad \text{and} \quad u_k(z_k) \geq u_k(y_k)$$

or

$$u_j(z_j) < u_j(y_j^*) \quad \text{and} \quad u_k(y_k) \leq u_k(y_k^*(\epsilon)).$$

If $k \in T \setminus S$, k accepts if and only if

$$u_j(z_j) \geq u_j(y_j^*) \quad \text{and} \quad u_k(z_k) \geq u_k(y_k^*).$$

¹⁰This relies on the fact that when j has a counterobjection, j also has one with some other player besides j . To see this, suppose that the only counterobjection that j has against i 's objection is as a singleton. Then, given that $y^* \in \mathcal{M}_1^i(\mathbf{E})$, it must be the case that $u_j(y_j^*) = u_j(\omega_j)$. If $S \cup \{j\} = N$, we would have a contradiction to the efficiency of y^* . So there must exist $k \in N \setminus (S \cup \{j\})$. Consider (S, y) as an objection of i against k . Since $y^* \in \mathcal{M}_1^i(\mathbf{E})$, there exists a counterobjection (T', z') by k . Of course, if $j \in T'$, this also serves as a counterobjection by j in which $T' \neq \{j\}$. If $j \notin T'$, then, given that $u_j(y_j^*) = u_j(\omega_j)$, it follows that $(T' \cup \{j\}, (z', \omega_j))$ is a counterobjection by player j through a coalition that is larger than $\{j\}$.

Suppose that k' is the first respondent following k who rejects the counterproposal. If $k \in T \cap S$, then k accepts if and only if $k' \in T \cap S$. If $k \in T \setminus S$, then k accepts if and only if $k' \in T \setminus S$.

This completes the description of all responses to a counterproposal in stage 2.

(vi) In stage 3, following a status quo y^* agreed upon in stage 0, a proposal (S, y, j) , and a counterproposal (T, z) made by j and accepted by $T \setminus \{j\}$, player j accepts if and only if $u_j(y_j^*) \leq u_j(z_j)$.

Obviously, (vi) is consistent with subgame perfection since stage 3 is a final stage with a one-person decision problem. By backward induction, it is also easy to see that (iii), (iv), and (v) are consistent with subgame perfection. Continuity of the utility functions and the fact that the strategies used by responders involve weak inequalities guarantee the existence of best responses in some of these subgames.

To see that (ii) corresponds to an equilibrium, it is clear this is the case if $y^* \notin \mathcal{M}_1^i(E)$. Again the observation at the end of the previous paragraph applies.

Suppose then that $y^* \in \mathcal{M}_1^i(E)$ and consider a deviation by player i . For this to be profitable it must mean that i makes the proposal (S, y, j) , which is accepted by all members of S as well as by j (if the game were to continue beyond stage 1, this would give player i either his endowment or the bundle $y_i^*(\epsilon)$). This clearly requires (S, y, j) to be an objection. However, since $y^* \in \mathcal{M}_1^i(E)$, there exists a counterobjection to this proposal. And by (iv), (v), and (vi), player j will reject (S, y, j) and make a counterproposal (which will be accepted). This contradicts that (S, y, j) is accepted by j , as we had established.

It is easy to see that (i) corresponds to best responses by all players in stage 0. It follows, therefore, that this profile constitutes a subgame perfect equilibrium whose outcome is $g(\bar{m}) = x^*$ and $x^* \in \text{SPE}(\Gamma, E)$.

We now proceed to show that if \bar{m} is a subgame perfect equilibrium of (Γ, E) , then $g(\bar{m}) \in \mathcal{M}_1^i(E)$. Consider a subgame perfect equilibrium \bar{m} such that $\bar{m}_i^0 = (x^i, \pi_i)$. Let $\bar{z} = g(\bar{m})$.

Claim 1.1. For all $i, j \in N$, $x^i = x^j \equiv x^*$.

Suppose that this is not so. Then player $j = n(\bar{m}^0)$ receives $\omega_j - \epsilon_j$. However, this player can gain by changing π_j to π'_j such that $j \neq p(\pi_{-j}, \pi'_j)_n$. A deviation from \bar{m}_j with just such a change in π_j will result in j receiving ω_j instead of $\omega_j - \epsilon_j$, which contradicts the hypothesis that \bar{m} is a subgame perfect equilibrium.

A consequence of the proof of Claim 1.1 is that any SPE outcome must be individually rational.

Claim 1.2. $u(\bar{z}) \geq u(x^*)$.

