Available online at www.sciencedirect.com
SCIENCE DIRECT®
Games and Economic Behavior 43 (2003) 196–213
GAMES and Economic Behavior
www.elsevier.com/locate/geb

Consistent solutions for cooperative games with externalities ☆

Chen-Ying Huang [a,*] and Tomas Sjöström [b]

[a] *Department of Economics, National Taiwan University, 21 Hsu-Chow Road, Taipei, Taiwan*
[b] *Department of Economics, Penn State University, University Park, PA 16802, USA*

Received 26 January 2000

Abstract

In order to calculate the worth of a coalition of players, the coalition needs to predict the actions of outsiders. We propose that, for a given solution concept, such predictions should be made by applying the solution concept to the "reduced society" consisting of the non-members. We illustrate by computing the *r-core* for the case of Bertrand competition with differentiated commodities.

JEL classification: C71

Keywords: Recursive core; Externalities

1. Introduction

Classical cooperative game theory studies situations where agreements are fully binding and enforceable (Aumann, 1989; Luce and Raiffa, 1957). The starting point is a *characteristic function* which specifies the *worth* of each coalition. The worth is what the coalition can achieve on its own without cooperating with outside players. If there are no externalities, i.e., if the payoffs to the members of a coalition do not depend on actions taken by non-members, then the worth can be defined without specifying the actions of non-members. But if externalities are present, then in order to calculate the worth of a coalition one must predict the actions of non-members. This article considers the problem of defining a characteristic function in the presence of externalities. No new *solution*

☆ An earlier version of this paper had the title *The p-core*.

* Corresponding author.
E-mail addresses: chenying@ccms.ntu.edu.tw (C.-Y. Huang), jts10@psu.edu (T. Sjöström).

0899-8256/03/$ – see front matter
doi:10.1016/S0899-8256(03)00011-3

concept is proposed. Since the classical literature contains a number of solutions for games in characteristic function form, we think a basic step in the exploration of games with externalities should be to look at the predictions made by the classical solution concepts.

Suppose a set of players, denoted N, play a normal form game with transferable utility. It will be efficient for the grand coalition N to form and choose a strategy profile which maximizes the joint surplus. But a coalition $S \subseteq N$ may refuse to cooperate with the other players. Since there are externalities among the coalitions, to calculate the worth of coalition S we must predict what the players in $N \backslash S$ will do once S has formed. Two well-known ways of defining the worth of a coalition, i.e., of constructing a characteristic function, are the α- and β-theories (Shubik, 1983, pp. 136–138). We introduce a new *recursive* method, called the r-theory. Once a method for defining the characteristic function has been found, all the classical solution concepts such as the core, the bargaining set, or the stable set can be applied. But a novel aspect of our theory is that the worth of a coalition depends on which solution concept is used to predict the behavior of the complement, i.e., the characteristic function will depend on the solution concept.

We assume utility is transferable and binding contracts can be written within coalitions. Given a coalition structure, we assume (following Ichiishi (1981) and Ray and Vohra (1997)) that each coalition acts as a "composite player" who maximizes the joint payoff of its members, and we look at the Nash equilibria of the resulting normal form game. Our main concern is to study the formation of coalitions. The main assumption is that coalition S, if it forms, predicts that the players in $N \backslash S$ will behave in a way which is consistent with the same solution concept that is proposed for the game itself (but now applied to the *reduced game* with player set $N \backslash S$). The definition is recursive. Thus, for example, the *r-core* is defined as follows. For $k = 1$, we define the r-core for k-player societies by individual payoff maximization. Suppose for any $k \geqslant 1$, the r-core is defined for any society which has no more than k players. Then the r-core for a $(k+1)$-player society N consists of those payoff vectors that give each coalition $S \subseteq N$ at least its worth. The worth of a proper subcoalition $S \subset N$ is defined by using the r-core to predict the behavior of the players in the remaining "reduced society" $N \backslash S$. This can be done because $N \backslash S$ has at most k members. The worth of N itself is defined by joint payoff maximization.

A possible ambiguity in this definition is that the r-core for $N \backslash S$ need not be a singleton. In such cases we assume the players in S pessimistically expect that r-core outcome for $N \backslash S$ which is worst for S. Thus, coalitions are pessimistic subject to the constraint that outcomes must be consistent with the solution concept. In contrast, according to the α-theory coalitions are pessimistic without paying attention to such a constraint. That is, the α-theory defines the worth of a coalition to be what it can guarantee itself *regardless* of the actions of outsiders. In the r-theory, the actions of the outsiders must be consistent with the solution concept. Since this makes coalitions less pessimistic about what they can achieve, the r-core is a subset of the α-core. (Since we will allow players in a coalition to use correlated strategies the distinction between the α- and β-cores will not be important, and we will refer only to the α-core.)

