ELSEVIER Journal of Mathematical Economics 23 (1994) 533–548

JOURNAL OF Mathematical ECONOMICS

Core implementation and increasing returns to scale for cooperation

Benny Moldovanu [a], Eyal Winter *,[b]

[a] *The Economics Department, The University of Bonn, Bonn, Germany*
[b] *The Economics Department and The Center for Rationality and Interactive Decision Theory, The Hebrew University of Jerusalem, Feldman Bldg., Gibat Ram, 91904 Jerusalem, Israel*

Submitted January 1993, accepted June 1993

Abstract

In this paper we analyze a simple non-cooperative bargaining model for coalition formation and payoff distribution for games in coalitional form. We show that under our bargaining regime a cooperative game is core-implementable if and only if it possesses the property of increasing returns to scale for cooperation, i.e. the game is convex. This offers a characterization of a purely cooperative notion by means of a non-cooperative foundation.

Keywords: Non-cooperative bargaining; Core implementation

JEL classification: C71; C78

1. Introduction

The core is a notion of collective stability. A core allocation or outcome is immune against deviations by coalitions. However, in non-cooperative and competitive environments it is not clear when and how can players be induced to play core outcomes. This is the question we wish to address in this paper. Collective

* Corresponding author. Part of this research was supported by the Deutscheforschungsgemeinschaft though the SFB 303 at the University of Bonn. The second author wishes to thank Reinhard Selten for his hospitality during the summer of 1992. We also thank Sergiu Hart Motty Perry and Reinhard Selten for many valuable discussions on this topic.

SSDI 0304-4068(94)E0077-9

decision making is often carried out through some procedure of multilateral bargaining. In general such bargaining games may yield equilibrium outcomes which are neither stable nor even efficient, as was indicated by a well known example due to Shaked (see Sutton (1986)). [1]

The motivation of constructing sequential bargaining games to sustain various cooperative solution concepts, and in particular the core, stems from the desire to explore the role of an arbitrator or a planner in multilateral situations. One of the consequences of our analysis is that in the framework and conditions of our core implementation no planer or arbitrator is needed in order to induce coalitional stable outcomes, since such outcomes can be directly sustained through a face to face (non-cooperative) bargaining.

We will use a simple multi-person sequential bargaining game for coalition formation, which is based on a cooperative game with side payments (in coalitional form). Our main objective here is to characterize the class of cooperative games for which the non-cooperative solution (subgame perfect equilibria) of the bargaining game necessarily yields outcomes which are coalitional stable, i.e, outcomes that are in the core of the underlying cooperative game. Such games will thus have the property that the bargaining behavior is not only stable with respect to unilateral deviations, but stability holds also with respect to multilateral deviations. i.e, no group of agents can be made better off by correlating a joint deviation. A cooperative game for which this property holds for all its restricted games will be said to be core implementable. Our main result asserts that a cooperative game is core implementable if and only if it is convex. [2] We thus provide a characterization of an important class of cooperative games by means of a non-cooperative notion.

The approach taken in this paper is in a way dual to that which is typically used in the literature on core implementation. Usually one restricts itself to a class of cooperative games, such as the class of totally balanced games (see Perry and Reny (1991) or Moldovanu and Winter (1991)), and then one introduces a non-cooperative bargaining model that implements the core within this class. We, alternatively, raise the following question: Which are the cooperative games that are core implementable under our bargaining regime?

Convex games have the property of increasing returns to scale for cooperation. This means that every player becomes more valuable as the coalition to which he joins grows. It is this property that made this class of games so much popular from the point of view of applications in economics. There is a wide range of interesting

[1] In Van Damme et al. (1990) it is shown that even in bilateral bargaining, when a smallest money unit exists inefficient subgame perfect equilibria may exist.

[2] The word convex in our context is perhaps misleading. Using the terminology 'games of increasing returns' or 'games of increasing individual contributions' could have been more informative. We however stick to the term convexity because this is the most common name for this class of games ever since Shapley (1971) introduced it.

exchange markets that give rise to a convex game (see Rosenmueller (1983) and Ichiishi (1992)). Public good consumption can typically be modeled as a game with increasing returns to scale for cooperation, and a situation of oligopoly can also be described as a convex game under some specifications on the demand functions of the consumers (Shubik (1987)).

Our main result implies that, in each of the environments described above, our non-cooperative bargaining model always yields outcomes which are core allocations, not only on the whole economy but also on every sub-economy. The example of a public good economy will be later more thoroughly discussed.

Without trying to give a comprehensive survey of the literature on non-cooperative foundations of cooperative game theory, the works which are closely related to ours, and in particular those which discussed core implementation, need to be mentioned. A pioneer attempt to investigate coalition formation in cooperative games by means of bargaining extensive form games is due to Selten (1981). Selten has used a bargaining model similar to ours to establish the relation between the set of semi-stable demand vectors (or aspiration vectors a la Bennett (1983)) of the cooperative game, and the equilibrium payoffs of the bargaining game. Later Chatterjee et al. (1993) examined a different model to explore the efficiency of multilateral bargaining outcomes. In Moldovanu and Winter (1991) and Winter (1992) the relation between bargaining mechanism robustness and order independence on one hand and the core on the other hand is established. Roughly, in the later paper it is shown that the only bargaining outcomes which are immune against procedural changes are core outcomes.