Notice that given π_{-j} , by a suitable choice of π'_j , player j can make sure that $j = p(\pi_{-j}, \pi'_j)_1$. Suppose that by such a choice j becomes the first player in the order p and then proposes (N, x^*) in stage 1. Irrespective of how the others respond, the outcome is x^* . Thus every player can make a unilateral deviation to ensure the status quo x^* . Since $\bar{z} \in \text{SPE}(\Gamma, E)$, the claim follows.

A consequence of Claim 1.2 and its proof is that any SPE outcome must be efficient.

Claim 1.3. $u(\bar{z}) = u(x^*)$.

Suppose that this is not so. Then, by Claim 1.2, $u(\bar{z}) > u(x^*)$. By monotonicity of preferences, there must exist $z' \in A_N$ and $i \in N$ such that $u(z') \gg u(x^*)$ and $u_i(z'_i) > u_i(\bar{z}_i)$. Suppose that i changes her strategy to become the first player according to the protocol and proposes (N, z') . This proposal will be accepted by all the other players since $u_j(z'_j) > u_j(x^*_j)$ for all $j \in N$ and is, therefore, a profitable deviation for player i . But this contradicts the hypothesis that \bar{z} is a SPE outcome.

The following claim regarding stage 3 is obvious from backwards induction.

Claim 1.4. Consider a subgame in stage 3 following a proposal (S, y, j) and a counterproposal (T, z) . (The proposal must have been accepted by S and the counterproposal by T for stage 3 to occur.) If $u_j(x^*_j) > u_j(z_j)$, then in stage 3, player j must reject, and the equilibrium outcome in this subgame must be $(x^*_j, (x^*_k(\epsilon))_{k \neq j})$.

To complete the proof, suppose that $\bar{z} \notin \mathcal{M}_1^i(E)$. Then there exists a justified objection (S, y) to \bar{z} (or, equivalently, to x^*) by i against j . Let player i deviate from the equilibrium strategies by changing his permutation in stage 0 so as to become the proposer in stage 1. Let player i announce (S, y, j) in stage 1. If this proposal is accepted by j , clearly all members of S would accept it, and player i would have a profitable deviation. But this contradicts the hypothesis that \bar{m} is a SPE. Thus, to complete the proof that $\bar{z} \in \mathcal{M}_1^i(E)$ we prove the following claim.

Claim 1.5. Suppose that (S, y) is a justified objection to \bar{z} made by i against j . Further, suppose that x^* is the status quo agreed upon in stage 0, $u(\bar{z}) = u(x^*)$, and that player i has made the proposal (S, y, j) in stage 1. Then player j 's strategy according to \bar{m} is to accept the proposal (S, y, j) .

Suppose that the claim is false. Then j counters with a proposal (T, z) . We will now show that this counterproposal will not be accepted by all the responders. This will prove that player j must accept (S, y, j) in a subgame perfect equilibrium, a contradiction to our first assumption. Recall that

there exists a player $k \in T$, $k \neq j$. There are now two cases to consider:

(a) there exists $k \neq j$, $k \in T \setminus S$ such that $u_k(x_k^*) > u_k(z_k)$ or there exists $k \neq j$, $k \in S \cap T$ such that $u_k(y_k) > u_k(z_k)$.

Let k be the last responder for whom the above condition holds. Suppose that the counterproposal is accepted by all the responders. Then the final outcome for player k is either z_k or $x_k^*(\epsilon)$. In either case, player k can do better by rejecting the proposal. This proves that the proposal will not be accepted by all the players in $T \setminus \{j\}$.

(b) For all $k \neq j$, $k \in T \setminus S$, $u_k(x_k^*) \leq u_k(z_k)$ and for all $k \neq j$, $k \in S \cap T$, $u_k(y_k) \leq u_k(z_k)$. Since (S, y) is a justified objection of i against j , this must mean that $u_j(z_j) < u_j(x_j^*)$. But then, we know from Claim 1.4 that in stage 3, the final outcome must be $(x_j^*, (x_k^*(\epsilon))_{k \neq j})$. By rejecting the counterproposal, any $k \in T$, $k \neq j$ can obtain either y_k or x_k^* , depending on whether $k \in S$ or $k \notin S$. In either case, rejecting the proposal is better than proceeding to stage 3 and receiving $x_k^*(\epsilon)$. Thus the counterproposal is not accepted by all members of T .