Ray and Vohra (1997) pioneered the application of consistency to cooperative games with externalities. They proposed a new solution concept called Equilibrium Binding Agreements (EBA). Since our focus is on sustaining cooperation in the grand coalition, we will mainly compare our predictions to the EBA for the grand coalition. In Ray and Vohra's theory a strategy vector for the grand coalition is blocked if there is a "leading perpetrator" $S \subset N$ that can induce another strategy vector under a finer coalition structure which would make itself better off. The new finer coalition structure must be stable in the sense that coalitions should not break apart due to the same type of internally credible (equilibrium) deviations. (It is assumed that coalitions can break apart but not re-merge, hence the finest possible coalition structure is always by definition stable.) Also, if the deviation involves other coalitions except S breaking apart, then all newly formed coalitions (except perhaps one) must be "secondary perpetrators" who are better off than they would have been if the coalition had not deviated. We do not use the notions of leading and secondary perpetrators. Instead we propose a recursive definition for the worth of a coalition which can be used together with the classical solution concepts. Moreover, we will follow the classical approach to cooperative games by assuming that when coalition S calculates its worth, it does *not* worry about destabilizing deviations *within* S. This is because, by assumption, coalition S can sign an agreement which is fully binding and enforceable.

A non-cooperative view of coalition formation with binding agreements is provided by Perry and Reny (1994). In their non-cooperative game, player i can propose that player j, for example, signs a binding agreement to form the coalition $\{i, j\}$. After signing the agreement himself, player i hands it over to player j. By signing, player j can guarantee that $\{i, j\}$ forms (as can player i, after player j has signed). Of course, the fact that $\{i, j\}$ *can* sign such a binding agreement does not mean that the players in $N\backslash\{i, j\}$ predict that $\{i, j\}$ *will* actually do so, even if the players in $N\backslash\{i, j\}$ should refuse to cooperate with players i and j. In Section 6 we prove that, with three players, a version of Perry and Reny's game gives a non-cooperative foundation for the r-core. The details of a full non-cooperative implementation are left for another paper. Several recent articles have studied the non-cooperative equilibria of different extensive form games of coalition formation. Major contributions are due to Bloch (1995), Ray and Vohra (1999), and Yi (1997). In practise one may have insufficient information about the appropriate extensive form, so the cooperative and non-cooperative approaches may be useful complements.

In the model presented in this article, it will always be efficient for the grand coalition to form. However, in general partition function form games it need not be efficient for the grand coalition to form. For example, transactions costs might make large coalitions costly to organize. In this case, the r-theory can still be used, and it may predict coalition structures that are strict refinements of the grand coalition. The analysis of general partition function form games is avoided in this article in order to simplify the exposition.

2. Definitions

There is a set $N = \{1, 2, \ldots, n\}$ of players who play a normal form game Γ. Let A_i denote player i's set of pure strategies, and $\Delta(A_i)$ his set of mixed strategies. For simplicity we assume $|A_i| < \infty$, where $|A_i|$ denotes the number of elements in A_i. Player i's payoff in Γ is $\pi_i(a)$, where $a = (a_1, \ldots, a_n) \in A \equiv \bigtimes_{i \in N} A_i$. If $\sigma_i \in \Delta(A_i)$ is a probability distribution over A_i, then $\sigma_i(a_i)$ denotes the probability of taking action a_i. We extend player i's payoff function to mixed strategy profiles in the usual way:

$$\pi_i(\sigma_1, \ldots, \sigma_n) = \sum_{a \in A} \prod_{j \in N} \sigma_j(a_j) \pi_i(a).$$

Before playing the game, the n players can organize themselves into coalitions. The result is a partition $\mathcal{P}_N = (S_1, \ldots, S_K)$ of N, where K is the number of coalitions. We assume the possibility of unrestricted side payments within a coalition. There are no transfers across coalitions. For $S_k \in \mathcal{P}_N$, let $t_{ij} \geqslant 0$ denote the side payment player $i \in S_k$ makes to player $j \in S_k$. The net transfer received by player $i \in S_k$ is

$$t_i \equiv \sum_{j \in S_k} (t_{ji} - t_{ij}).$$

Player i's utility if action profile $a \in A$ is taken and his net transfer is t_i is $\pi_i(a) + t_i$. Within the coalition transfers must balance:

$$\sum_{i \in S_k} t_i = 0. \tag{1}$$

For a given partition $\mathcal{P}_N$, let $T_{\mathcal{P}_N}$ denote the set of feasible net transfer profiles $(t_1, \ldots, t_n)$, i.e., transfers that satisfy Eq. (1) for all $S_k \in \mathcal{P}_N$. A partition of a subset $S \subset N$ is denoted $\mathcal{P}_S$, and $T_{\mathcal{P}_S}$ denotes the set of feasible net transfers $(t_i)_{i \in S}$, i.e., transfers that satisfy Eq. (1) for all $S_k \in \mathcal{P}_S$.