A paper which is related to ours is Perry and Reny (1991). Perry and Reny consider a bargaining model where time is a continuous variable, and where a distinction is made between the time an agreement is reached and the time utility is being consumed. They show that under certain restrictions on players strategies the core can be implemented by stationary subgame perfect equilibria on the domain of totally balanced games.

Our paper differs from Perry and Reny's (P&R) both in motivation and in the results. Their objective was to describe a bargaining game based on Edgeworth's justification for the core and provide a non-cooperative view of this solution concept. P&R's equivalence result is strongly dependent on the fact that they use a continuous time model. They indicate informally that their results cannot be obtained by using a discrete time framework. This is indeed the case when one considers their domain of totally balanced games. One thing we show in this paper is that this can be done when the considered domain is smaller, namely, the set of all convex games.

Our insisting on a discrete time model is driven by the fact that our approach is less descriptive than that of P&R. Our concern is more with practical implementation of the core. Simplicity thus plays a much more important role here. The model we use is in fact the simplest generalization of the two person alternating offer model. We impose no restriction on players strategies, but only on the equilibrium outcomes. From the point of view of mechanism design, we feel that this simplicity is an important distinction between our paper and other related papers in the field. Admittedly we pay back for this simplicity in terms of the fact that the core is implemented on a domain smaller than that of all totally balanced game, namely the domain of convex games but given the economic importance of this domain we feel that this is a cost worthwhile paying.

Another important distinction between our results and the remaining literature on core implementation, is that our intention is not only to implement the core but also to exactly identify the cooperative games which admit such simple core implementation. What we will show is that if the underlying game is not convex, then at least some stationary subgame perfect equilibria will yield outcomes which are not in the core. So the property of increasing returns to scale for cooperation is not only sufficient for guaranteeing coalitional stable bargaining outcomes, it is also necessary. Our results therefore not only characterize the core but they also provide a meaningful characterization for the class of convex games in terms of core implementability. This follows from a result, with some importance of its own, which characterizes the class of convex games in terms of the size of the cores of its restricted games. This is an interesting parallel to Shapley and Shubik's characterization of market games.

2. The model

Consider a set of n players $N = \{1, 2, \dots, n\}$ who are involved in some interaction which gives rise to utilities for all possible coalitions. Without specifying precisely the environment in which this interaction takes place, it is represented as a cooperative game in coalition function form $v{:}2^N \to R$. Thus for each S, $v(S)$ is the total utility that the coalition S can allocate among its members. We therefore assume that utility is transferable. We also assume that cooperation pays off in the strict sense. Formally, v is strictly super-additive, i.e. $V(S \cup T) > v(S) + v(T)$ for all S,T with $S \cap T = \emptyset$, and that $v(S) > 0$ for all S.

The cooperative game (C-game) is an exogenous element of the model, and the precise source of the utilities is not specified. However, unlike the classical approach to cooperative games, we do not assume that coalition formation and payoff distribution can be predicted based solely on the C-game v. We thus entertain a simple non-cooperative bargaining procedure upon which such predictions can be based. We first start with the informal description of this procedure, which is a simplified version of Selten's (1981) proposal model.

The bargaining starts by randomly choosing a player i from N who has to initiate a proposal. A proposing player submits publicly a proposal (S, x_S, j) which consists of: (1) The proposed coalition of agents S, which contains i; (2) a feasible payoff vector x_S for the players in S, and (3) a player j in $N \setminus i$ which has to respond to i's proposal. A responding player either accepts the proposal or rejects it. Upon acceptance the responder chooses a new responder among those in S who have not yet responded. If the responder rejects a proposal then this proposal is removed and the responder initiates a new proposal. An agreement in this bargaining game is a pair (S, x_S) where $x(S) = v(S)$.

We 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. Formally, the players have lexicographic preferences over agreements. Thus if (S, x_S) and (T, y_T) are two agreements, then every i in $S \cap T$ prefers (S, x_S) to (T, y_T) if $x_i > y_i$ or $x_i = y_i$ and $T \subsetneq S$. This assumption is merely used in order to cope with the technical problem of ties, which can appear if players are indifferent between different agreements yielding the same payoff. Such ties may induce inefficiencies which, intuitively, are not very reasonable. In the last section we will demonstrate by means of an example why such an assumption is needed.

Finally, the bargaining terminates when an agreement is reached. Each player of the formed coalition receives his bargained payoff, and the rest are assigned their individually rational payoffs (i.e. $v(i)$). If the bargaining never terminates all the players are assigned their individually rational payoffs. Thus the bargaining model described here fits best environments where no further bargaining takes place after the formation of the first coalition. This is often the nature of political bargaining, where only one winning coalition can form, but it also resembles situations where a single project is being auctioned among a group of competing constructors who may also cooperate for the purpose of winning the project. Nevertheless, this model is also applicable to situations where sequential formations of coalitions takes place, only that in these cases one should view the formation of each coalition as a separate game.