We have shown that according to the equilibrium strategies in \bar{m} , a counterproposal by j will be rejected. And this will yield 0 to player j . On the other hand, accepting the proposal (S, y, j) allows player j to receive ω_j . Clearly then (S, y, j) will be accepted by j , and this completes the proof of Claim 1.5. ■

5. CONCLUDING REMARKS

The main difficulty involved in using the mechanism of the previous section (or any obvious modification thereof) to implement the Mas-Colell bargaining set stems from the way in which counterobjections are specified. Recall that a counterobjection used in defining $B(E)$ is required to make some member of the coalition strictly better off. Suppose that $x \notin B(E)$ and (S, y) is a justified objection to x . While there does not exist a counterobjection (in the sense of $B(E)$) to (S, y) , it is possible that there exists a coalition T and $z \in A_T$ such that $u_i(z_i) = u_i(y_i)$ for all $i \in S \cap T$ and $u_j(z_j) = u_j(x_j)$ for all $j \in T \setminus S$. In the kind of mechanism described in the previous section, x will be an equilibrium outcome, supported by strategies in which members of T accept the counterproposal (T, z) if the objection (S, y) is made, and the justified objection is, therefore, not made. It should be clear that imposing a small cost on the counterproposer can deter such frivolous counterproposals.¹¹

¹¹This problem can be avoided by modifying the definition of $B(E)$ to require counterobjections to hold with a weak inequality " \geq " instead of " \gg " (and imposing other restrictions on a counterobjection), as in Zhou (1994).

In Serrano and Vohra (2001), we formalize this idea by introducing an extensive form mechanism with discounting. There we show that for δ close to 1 (where $\delta \in (0, 1)$ denotes the common discount factor), this mechanism approximately implements the Mas-Colell bargaining set in subgame perfect equilibrium. The set of limit points of SPE outcomes as δ converges to 1 yields the closure of $B(E)$. Indeed, one cannot do better. While the SPE payoff correspondence is upper hemicontinuous at $\delta = 1$, the Mas-Colell bargaining set is not closed in general. To illustrate this latter point, consider a three-player transferable utility game (N, v) , where $v(N) = v(S) = 1$ for all two-player coalition S and $v(\{i\}) = 0$. The reader can check that $(1/2, 1/2, 0) \notin B(N, v)$, but for $\epsilon > 0$ small, $(1/2 - \epsilon, 1/2 - \epsilon, 2\epsilon) \in B(N, v)$.

The new features in the mechanism in Serrano and Vohra (2001) are two-fold. First, if a proposal is accepted by the responders, it takes until the next period for the proposer to ratify it or not. And second, although the first proposal may be made to subcoalitions, every player in N votes on it; recall that in defining $B(E)$ an objection is not directed against a particular player j . The rules also incorporate some asymmetries in treating responders to a proposal or a counterproposal. These are related to the asymmetries of the bargaining set in treating objections and counterobjections.

The implementation of the consistent bargaining set seems to be a more ambitious task, one which is still an important open problem for future research. The reason is that the game called for should not fix a finite horizon. An extension of the mechanism in Serrano and Vohra (2001) seems to take us part of the way, but its full analysis is far from being straightforward. To conclude, our results underscore the crucial importance of small details in the non-cooperative foundations of bargaining sets. This suggests as an alternative research agenda the exploration of natural variants of coalitional bargaining procedures. The insights obtained in their study will also add to our understanding of coalition formation models.

REFERENCES

- Abreu, D., and Sen, A. (1990). "Subgame Perfect Implementation: A Necessary and Almost Sufficient Condition," *J. Econ. Theory* **50**, 285–299.
- Aumann, R. J., and Maschler, M. (1964). "The Bargaining Set for Cooperative Games," in (M. Dresher, L. S. Shapley, and A. W. Tucker, Eds.), *Advances in Game Theory*, pp. 443–476. Princeton, NJ: Princeton Univ. Press.
- Bloch, F. (1996). "Sequential Formation of Coalitions with Fixed Payoff Division," *Games Econ. Behav.* **14**, 90–123.
- Chatterjee, K., Dutta, B., Ray, D., and Sengupta, K. (1993). "A Non-Cooperative Theory of Coalitional Bargaining," *Rev. Econ. Stud.* **60**, 463–477.