A partition $\mathcal{P}_N$ induces a normal form game $\Gamma(\mathcal{P}_N)$ among the coalitions as follows. Each coalition $S_k \in \mathcal{P}_N$ is a "composite player." Composite player S_k has a pure strategy set $A_{S_k} \equiv \bigtimes_{i \in S_k} A_i$, so a strategy specifies an action for each coalition member. Denote a typical element of A_{S_k} by a_{S_k}. Since there is transferable utility within a coalition, each coalition should aim to maximize the sum of the payoffs of its members. The payoff for composite player S_k is therefore the sum of the members' payoffs:

$$\pi_{S_k}(a_{S_1}, \ldots, a_{S_K}) \equiv \sum_{i \in S_k} \pi_i(a_{S_1}, \ldots, a_{S_K}).$$

Let $\sigma_{S_k} \in \Delta(A_{S_k})$ denote a probability distribution over the set A_{S_k}. The members of coalition S_k can use correlated strategies, i.e., strategies in $\Delta(A_{S_k})$. Let $\sigma = (\sigma_{S_1}, \ldots, \sigma_{S_K})$ and $(\sigma'_{S_k}, \sigma_{-S_k}) \equiv (\sigma_{S_1}, \ldots, \sigma_{S_{k-1}}, \sigma'_{S_k}, \sigma_{S_{k+1}}, \ldots, \sigma_{S_K})$. Extend the payoff function to mixed strategies in the usual way:

$$\pi_{S_k}(\sigma) = \sum_{a_{S_1} \in A_{S_1}} \cdots \sum_{a_{S_K} \in A_{S_K}} \sigma_{S_1}(a_{S_1}) \ldots \sigma_{S_K}(a_{S_K}) \pi_{S_k}(a_{S_1}, \ldots, a_{S_K}).$$

We identify the original game Γ with $\Gamma(\{1\}, \{2\}, \ldots, \{n\})$, i.e., it is the game where each coalition is a singleton.

Let $E(\mathcal{P}_N)$ denote the set of Nash equilibria of the game $\Gamma(\mathcal{P}_N)$. That is, $\sigma \in E(\mathcal{P}_N)$ iff $\sigma \in \bigtimes_{S_k \in \mathcal{P}_N} \Delta(A_{S_k})$, and for every $S_k \in \mathcal{P}_N$ and every $\sigma'_{S_k} \in \Delta(A_{S_k})$,

$$\pi_{S_k}(\sigma_{S_k}, \sigma_{-S_k}) \geqslant \pi_{S_k}\left(\sigma'_{S_k}, \sigma_{-S_k}\right).$$

Following Ichiishi's (1981) notion of non-cooperative play between coalitions, we will assume that if the coalition structure is $\mathcal{P}_N$, the outcome will be some strategy $\sigma \in E(\mathcal{P}_N)$.

3. The r-theory

For any $S \subseteq N$, let 2^S denote the set of all subsets of S, and let R^S denote the payoff space pertaining to S. Let $V : 2^S \to R$ be a *characteristic function for the reduced society* S, i.e., a function which assigns a value $V(T)$ to each $T \subseteq S$. Let $\mathcal{V}^S$ denote the set of all such functions. A *presolution for society* S is a correspondence $G^S : \mathcal{V}^S \to R^S$. To illustrate, we give two examples of presolutions.

Example A. The core presolution:

$$G^S(V) = \left\{ \{u_i\}_{i \in S} \colon \sum_{i \in T} u_i \geqslant V(T) \text{ for all } T \subseteq S \right\}. \tag{2}$$

Example B. The Zhou (1994) bargaining set presolution. Here $\{u_i\}_{i \in S} \in G^S(V)$ if and only if the following is true. For each $T \subseteq S$, if we can find a vector $\{y_k\}_{k \in T}$ so that for every $k \in T$

$$u_k < y_k \quad \text{and} \quad \sum_{k \in T} y_k = V(T)$$

then there must exist a coalition $U \subseteq S$ and a vector $\{z_h\}_{h \in U}$ such that

$$\begin{aligned} &U \backslash T \neq \emptyset, \qquad T \backslash U \neq \emptyset, \quad \text{and} \quad U \cap T \neq \emptyset, \\ &u_h \leqslant z_h, \quad \text{for all } h \in U \backslash T, \qquad y_h \leqslant z_h, \quad \text{for all } h \in U \cap T, \end{aligned} \tag{3}$$

and

$$\sum_{h \in U} z_h = V(U).$$

In our theory, when a coalition $U \subseteq N$ calculates its "worth," it will use some presolution $G^{N \backslash U}$, together with feasibility constraints, to predict the outcome of the "reduced game" made up of the players in the $N \backslash U$. In order to use $G^{N \backslash U}$ to solve the reduced game, coalition U needs to know the worth of every coalition $T \subseteq N \backslash U$, i.e., it needs to know a characteristic function $V \in \mathcal{V}^{N \backslash U}$. We will now show how to derive such a characteristic function. As a first step we assume a presolution G^S has been exogenously assigned to every $S \subseteq N$.

Fix a coalition $S \subseteq N$, and consider the reduced game among the players in S, given that the coalition structure of the complement is given by the partition $\mathcal{P}_{N \backslash S}$. Let $C(S \mid S, \mathcal{P}_{N \backslash S})$ denote the set of strategies that can possibly be played in this situation. To determine this set, we will first estimate the worth of each subcoalition $T \subseteq S$ in this reduced game, denoted $V(T \mid S, \mathcal{P}_{N \backslash S})$. Then, we will apply the exogenously given presolution G^S. We do this recursively in a consistent way.