Each choice of the first proposer in the above bargaining model gives rise to a non-cooperative game (NC-game) in extensive form with perfect information. Our non-cooperative solution concept is stationary subgame perfect equilibrium in pure strategies. To define the notion of stationary strategies one needs to divide the set of decision points in the game tree into equivalence classes and then require that players behave similarly at every two decision points within the same class. Thus, in stationary strategies players behavior is history dependent only in a limited way. To be more specific, we will use the term position when we refer to these equivalence classes.

Take i in N. A proposal position for i, denoted by (i), is the set of all i's decision points in the game tree in which i has to initiate a proposal. The responder position (S, x_S, T) for i is the set of all decision points for i in which i has to respond to the proposal (S, x_S) after the players in $T \subset S$ have already accepted the proposal. A stationary strategy for i is a function which assigns a proposal by player i to every proposal position, and assigns a response, i.e., an element of the set {yes, no}, to every responder position (S, x_S, T). In a stationary strategy a player acts in the same way at any decision point which belongs to the same position. A stationary subgame perfect equilibrium (SSPE), is a subgame perfect equilibrium that uses only stationary strategies. Note that we do not, a-priori, restrict players' strategies in any way.

3. The core and the SSPE payoffs vectors

As a counterpart to the non-cooperative solution, we will focus on the core as the cooperative solution concept. If v is a C-game on N, the core of v, denoted by $C(v)$, is given by

$$C(v) = \{x \in R^n;\ x(S) \geq v(S) \forall S \subset N, \text{ and } x(N) = v(N)\},$$

where $x(S) = \sum_{i \in S} x_i$.

Let v be a C-game and let $G^{j,N}$ be the corresponding NC-game, when N is the set of players and j opens the bargaining (proposes first). Let b be an SSPE of $G^{j,N}$. For each i in N we denote by $E_i(b)$ the payoff for i when b is played, and $E(b) = (E_i(b))_{i \in N}$. We will use $E(G^{j,N})$ to denote the set of all SSPE payoff vectors of $G^{j,N}$. As was already mentioned our main interest in this paper is to explore the relation between the core payoffs of the C-game v, and the SSPE payoff vector of the corresponding NC-game.

To obtain an equivalence theorem in terms of this relation we need to consider not only the C-game v on the set of players N but also all of its restricted games [3]. For each coalition $S \subset N$ the restricted game v_S is the C-game on the set of player S given by $v_S(T) = v(T)$ for all $T \subset S$. We will denote by $G^{j,S}$ the NC-game induced by v_S, where $j \in S$ opens the game.

We can now define the notion of core implementability for C-games on which our result is based.

A C-game v is said to be *Core Implementable* if for all $S \subset N$, and $j \in S$, $E(G^{j,S}) = C(v^s)$. [4]

Thus, a core implementable C-game is one in which for every possible forum of negotiation $S \subset N$, the set of all the SSPE outcomes of the bargaining game coincides with the core of the relevant C-game. Note that we use the regular notion of subgame perfection but our notion of core implementation is somewhat stronger than the usual requirement in the sense that it imposes a condition not only on the C-game v but also on all of its restricted games. What we want to guarantee by this notion is that no group of players refusing to take part in the negotiation (or leaving with some pre-play agreement) will cause 'instability' on

[3] In our terminology we distinguish between restricted games which refer to cooperative games, and subgames which are used as usually for extensive form games.

[4] In the sequel we will allow ourselves to abuse notation by writing G^S when we refer to $G^{i,S}$ for some arbitrary player i in S.

the set of remaining players, i.e., some equilibria of the resulting C-game will yield non-core outcomes. Note, in particular, that in core implementable games there is no way also for a group of players to profitably deviate from an SSPE by jointly correlating an alternative behavior. Hence all SSPE are also strong equilibria in the sense of Aumann (1959).

We are now ready to state our main result.

Theorem 1. A C-game v is core implementable if and only if it is convex.

Convex games have the property of 'snowballing' cooperation. In such games players become more essential when they join large coalitions. Formally, it means that players marginal contribution to coalitions is monotone with respect to inclusions, i.e v is convex if for all $i \in N$ and T, $S \subset N \setminus i$ with $T \subset S$ we have, $v(S \cup i) - v(S) \geq v(T \cup i) - v(T)$.

We will in fact show more than the assertion claimed in Theorem 1. Our first lemma demonstrates that for any game with a non-empty core (not necessarily convex), every core outcome is sustainable by some stationary subgame perfect equilibrium of our bargaining game. Later we will use an auxiliary result about the relation between semi-stable vectors and the core to show that only in convex games all equilibrium outcomes are core outcomes.

Lemma 3.1. Let v be a C-game with a non-empty core. For each payoff vector x in the core of v there exists some SSPE b of the game G^N such that x is the equilibrium payoff of b.