- Davis, M., and Maschler, M. (1963). "Existence of Stable Payoff Configurations for Cooperative Games: Abstract," *Bull. Amer. Math. Soc.* **69**, 106–108.
- Davis, M., and Maschler, M. (1967). "Existence of Stable Payoff Configurations for Cooperative Games," in Shubik, M. (Ed.), *Essays in Mathematical Economics in Honor of Oskar Morgenstern*, pp. 39–52. Princeton, NJ: Princeton Univ. Press.
- Dutta, B., Ray, D., Sengupta, K., and Vohra, R. (1989). "A Consistent Bargaining Set," *J. Econ. Theory* **49**, 93–112.
- Einy, E., and Wettstein, D. (1999). "A Non-Cooperative Interpretation of Bargaining Sets," *Rev. Econ. Design* **4**, 219–230.
- Gul, F. (1989). "Bargaining Foundations of the Shapley Value," *Econometrica* **57**, 81–95.
- Hart, S., and Mas-Colell, A. (1996). "Bargaining and Value," *Econometrica* **64**, 357–380.
- Hurwicz, L. (1979). "Outcome Functions Yielding Walrasian and Lindahl Allocations at Nash Equilibrium Points," *Rev. Econ. Stud.* **46**, 217–225.
- Hurwicz, L., Maskin, E., and Postlewaite, A. (1995). "Feasible Nash Implementation of Social Choice Rules When the Designer Does Not Know Endowments or Production Sets," in *The Economics of Informational Decentralization: Complexity, Efficiency and Stability*, (J. O. Ledyard Ed.), pp. 367–433, Amsterdam: Kluwer Academic.
- Lagunoff, R. (1994). "A Simple Noncooperative Core Story," *Games Econ. Behav.* **7**, 54–61.
- Maschler, M. (1992). "The Bargaining Set, Kernel and Nucleolus," in *Handbook of Game Theory with Economic Applications*, Vol. I, (R. J. Aumann and S. Hart, Eds.), pp. 591–667. New York: North-Holland.
- Mas-Colell, A. (1989). "An Equivalence Theorem for a Bargaining Set," *J. Math. Econ.* **18**, 129–139.
- Maskin, E. (1977). "Nash Equilibrium and Welfare Optimality," mimeo, MIT. (1999) *Rev. Econ. Stud.* **66**, 23–38.
- Maskin, E. (1985). "The Theory of Implementation in Nash Equilibrium: a Survey," in *Social Goals and Social Organization*, (L. Hurwicz, D. Schmeidler, and H. Sonnenschein, Eds.), Cambridge, UK: Cambridge Univ. Press.
- Moore, J., and Repullo, R. (1988). "Subgame Perfect Implementation," *Econometrica* **56**, 1191–1220.
- Perez-Castrillo, D. (1994). "Cooperative Outcomes through Non-Cooperative Games," *Games Econ. Behav.* **7**, 428–440.
- Perez-Castrillo, D., and Wettstein, D. (2000). "Implementation of Bargaining Sets via Simple Mechanisms," *Games Econ. Behav.* **31**, 106–120.
- Perry, M., and Reny, P. J. (1994). "A Non-Cooperative View of Coalition Formation and the Core," *Econometrica* **62**, 795–817.
- Ray, D., and Vohra, R. (1999). "A Theory of Endogenous Coalition Structures," *Games Econ. Behav.* **26**, 286–336.
- Schmeidler, D. (1980). "Walrasian Analysis via Strategic Outcome Functions," *Econometrica* **48**, 1585–1594.
- Serrano, R. (1993). "Non-Cooperative Implementation of the Nucleolus: The 3-Player Case," *Int. J. Game Theory* **22**, 345–357.
- Serrano, R. (1995). "A Market to Implement the Core," *J. Econ. Theory* **67**, 285–294.
- Serrano, R. (1997). "Reinterpreting the Kernel," *J. Econ. Theory* **77**, 58–80.
- Serrano, R., and Vohra, R. (1997). "Non-Cooperative Implementation of the Core," *Soc. Choice Welfare* **14**, 513–525.

- Serrano, R., and Vohra, R. (2001). "Implementing the Mas-Colell Bargaining Set," Working Paper No. 2001-24 Brown University.
- Shapley, L. S., and Shubik, M. (1969). "On Market Games," *J. Econ. Theory* **1**, 9-25.
- Thomson, W., and Tadenuma, K. (1995). "Games of Fair Division," *Games Econ. Behav.* **9**, 191-204.
- Vohra, R. (1991). "An Existence Theorem for a Bargaining Set," *J. Math. Econ.* **20**, 19-34.
- Zhou, L. (1994). "A New Bargaining Set of an N-Person Game and Endogenous Coalition Formation," *Games Econ. Behav.* **6**, 512-526.