For any $i \in N$, and any coalition structure $\mathcal{P}_{N \backslash \{i\}}$ for the players in $N \backslash \{i\}$, the set of strategies that can possibly be played in the game $\Gamma(\{i\}, \mathcal{P}_{N \backslash \{i\}})$ is the set of Nash equilibria of this game, so

$$C\big(\{i\} \bigm| \{i\}, \mathcal{P}_{N \backslash \{i\}}\big) \equiv E\big(\{i\}, \mathcal{P}_{N \backslash \{i\}}\big). \tag{4}$$

Now make the induction hypothesis that we have defined a set of possible strategies $C(S \mid S, \mathcal{P}_{N\setminus S})$ for all S such that $1 \leqslant |S| \leqslant s-1$, and all partitions $\mathcal{P}_{N\setminus S}$ of $N\setminus S$. Consider S such that $|S| = s$. Define

$$V(S \mid S, \mathcal{P}_{N\setminus S}) \equiv \min\{\pi_S\colon \pi_S = \pi_S(\sigma) \text{ for some } \sigma \in E(S, \mathcal{P}_{N\setminus S})\} \tag{5}$$

and for any non-empty proper subset $T \subset S$ define

$$V(T \mid S, \mathcal{P}_{N\setminus S}) \equiv \min\{\pi_T\colon \pi_T = \pi_T(\sigma) \text{ for some } \sigma \in C(S\setminus T \mid S\setminus T, T, \mathcal{P}_{N\setminus S})\}. \tag{6}$$

The right-hand side of (6) is well-defined by the induction hypothesis, because $S\setminus T$ is a non-empty set which has strictly fewer players than S, and $(T, \mathcal{P}_{N\setminus S})$ is a partition of $N\setminus(S\setminus T)$. We will calculate $C(S \mid S, \mathcal{P}_{N\setminus S})$ by using the (exogenously given) presolution G^S together with the characteristic function for S defined by:

$$V(T) = V(T \mid S, \mathcal{P}_{N\setminus S}), \quad \text{for all } T \subseteq S, \tag{7}$$

where $V(T \mid S, \mathcal{P}_{N\setminus S})$ is given by Eq. (5) for $T = S$, and by Eq. (6) for $T \subset S$. So let $C(S \mid S, \mathcal{P}_{N\setminus S})$ be the set of strategies σ such that, for some partition $\mathcal{P}_S$ of S and some net transfer profile $(t_i)_{i\in S} \in T_{\mathcal{P}_S}$, (8) and (9) hold:

$$\sigma \in E(\mathcal{P}_S, \mathcal{P}_{N\setminus S}), \tag{8}$$

$$\{\pi_i(\sigma) + t_i\}_{i\in S} \in G^S(V), \tag{9}$$

where $V \in \mathcal{V}^S$ is defined by Eq. (7). Notice that in the definition of $C(S \mid S, \mathcal{P}_{N\setminus S})$ we allow coalition S to divide itself into a non-trivial partition $\mathcal{P}_S$. (For example, in the quantity setting Cournot oligopoly studied in Huang and Sjöström (1998), the firms in a coalition $S \neq N$ may prefer to break apart, because the output of the firms in $N\setminus S$ is lower when S is divided into several firms than when S is one big firm.)

We have now defined $C(S \mid S, \mathcal{P}_{N\setminus S})$ for S such that $|S| = s$. Continuing this way, we can define $C(S \mid S, \mathcal{P}_{N\setminus S})$ for larger and larger S. The final step occurs when $S = N$. Although the method in the final step is the same as in the preceding steps, it may be helpful to describe this last step explicitly. First, to calculate $V(N \mid N)$ according to (5), notice that $E(N)$ is simply the set of strategies that maximize the joint payoffs of the players in N; thus

$$V(N \mid N) \equiv \max\left\{\sum_{i\in N} \pi_i(\sigma)\colon \sigma \in \Delta(A)\right\}. \tag{10}$$

Strictly speaking, $V(N \mid N)$ should be written $V(N \mid N, \mathcal{P}_{N\setminus N})$ to conform to (5). But $\mathcal{P}_{N\setminus N}$ is the trivial partitioning of the empty set since $N\setminus N = \emptyset$, and to simplify notation we write $V(N \mid N)$ instead of $V(N \mid N, \mathcal{P}_{N\setminus N})$.

For any non-empty proper subset $T \subset N$ define $V(T \mid N)$ according to (6), by

$$V(T \mid N) \equiv \min\{\pi_T\colon \pi_T = \pi_T(\sigma) \text{ for some } \sigma \in C(N\setminus T \mid N\setminus T, T)\}. \tag{11}$$

Finally, let

$$V(T) = V(T \mid N), \quad \text{for all } T \subseteq N. \tag{12}$$

Then, $C(N \mid N)$ is the set of strategies $\sigma \in \Delta(A)$ such that there exists a net transfer profile $(t_1, \ldots, t_n) \in T_N$ such that:

$$\{\pi_i(\sigma) + t_i\}_{i \in N} \in G^N(V), \tag{13}$$

where $V \in \mathcal{V}^N$ is defined by Eq. (12).