Proof. Let x be some core outcome of v. Consider the following strategy combination:

(1) Each player i as a proposer proposes (N, x, j), where j is an arbitrary player in N.
(2) Each player i as a responder accepts any proposal which yields him a payoff greater than x_i, and the proposal (N, x). All other proposals are rejected by i.

Obviously the above strategy combination yields the outcome (N, x). It is thus left to show that this strategy combination is indeed and SSPE. Consider first proposal positions. Since x is in the core, any alternative proposal which yields i more than x_i must yield some other player, say j, less than x_j. Such a player must reject that proposal according to the specified behavior of the strategy combination. So no proposer can profitably deviate. We will now consider response positions. Obviously a player cannot accept a payoff less than x_i, because he is better off rejecting it and proposing (N, x) which will then be accepted. For the same reason no responder can accept x_i with a coalition that is a strict subset of N since he prefers (N, x). It is thus left to show that accepting (N, x) or any offer of more than x_i is a best response for i. Suppose first that i is the last player to respond. If i rejects the proposal then he becomes a proposer and he can get no more than x_i at the continuation of the game. Finally if i is not the last to respond the last assertion is obtained by backward induction on the number of players proceeding i in their response. Q.E.D.

Our next Lemma characterizes the equilibrium payoffs of players at positions in which they initiate proposals (see also Selten (1981)).

Lemma 3.2. For each SSPE b of G^N, and for each i in N, let $d_i(b)$ be i's payoff when b is played at any subgame that starts with i initiating a proposal. Let $d(S) = \sum_{i \in S} d_i(b)$ for $S \subset N$. Then:

(1) *For each $S \subset N$ $d(S) \geq v(S)$, and*
(2) *For every i in N there exists some $S \subset N$, such that i belongs to S and $d(S) = v(S)$.*

Proof. We first show (1). Suppose in negation that there exists some $S \subset N$ such that $d(S) < v(S)$. Let $\epsilon = [v(S) - d(S)]/|S|$, and consider the following behavior by i:

1. At a proposer position always propose $(S, d_S + \epsilon_S, j)$, where j is some arbitrary player in S.
2. At a responder position, stick to what b specifies.

We will show that i improves his payoff by adopting the behavior described above. To show this, it is enough to prove that i's proposal as specified above must be accepted. Indeed consider the behavior of the last responder to i's proposal, say j. If j rejects the proposal then, because of the stationarity of b, he is going to realize a payoff of $d_j(b)$ at the resulting subgame. If j however accepts the proposal then he earns $d_j(b) + \epsilon$. So j must accept the proposal. Using a backwards induction argument we can conclude that all the members in S must accept the proposal as well. We thus have shown (1).

To show (2) simply note that for any coalition that forms as a result of some SSPE, each player i must be paid at least d_i, otherwise i will be better off rejecting this proposal. So if $j \in N$ can guarantee d_j at the subgame starting with his proposal then there must be some coalition S for j where $\sum_{i \in S} d_i \leq v(S)$. Q.E.D.

Conditions (1) and (2) of Lemma 3.2 define a set of vectors known as the set of aspiration vectors or semi-stable vectors, which has been already discussed in the literature by both Albers (1979) and Bennett (1983), and was shown to be non-empty. Selten (1981) introduced a bargaining model similar to ours to establish the relation between this set of vectors and the equilibrium outcomes of the bargaining game. We will adopt Selten's terminology as this reference is the most relevant to our treatment. Thus, any vector which satisfies (1) and (2) is called a semi stable vector of v.

Denote by $X(v)$ the set of all semi-stable vectors of the C-game v, and define $F_i(x) := \{S \subset N;\ i \in S \text{ and } x(S) = v(S)\}$, we obtain $X(v) = \{x \in R^n;\ \forall i \in N\ F_i(v) \neq \emptyset, \text{ and } \forall S \subset N\ x(S) \geq v(S)\}$. Note that $C(v) \subset X(v)$.

Our next auxiliary result has an importance of its own. It, in fact, provides a characterization of convex games in terms of the relation between their core payoffs and the semi-stable payoff vectors. [5]

Proposition 3.3. Let v be a C-game on the set of player N. Then v is convex if and only if for every $S \subset N$, $C(v_s) = X(v_s)$. Namely, the core of the restricted game on S coincides with the set of semi-stable vectors of that restricted game.

There is a surprising relation between Proposition 3.3 and the famous Shapley and Shubik (1969) Theorem on market games. Shapley and Shubik have shown that a game is a market game if and only if each of its restricted games has a non-empty core. Proposition 3.3 asserts that a game is convex if and only if each of its restricted games has a 'large' core, where large means $X(v_S) = C(v_s)$.

Proposition 3.3 will follow immediately from the following two lemmas.

Lemma 3.4. If v is not convex then there exists some $S \subset N$ for which $C(v_S) \subsetneq X(v_s)$.