In payoff space, the solution to the game is the set of payoff vectors that can be expressed as $(\pi_1, \ldots, \pi_n) = (\pi_1(\sigma) + t_1, \ldots, \pi_n(\sigma) + t_n)$, for some σ and $(t_1, \ldots, t_n)$ that satisfy (13). For example, the *r-core* is the set of payoff vectors that can be expressed as $(\pi_1, \ldots, \pi_n) = (\pi_1(\sigma) + t_1, \ldots, \pi_n(\sigma) + t_n)$, for some σ and $(t_1, \ldots, t_n)$ that satisfy (13), where at each step of the recursive argument we use the core presolution (i.e., G^S is always given by (2)). The *r-bargaining set* is defined in the same way, but replacing the core presolution by the Zhou bargaining set presolution.

Notice also that if we apply the core (say) presolution to all societies, we may talk about the r-core for a reduced society in a natural way. More precisely, for any $S \subset N$, the r-core for S given the partition $\mathcal{P}_{N \setminus S}$ is the set of payoff vectors that can be expressed as $(\pi_i)_{i \in S} = (\pi_i(\sigma) + t_i)_{i \in S}$, for some σ and $(t_i)_{i \in S}$ that satisfy (8) and (9) (where G^S is given by (2)).

In contrast to our theory, in the α-theory coalitions apply maximal pessimism (Shubik, 1983, pp. 136–138). The worst that can happen to coalition S is that the complement $N \setminus S$ forms and "maximins" S. Hence, in the α-theory the worth of S is

$$V^\alpha(S) = \max_{\sigma_S \in \Delta(A_S)} \min_{\sigma_{N \setminus S} \in \Delta(A_{N \setminus S})} \pi_S(\sigma_S, \sigma_{N \setminus S}). \tag{14}$$

Now consider a payoff vector $(\pi_1, \ldots, \pi_n) = (\pi_1(\sigma) + t_1, \ldots, \pi_n(\sigma) + t_n)$, where $\sigma \in \Delta(A)$ and $(t_1, \ldots, t_n) \in T_N$. To take the core as an example, $(\pi_1, \ldots, \pi_n)$ is in the α-core iff $\sum_{i \in S} \pi_i \geqslant V^\alpha(S)$ for all $S \subseteq N$. But $(\pi_1, \ldots, \pi_n)$ is in the r-core iff $\sum_{i \in S} \pi_i \geqslant V(S)$ for all $S \subseteq N$, where V is defined by our recursive procedure. Certainly $V(S) \geqslant V^\alpha(S)$, so if the r-core exists then it must be a subset of the α-core. The characteristic function defined by (14) was introduced by von Neumann and Morgenstern (1944). They showed that V^α is superadditive. In contrast, our characteristic function V need not be superadditive. For example, superadditivity fails in the Cournot oligopoly discussed by Huang and Sjöström (1998).

The r-solution suffers from two kinds of existence problems. First, it may happen that $C(N \setminus T \mid N \setminus T, T)$ is empty for some $T \subset N$, in which case there is no prediction that coalition T can make about the behavior of its complement. Then $V(T) = V(TN)$ cannot even be defined as in (11). In this case we may say that the r-solution does not exist, since we are not able to define characteristic functions in a consistent way for all reduced societies. Second, even if $V(T)$ is well defined for each $T \subseteq N$, it may still not be possible to satisfy (13). For example, in the case of the r-core, we may be able to define the function V, but it may not be balanced. In such a case we may say that the r-solution exists but is empty.

Thus, for the r-solution to exist, it must exist and be non-empty for every reduced society $T \subset N$, or else the complement $N \setminus T$ cannot calculate its worth. This requirement may be too strong. The following argument was made by D. Ray. Suppose $N = \{1, 2, 3, 4\}$, and consider the reduced society $T = \{2, 3, 4\}$. Suppose we have recursively calculated the worth of T and all its subcoalitions, using the core presolution, and found that T itself is worth six, and any of the two player sub coalitions $\{2,3\}$, $\{2,4\}$, and $\{3,4\}$ is worth five. Player 2, 3 or 4 on his own is worth zero. Then, the r-core of the reduced society T is empty, because in this society it would be efficient for all of T to cooperate and divide up the six, but there is no way to satisfy all two-player subcoalitions of T (each two-player subcoalition would insist on at least five). Therefore, player 1 cannot make any prediction which is consistent with the r-core about what his complement $N\backslash\{1\} = T$ would do if he went on his own. Hence he cannot calculate his worth, and the r-core of this game does not exist. But suppose player 1 has a dominant strategy that gives him five no matter what the other players do. It would seem reasonable to define the worth of player 1 to be five, and the α-core does this, but not the r-core. It is a topic for future research to develop solution concepts that relax the requirement of existence of a non-empty solution set for each reduced society. (An anonymous referee suggested that by defining a Shapley value presolution we could generate a prediction even when the r-core does not exist.) We note, however, that in the non-cooperative approach of Perry and Reny (1994), discussed in Section 6, the requirement that the r-core is non-empty for every reduced society corresponds to the requirement that an equilibrium exists in every subgame, even those that are not reached along the equilibrium path. Thus, the requirement has a natural interpretation in the non-cooperative setting.

Remark 1. There are cases when it is not reasonable to expect a particular coalition to form. Fortunately, this does not matter for the r-core. To illustrate, suppose $n = 4$, and suppose in our analysis of the r-core we have derived the following characteristic function: $V(N) = 20$, $V(\{1,2,3\}) = 7$, $V(\{1,2\}) = V(\{1,3\}) = V(\{2,3\}) = 6$, and $V(S) = 0$ otherwise. The payoff vector $x = (1,1,1,17)$ can be blocked by coalition $S = \{1,2,3\}$. However, if S forms, how would they divide up the 7? At least some pair $\{i,j\} \subset S$ would get strictly less than six, but players i and j would not agree to this because they can block x on their own and take $V(\{i,j\}) = 6$. In the terminology of Ray (1989), $\{1,2,3\}$ does not *credibly* block $x = (1,1,1,17)$. On the other hand, any two-player subcoalition of $\{1,2,3\}$ can credibly block x. Ray (1989) showed that this situation is general. That is, whenever a coalition $S \subset N$ blocks a payoff vector x and the blocking is not credible, then there exists a subcoalition $T \subset S$ that credibly blocks x. Thus, restricting attention to credibly blocking coalitions would not make any difference to the r-core.

4. An example

In this section we provide an example to illustrate the role of binding agreements. Player 1 chooses a row, player 2 chooses a column, player 3 chooses a matrix. Note that players 1 and 2 are symmetric.

	L				*R*		
	ℓ	m	r		ℓ	m	r
U	5, 5, 5	−10, 1, 9	2, 6, −1	U	7, 7, 0	0, 0, 0	0, 8, 0
M	1, −10, 9	−11, −11, 21	1, −9, 9	M	0, 0, 0	0, 0, 0	0, 1, 0
D	6, 2, −1	−9, 1, 9	4, 4, −1	D	8, 0, 0	1, 0, 0	6, 6, 0

We first derive the α-core and the r-core, using $V^\alpha(S)$ and $V(S)$ to denote the worth of coalition S according to the α- and r-theories, respectively. Strategy D is strictly dominant for player 1, strategy r is strictly dominant for player 2, and player 3's best response against Dr is R. Thus, for the finest coalition structure $(\{1\},\{2\},\{3\})$ we predict the outcome (D,r,R) with payoffs $(6,6,0)$.

The grand coalition $N=\{1,2,3\}$ maximizes its joint payoff by choosing $U\ell L$. Thus, $V^\alpha(N)=V(N)=15$. What is the worth of coalition $\{2,3\}$? The greatest joint payoff players 2 and 3 can get is ten, and by playing mL they get ten *regardless* of player 1's action. Thus, $V^\alpha(\{2,3\})=V(\{2,3\})=10$. By symmetry, $V^\alpha(\{1,3\})=V(\{1,3\})=10$. For coalition $\{1,2\}$, $U\ell$ is a strictly dominant strategy which gives them at least 10, so $V^\alpha(\{1,2\})=10$. Since player 3's best response against $U\ell$ is L, in fact also $V(\{1,2\})=10$.

Next, consider the singleton coalition $\{3\}$. Player 3 can guarantee himself at most 0, by playing R, so $V^\alpha(\{3\})=0$. To calculate $V(\{3\})$, player 3 has to predict how players 1 and 2 would behave if player 3 refuses to cooperate with them. Would 1 and 2 stay together or break up? If coalition $\{1,2\}$ forms, they get 10 as argued above. On the other hand, if player 1 refuses to cooperate with player 2, the result is the finest coalition structure $(\{1\},\{2\},\{3\})$ in which case player 1 gets 6. Similarly, player 2 can get 6 by separating from 1. Since $6+6>10$, the r-core prediction for the reduced society $\{1,2\}$ is that players 1 and 2 split up. The outcome will be (D,r,R), so $V(\{3\})=0$.

For the other two non-trivial reduced societies, $\{1,3\}$ and $\{2,3\}$, the r-core predicts that the coalition stays together. Hence, we calculate $V(\{1\})=V(\{2\})=-9$. Also, $V^\alpha(\{1\})=V^\alpha(\{2\})=-9$.

Since $V^\alpha(S)=V(S)$ for all S, the α-core equals the r-core in this example. In either case, the payoffs must satisfy:

$$\pi_1+\pi_2+\pi_3=15,\qquad \pi_1+\pi_3\geqslant 10,\qquad \pi_2+\pi_3\geqslant 10,\qquad \pi_1+\pi_2\geqslant 10,$$
$$\pi_1\geqslant -9,\qquad \pi_2\geqslant -9,\qquad \pi_3\geqslant 0.$$

The unique payoff in the r-core is, therefore, $(5,5,5)$.

Notice that, since each two-player coalition can *guarantee* itself ten, ten is their worth according to the α-theory. Thus, each pair of players must get at least ten in the α-core. Since we allow for less pessimistic expectations, it must be true that $V(\{i,j\})\geqslant V^\alpha(\{i,j\})$. In fact, in this example, a two-player coalition cannot realistically hope for *more* than ten, so its worth according to the r-theory is ten as well. There are EBA for the grand coalition that give two-player coalitions much less than ten, however, as we now discuss.