Proof. Let S be a minimal coalition for which v_S is not convex i.e., for every $T \subseteq S$ v_T is a convex game. (Note that a one-person game is always convex). Since v_S is not convex there exists $i \in S$, R, $T \subset S$ such that $i \in R \subset T$, with $v(T \cup i) - v(T) < v(R \cup i) - v(R)$. By the minimality of S we must have $S = T$. Otherwise v_T is also not convex contradicting S's minimality. Let us now take a permutation π of the players in S such that i is placed last. Without loss of generality rename the players in S according to their position in π, i.e., $\pi = (1, 2, \ldots, s)$ for $s = |S|$, (i is now named s). For each $j \in \{1, 2, \ldots, s\}$ write $m_j = v(\{1, 2, \ldots, j\}) - v(\{1, 2, \ldots, j-1\})$, and

$$b_S = \max_{\substack{R \subset S \\ s \in R}} \left(v(R) - \sum_{j \in R \setminus \{s\}} m_j \right).$$

[5] See also Sharkey (1982) and Moulin (1990) for a different characterization of convex games in terms of the property of their cores.

Since

$$\sum_{j \in R \setminus \{s\}} m_j = v(R \setminus \{s\}),$$

our assumption on S implies $b_S > m_S$. Consider now the vector $x = (m_1, \ldots, m_{s-1}, b_S)$.

Claim 1. *x is a semi-stable payoff of v_s.*

Proof. Since the restricted game on $\{1, 2 \ldots, s-1\}$ is assumed to be convex (the minimality of S), we have $x(R) = \sum_{j \in R} m_j \geq v(R)$ for all $R \subset \{1, 2, \ldots, s-1\}$. In fact $(x_1, \ldots, x_{s-1})$ is an extreme point of the core of $v_{\{1,2,\ldots,s-1\}}$. (See Shapley (1971)). Since $b_S \geq m_S$ we have $x(R) \geq v(R)$ for all $R \subset S$. This shows that the first condition for semi-stability is satisfied. We will now show that every player i can realize his payoff x_i in at least one coalition. Indeed if $i < s$, then $x(\{1, 2, \ldots, i\}) = v(\{1, 2, \ldots, i\})$. If $i = s$, then let R be the coalition for which the maximum in the definition of b_S is obtained and observe that $x(R) = v(R)$. So x is a semi-stable vector of v_S.

Claim 2. *x is not in the core of v_S.*

Proof. Since $\sum_{i=1}^{s} m_j = v(S)$, and since $b_s > m_s$, we have $x(S) > v_S(S)$. So x is not feasible with respect to the grand coalition in v_S, and thus is not a core payoff of this game. Claims 1 and 2 show that $x \in X(v_S) \setminus C(v_S)$. This completes the proof of the Lemma. Q.E.D.

Lemma 3.5. *If v is convex then $X(v_S) = C(v_S)$ for all $S \subset N$.*

Proof. The fact that $C(v_S) \subset X(v_S)$ is straightforward. We will thus show that $X(v_S) \subset C(v_S)$. Since all the restricted games of a convex game are also convex it is sufficient to show this inclusion for the grand coalition. Let $x \in X(v)$. To show that x is in $C(v)$ it is sufficient to demonstrate that x is feasible with respect to the grand coalition. For each i in N let $C_i(x)$ be a coalition which is minimal in $F_i(x)$ with respect to the inclusion relation, i.e, if $S \subsetneq C_i(x)$, then $S \notin F_i(x)$. We will first show that $C_i(x)$ is unique. Since x is kept fixed we set $C_i(x) = C_i$.

Suppose that C_i^1 and C_i^2 are both minimal in $F_i(x)$, with $C_i^1 \neq C_i^2$. Since $i \in C_i^1$ and $i \in C_i^2$, we have $C_i^1 \cap C_i^2 \neq \emptyset$. By the minimality with the fact that $x(S) \geq v(S)$ for all S, we get $x(C_i^1 \cap C_i^2) > v(C_i^1 \cap C_i^2)$. We now obtain

$$\begin{aligned} x(C_i^1 \cup C_i^2) &= x(C_i^1) + x(C_i^2) - x(C_i^1 \cap C_i^2) \\ &= v(C_i^1) + v(C_i^2) - x(C_i^1 \cap C_i^2) \\ &< v(C_i^1) + v(C_i^2) - v(C_i^1 \cap C_i^2) \\ &\leq v(C_i^1 \cup C_i^2). \end{aligned}$$

The last inequality follows from the convexity of v. We thus obtain $x(c_i^1 \cup C_i^2) < v(C_i^1 \cup C_i^2)$ which contradicts the fact that x is semi-stable.

We next show that for i,j in N with $i \neq j$, we have either $C_i = C_j$ or $i,j \in N \backslash (C_i \cap C_j)$. Suppose that $C_i \neq C_j$. If $i \in C_i \cap C_j$, then by the minimality of C_i we have $v(C_i \cap C_j) < x(C_i \cap C_j)$, and again with the same chain of inequalities as above we obtain $x(C_i \cup C_j) < v(C_i \cup C_j)$, which contradicts the fact that $x \in X(v)$. Similarly, we show that $j \in C_j$ is impossible. We thus get that for each i, and each $j \in C_i$ we have $C_i = C_j$, which means that the relation $i \sim j$ if and only if $i \in C_j$ is an equivalence relation on N. For each equivalence class choose some player k and let K denote the set of all these players. $\{C_k\}_{k \in K}$ is partition of N. Since each C_k satisfies $x(C_k) = v(C_k)$ we obtain by super-additivity (which is implied by the convexity):

$$x(N) = \sum_{k \in K} x(C_k) = \sum_{k \in K} v(C_k) < v(N).$$

So x is feasible for N and thus in the core of v. Q.E.D.