According to Ray and Vohra (1997), every coalition has to worry about destabilizing internal deviations, which can prevent a coalition from forming. In this example, the set of EBA for the grand coalition consists of all payoff vectors (π_1,π_2,π_3) such that $\pi_1+\pi_2+\pi_3=15$, $\pi_1+\pi_3\geqslant 10$, $\pi_2+\pi_3\geqslant 10$, $\pi_1\geqslant -9$, $\pi_2\geqslant -9$, $\pi_3\geqslant 0$. Thus, there is an EBA for the grand coalition which gives minus nine each to players 1 and 2, and thirty-three to player 3. While a single player (either player 1 or 2) who leaves the grand coalition expects only -9, a single player who leaves the coalition $\{1,2\}$ expects 6. This is because if coalition $\{1,2\}$ breaks up then neither player is expected to merge with player 3, so that the result will be the finest coalition structure $(\{1\},\{2\},\{3\})$ with outcome (D,r,R). Since coalition $\{1,2\}$ only expects 10, and $6+6>10$, the coalition $\{1,2\}$ is not internally stable and is not allowed to block $(-9, -9, 33)$ according to the theory behind the EBA. Neither can any player $i \in \{1, 2\}$ block $(-9, -9, 33)$ on his own, because refusing to cooperate with the other two players would yield him only -9 (recall that player $i \in \{1, 2\}$ fears that if he defects from the grand coalition then the other two players will form a two-person coalition). Thus, according to the EBA, players 1 and 2 have no blocking power, either jointly or individually.

We agree with Ray and Vohra that *individually* players 1 and 2 are weak. However, we argue that if binding contracts are possible, then together players 1 and 2 are quite strong. They can and should insist on getting at least $V(\{1, 2\}) = 10$, because—by assumption—they can sign a binding contract which *guarantees* ten. If $(-9, -9, 33)$ is proposed as a payoff vector for the grand coalition, then players 1 and 2 can do better by signing a binding agreement which gives them five each. (Notice in passing that such blocking is credible in the sense of Ray (1989) because each player $i \in \{1, 2\}$ prefers to sign *this* agreement rather than trying to block on his own—recall that $V(\{i\}) = -9$.)

In this example, even if a break-up of coalition $\{1, 2\}$ results in the finest coalition structure and strategy profile (D, r, R), then this should not prevent players 1 and 2 from blocking a payoff vector such as $(-9, -9, 33)$. Indeed, if players 1 and 2 can conspire to induce (D, r, R), then their blocking power in the grand coalition is even increased, since (D, r, R) gives them six each. A modified version of EBA would indeed allow players 1 and 2 to induce (D, r, R), by treating both as leading perpetrators (cf. (Vohra, 1997, p. 137)). However, the modified EBA will still rule out blocking by a single leading perpetrator consisting of one internally unstable coalition. In contrast, the r-theory allows blocking regardless of the internal instability of the blocking coalition.

To see the distinction more clearly, consider a modified example suggested by an anonymous referee. Let us change the payoffs for (D, r, R) to $(10.1, 0.1, -0.75)$. Then, (D, r, R) is still the unique Nash equilibrium for the finest coalition structure. But now the internal instability of coalition $\{1, 2\}$ is more serious for player 2, who fears that a break-up will leave him with 0.1. Now there is an EBA for the grand coalition with payoff vector $(1, 1, 13)$, even if we use Vohra's (1997) modified definition of EBA. Indeed, by the logic of EBA, player 2 will not join together with player 1 to block $(1, 1, 13)$, since player 2 fears being "double-crossed" and getting 0.1. In contrast, according to the r-theory the coalition $\{1, 2\}$ will block any agreement in the grand coalition where the sum of their payoffs is less than ten, even in the modified example, because they can sign a binding contract which guarantees them ten. It is true that player 2 might suffer if coalition $\{1, 2\}$ breaks apart. But if player 2 joins the coalition $\{1, 2\}$ only after having obtained player 1's signature on a contract, then there is no downside risk for player 2. Player 2 simply has to hold on to his copy of the contract. By assumption, the binding contract rules out any unilateral defection by player 1. It is especially in cases where a break-up would hurt some coalition member that the possibility of binding contract is crucial. At the very least, a binding agreement would eliminate break-ups that are not supported unanimously. In the modified example, coalition $\{1, 2\}$ will not break up by unanimous consent, so we think it is reasonable to assign the value of ten to this coalition.

The notion that internal instability should not impede the blocking ability of a coalition is also implicit in Perry and Reny's (1994) non-cooperative game. In their model, as in ours, coalitions can sign binding contracts which prevent members from defecting further.

Thus, it is not surprising that a version of Perry and Reny's (1994) game can be used for a non-cooperative implementation of the r-core (see Section 6).

Remark 2. In general, when there are many possible predictions about the behavior of a reduced society, the r-theory assumes a degree of pessimism. Ray and Vohra (1997) assume players are optimistic, in the sense that when a coalition deviates it can name the best equilibrium under the best coalition structure it can induce. In the example in this section, however, there is a unique consistent prediction for each reduced society, so optimism versus pessimism is not an issue. An "optimistic r-core" would also predict $(5, 5, 5)$.