The Proof of Theorem 1. First, if the C-game v is convex, then by Proposition 3.3, for all $S \subset N$ $X(v_S) = C(v_S)$. Take $S \subset N$, and an SSPE b of G^S. Let $d_S = (d_1, \ldots, d_{|S|})$, where d_i is the payoff for i when b is played at any subgame starting with i proposing (see the proof of Lemma 3.2). By Lemma 3.2 $d_S \in X(v_S)$, and thus $d_S \in C(v_S)$. This means that d_S is feasible for the coalition S. By our assumption on players preferences on agreements, (i.e, that players prefer among two agreements in which they attain the same payoff the one with the larger coalition), we can conclude that S must form. Thus d_S is b's equilibrium payoff. This means that $E(G^S) \subset C(v_S)$ for all $S \subset N$, and together with Lema 3.1 this yields that v is core implementable. Suppose now that v is not convex, then, by Lemma 3.4, there exists some $S \subset N$ for which $C(v_S) \subsetneq X(v_S)$. Take $x \in X(v_S) \backslash C(v_S)$, and for each i in S let $T_i \subset S$ be some coalition for which $i \in T_i$ and $x(T_i) = v(T_i)$. Since $x \notin C(v_S)$, $T_i \neq S$ for all i in S. Consider the following stationary strategy profile $b = (b_i)_{i \in S}$ in G^S, where each b_i is given as follows:

(1) At each position in which i is a proposer i proposes (T_i, x_{T_i}).
(2) At each responder position (R, y_R, T) in which i is a responder, i accepts the proposal (R, y_R) if and only if $y_j \geq x_j$ for all $j \in S \backslash T$ (note that $i \in S \backslash T$ since i hasn't responded yet).

We will now show that b is an SSPE in G^S. Let us consider first the responder behavior. Assume that i is the last to respond to an existing proposal (R, y_R). If i rejects the proposal then, in the continuation of the game, i cannot attain more than x_i. Hence if $y_i \geq x_i$, i's optimal behavior is to accept the proposal as specified by b. If $x_i > y_i$ i's optimal behavior is to reject the proposal and to propose a different proposal which yields him his x_i. Now if i is not the last to respond, then inducting backwards on the number of players who still need to respond shows that the action specified for a responder by b is indeed a best response for every responder. Suppose now that i is a proposer, and that i deviates by proposing an outcome (R, y_R) with $y_i > x_i$. Since $x(R) \geq v(R)$, this means that for some $j \in R$, $y_j < x_j$, and according to the action specified by b for a responder, such proposal is going to be rejected by the first responder. If i as a proposer submits a proposal with $y_i < x_i$, then either this proposal is accepted and i has lost by deviating, or this proposal is rejected by some of the concerned players, say j. In the later case the rejecting player will propose again (T_j, x_{T_i}) which will then be approved as b specifies. Thus no proposer i can attain more than x_i by deviating at some subgame, which shows that b is an SSPE. Now, when b is played the resulting payoff vector is $z = (x_{T_i}, (v(j))_{j \in S \setminus T_i})$ for some i in S. (recall that $S \setminus T_i \neq \emptyset$). Since v is supper additive, $z(S) < v(S)$ and z is not a core payoff of v_S and the proof is completed. Q.E.D.

Remark. The second part of the proof of Theorem 1 shows that unlike other models of bargaining in characteristic function games, ours always admits stationary subgame perfect equilibria also in games with empty core. This follows from the fact that the set of semi-stable vectors is always non-empty. In fact every semi-stable vector x gives rise to an SSPE in which each player's equilibrium payoff at any decision point in which he is a proposer is x_i.

4. Examples

The following example shows that subgame perfect equilibria which employ non-stationary strategies may sustain outcomes which are not in the core even when the C-game is convex.

Example 4.1

Consider the following 2-person convex game: $v(1) = v(2) = 1$, $v(\{1, 2\}) = 3$. Consider first the following three auxiliary strategy combinations:

b_1. Player 1 proposes always $(N, (2, 1))$ and accepts a proposal if and only if $x_1 \geq 2$.
Player 2 proposes always $(N, (2, 1))$ and accepts a proposal if and only if $x_2 \geq 1$.

b_2. Player 1 proposes always $(N, (1, 2))$ and accepts a proposal iff $x_1 \geq 1$.
Player 2 proposes always $(N, (1, 2))$ and accepts a proposal iff $x_2 \geq 2$.

b_3. Player 1 proposes always $(N, (1.2, 1.2))$ and accepts a proposal iff $x_1 \geq 1.2$.
Player 2 proposes always $(N, (1.2, 1.2))$ and accepts a proposal iff $x_2 \geq 1.2$.

b_1, b_2 and b_3 are all stationary strategies. b_1 and b_2 are also SSPEs but b_3 is not.
Consider now the following non-stationary strategy combination:

b^*. Both players play b_3 as long as nobody deviates from b_3. Whenever some player $i \in \{1, 2\}$ deviates from b_3 both players move to play b_{3-i}.

b^* is a non-stationary equilibrium which, when played, yields the payoff (1.2, 1.2) which is not in the core of v (since it is not Pareto optimal). Moreover b^* is a subgame perfect equilibrium. To realize that, note that on subgames which occur after some player has deviated from b_3, b^* yields an equilibrium, since both b_1 and b_2 are equilibria. Furthermore, deviating from the path induced by b_3 cannot be optimal for either players, because this will yield the deviator a payoff of 1 which is less than 1.2.

The next example demonstrates that strictly super-additive games which are not convex may admit SSPE outcomes which are not in the core, even when the core is non-empty.

Example 4.2

Consider the three person game given by $v(i) = 1/5$ $(i = 1, 2, 3)$, $v(1, 2) = 1$, $v(1, 3) = 1$, $v(2, 3) = 1/2$ and $v(1, 2, 3) = 1.25$. The game v is strictly super-additive with a non-empty core but it is not convex. Consider now the following stationary strategy combination:

Player 1 proposes $(\{1, 2\}, (1/2, 1/2))$ and rejects an agreement if and only if it yields him less than $1/2$.

Player 2 proposes $(\{1, 2\}, (1/2, 1/2))$, and rejects an agreement if and only if it yields him less than $1/2$.

Player 3 proposes $(\{1, 3\}, (1/2, 1/2))$ and rejects an agreement if and only if it yields him less than a half.

The strategy combination described above is a stationary subgame perfect equilibrium. This follows from the fact that for a player to get more than $1/2$, would mean that some other player must be satisfied with less than $1/2$, but everyone rejects such agreements. Similarly to get $1/2$ within the grand coalition would mean a consent by some other player to obtain less than $1/2$.

Suppose that player 1 opens the game, then the equilibrium strategies induce the agreement $(\{1, 2\}, (1/2, 1/2))$. The payoff outcome of the game is thus $(1/2, 1/2, 0)$ which is outside the non-empty core. So v is not core-implementable. [6] Note that every restricted game of v is convex, so all the SSPE outcomes in

[6] Note that the same result can be achieved when either player 2 or player 3 open the game.

G^S for $S \subsetneq N$ must be core outcomes in the corresponding C-game. In fact the strategy combination above is an SSPE whenever $v(N) < 3/2$. For $v(N) \geq 3/2$ this strategy combination in not an SSPE any more, since every proposer can propose a better agreement consisting of the grand coalition, which will be accepted by the rest of the players. In particular for $v(N) = 3/2$ the agreement $(N, (1/2, 1/2, 1/2))$ is preferred by player 1 to $(\{1, 2\}, (1/2, 1/2))$. However to guarantee that all SSPE outcomes of G^N will be in the core of v we need to have $v(N) \geq 7/4$. For these values of $v(N)$ the game v is convex.

We conclude by illustrating the implications of our results for a model of public good provision.

Example 4.3. The provision of a public good

A group of agents has to decide on the provision and cost sharing of a public good. Let $g(\alpha)$ denote the cost of producing the amount α of the public good. Each agent has an initial endowment of money w_i, and a utility function $U_i(\alpha)$. We assume that $g(0) = 0$, and that $g(\alpha)$ and $U_i(\alpha)$ are increasing and continuous. To obtain a bounded set of payoffs we also assume that $g(\alpha)/\alpha$ is bounded above and that $U_i(\alpha)/\alpha$ approaches 0 as α approaches infinity.

If the amount α is produced with the cost allocation $(c_1, \ldots, c_n)$, then agent i's utility level is $U_i(\alpha) - c_i$. This public good economy induces the following TU game: $v(S) = \max_\alpha\{\sum_{i \in S} U_i(\alpha) - g(\alpha);\ g(\alpha) \leq \sum_{i \in S} w_i\}$. This definition is based on the assumption that each coalition can produce any quantity with total cost that does not exceeds the total initial endowment available to its members, and then make side payments among them.

It can easily be shown that the game v is convex. (see also Demange (1987)). This means first that, if the quantity of production and the cost allocation are determined by means of the bargaining model described above, then all the utility levels vectors which are sustainable by some stationary subgame perfect equilibrium are in the core of the economy. In particular, Demange (1987), (Proposition 1) also implies that the correspondence SSPE (U_N) [7] which associates to each vector of utility functions the set of all utility levels which are sustainable by some SSPE of the bargaining game is coalitional non-manipulable. This means that no coalition has an incentive to misrepresent [8] the preferences of its members.

[7] $U_N = \{U_i\}_{i \in N}$.

[8] Formally this means the following. Let $G(U_N)$ be the bargaining game based on the utility functions U_N. Let U'_S be some alternative utility functions for some coalition $S \subset N$. There exists no SSPE b of $G(U'_S, U_{N \setminus S})$ such that for each SSPE b' of $G(U_N)$, and for each i in S i's equilibrium payoff in b is greater than his equilibrium payoff in b'.

5. Conclusions

1. In this paper we have analyzed a bargaining game which implements the core of every game with increasing returns to scale for cooperation, i.e, convex game. We have also characterized the class of convex games by this property of core implementation. Typically implementation results are strongly dependent on the precise mechanism of negotiation which is applied. The one we have used here is the simplest and most straightforward generalization of the natural two person alternating bid model (without discounting), which includes coalition formation.

The fact that it is *exactly* the class of convex games which admits core implementation using our simple bargaining game, indicates the relation between the property of increasing returns for cooperation and the stability of bargaining outcomes. What we have shown is that the property of convexity is conducive for stability. This indeed seems to fit our intuition about real life environments of interactive decision making. In situations where every individual contribution is essential, and there is a lot to loose by an antagonistic formation of small groups, even when players behave competitively (or non-cooperatively) the collective interaction will necessarily induce a stable result.

2. We have imposed a weak condition on players preferences over agreements, namely that players prefer large coalitions to small ones provided that they earn the same payoff. This is done in order to prevent players from excluding others from the formed coalition just because they are indifferent between them joining the coalition or staying out. One might tend to think, that the notion of subgame perfection alone, will take care of breaking ties in favor of large coalitions. This is unfortunately not the case in our context. Consider for example the following 3-person convex game;

$v(\{i\}) = 0$ $(i = 1, 2, 3)$
$v(S) = 1$ for $|S| = 2$, and $v(N) = 3$.

Consider the following SSPE of our bargaining game on the set of players N.

Player 1: Proposes $(\{1, 2\}, (0, 1))$, and accepts any proposal yielding him a non-negative payoff.
Player 2: Proposes $(\{1, 2\}, (0, 1))$, and accepts a proposal if and only if it yields him at least 1.
Player 3: Proposes $(N, (0, 1, 2))$, and accepts a proposal if and only if it yields him at least 2.

It is easily shown that the above strategy combination is an SSPE. (See the second part of the proof of Theorem 1.) Now, if the bargaining starts with either player 1 or 2, then the equilibrium outcome involves the coalition $\{1, 2\}$ and not the grand coalition. The equilibrium outcome thus cannot be in the core. This rather technical problem will arise whenever the strategy combination is based on an extreme point of the core.

References

Albers, W., 1979, Core – and kernel variants based on imputations and demand profiles, in: Moeschin and Pallaschke, eds., Game theory and related topics (North-Holland, Amsterdam).

Aumann, R.J., 1959, Acceptable points in general cooperative *n*-person games, in: H.W. Kuhn and R.D. Luce, eds., Contribution to the theory of games, Vol. 4, 287–324. (Princeton University Press, Princeton, NJ).

Bennett, E., 1983, The aspirations approach to predicting coalition formation and payoff distribution in side payment games, International Journal of Game Theory 12, 1–28.

Chatterjee, K., B. Dutta, D. Ray and K. Sengputa 1993, A non-cooperative theory of coalitional bargaining, Review of Economic Studies 60, 463–447.

Demange, G., 1987, Non manipulable cores, Econometrica 55, 1057–1074.

Ichiishi, T., 1992, The cooperative nature of the firm, forthcoming.

Moldovanu, B. and E. Winter, 1991, On order independent equilibria, Games and Economic Behavior, forthcoming.

Moulin, H., 1990, Cores and large cores when population varies, International Journal of Game Theory 19, no. 2, 219–232.

Perry, M. and P. Reny, 1991, A non-cooperative view of coalition formation and the core, Discussion paper (The Center for Rationality and Interactive Decision Theory, The Hebrew University of Jerusalem, Jerusalem).

Rosenmueller, J., 1981, The theory of games and markets (North-Holland, Amsterdam).

Selten, R., 1981, A non-cooperative model of characteristic function bargaining, in: V. Bohm and H.H. Nachtkamp, eds., Essays in Game Theory and Mathematical Economics in Honor of Oscar Morgenstern.

Shapley, L., 1971, Cores of convex games, International Journal of Game Theory 1, 11–26.

Shapley, L. and M. Shubik, 1969, On market games, Journal of Economic Theory 1, 9–25.

Sharkey, W., 1982, Cooperative games with large cores, International Journal of Game Theory 11, 175–182.

Shubik, M., 1987, A game theoretic approach to political economy (The MIT Press, Cambridge, MA).

Van Damme Eric, R. Selten and E. Winter, 1990, Alternating bid bargaining with a smallest money unit, Games and Economic Behavior 2, 188–201.

Winter, E., 1992, Mechanism robustness in multilateral bargaining, Discussion paper no. 7 (The Center of Rationality and Interactive Decision Theory, The Hebrew University of Jerusalem, Jerusalem).