Remark 3. In this example the *r-bargaining set* consists of all payoff vectors that satisfy

$$\pi_1 + \pi_2 + \pi_3 = 15, \qquad \pi_1 \geqslant -9, \qquad \pi_2 \geqslant -9, \qquad \pi_3 \geqslant 0,$$

and either $\pi_1 = \pi_2 \geqslant 5$, $\pi_2 = \pi_3 \geqslant 5$ or $\pi_1 = \pi_3 \geqslant 5$. (For calculations, see (Huang and Sjöström, 1998).) The α-bargaining set is the same as the r-bargaining set.

5. Bertrand competition with differentiated commodities

There are n symmetric price-setting oligopolistic firms, each with constant marginal and average cost equal to zero. Each firm produces a unique product. Firm i's demand function is

$$q_i(p_1, \ldots, p_n) = 1 - p_i - r\left(p_i - \frac{1}{n}\sum_{j=1}^{n} p_j\right),$$

where p_i is the price set by firm i, q_i is the quantity demanded of firm i's product, and $r > 0$ is a parameter of substitutability. The market is more competitive, the greater is r. The goods become perfect substitutes as $r \to \infty$. Notice that for any $r > 0$, the grand coalition N has a unique joint payoff maximizing strategy $\sigma^m = (1/2, 1/2, \ldots, 1/2)$. That is, each firm i should charge the monopoly price $p_i = 1/2$. With monopolistic pricing, the total industry profit is $n/4$. We consider the r-core of this game. Our result is that the monopolistic outcome can be achieved by the grand coalition if and only if $n \leqslant 9$ and r exceeds a critical value $\hat{r}(n)$. The function $\hat{r}(n)$ is increasing in n because it is harder to sustain cooperation when there are more players. Notice that the symmetry of the game implies that transfers will not be of any value in supporting the grand coalition.

For the grand coalition to be sustained, no coalition should profit from a deviation. This means no coalition S should be able to profit by free riding on the collusion of other firms while S itself captures a larger market share. If goods are very close substitutes, then free riding is unattractive because a defection by S triggers severe competition. Thus, when r is large the potential for ruinous competition should make it easier for the firms to collude. Recall that more competitive market conditions also help firms collude in the theory of repeated games (Shapiro, 1989).

Deneckere and Davidson (1985) studied the incentives for mergers in the differentiated Bertrand model, assuming all other non-merging coalitions stay fixed. Notice that this game has "positive externalities" in the sense that a merger of two firms reduces the competition and thus benefits outside firms. Deneckere and Davidson derived the following formulas. Suppose in a market with n firms a merger of size m happens. This leads to a merged entity consisting of these m players, and $n - m$ outside firms playing non-cooperatively. The per member profit of the merged firm is

$$\pi_i^c(m) = \left[\frac{2n + r(2n-1)}{4n + 2r(3n - m - 1) + r^2((n-m)/n)(2n + m - 2)}\right]^2 \times \left[1 + r\frac{n-m}{n}\right] \tag{15}$$

and each outsider earns profit equal to

$$\pi_i^0(m) = \left[\frac{2n + r(2n-m)}{4n + 2r(3n - m - 1) + r^2((n-m)/n)(2n + m - 2)}\right]^2 \times \left[1 + r\frac{n-1}{n}\right]. \tag{16}$$

For any given coalition structure, a merger of two coalitions will benefit all outsiders. Deneckere and Davidson showed that the game also has a superadditive property in the sense that for any given coalition structure, a merger of two coalitions results in a joint after-merger profit for them which is greater than the sum of their pre-merger profits. Thus, they showed:

Proposition 1. *Let $\{B_1, B_2, \ldots, B_k\}$ be a partition of N. Let $\pi(B_i \cup B_j)$ be the profit of $B_i \cup B_j$ after a merger of B_i and B_j occurs* (*and nothing else happens to the coalition structure*) *and let $\pi(B_i)$ denote the pre-merger profits. Then $\pi(B_i \cup B_j) > \pi(B_i) + \pi(B_j)$.*

Proposition 1 implies that it is never efficient for a coalition to break up. (In contrast, in the Cournot model analyzed in Huang and Sjöström (1998) it may be to the advantage of a coalition $S \subset N$ to break apart, as it will reduce the output of the complement.) This implies that if the r-core exists and is non-empty in reduced society S, then coalition S must form. There cannot exist a finer r-core partition $\mathcal{P}_S \neq \{S\}$ because the definition of the r-core implies that the sum of the payoffs for the players in S under the finer partition $\mathcal{P}_S$ would have to at least equal what they could get by sticking together, but this is impossible in view of Proposition 1. This result greatly simplifies the calculation of coalitional values in any reduced society. Deneckere and Davidson also showed that in any given coalition structure, a member of a large coalition earns strictly less than a member of a small coalition:

Proposition 2. *Let $\{B_1, B_2, \ldots, B_k\}$ be a partition of the N with $n_1 \geqslant n_2 \geqslant \cdots \geqslant n_k$, where $n_i = |B_i|$. Then $\pi_i \leqslant \pi_{i+1}$ (with equality if and only if $n_i = n_{i+1}$), where π_i denote the per member profit of coalition B_i.*

This result implies that if the r-core exists, the most difficult blocking constraint will pertain to singleton coalitions. But notice that to verify that the r-core is non-empty, we must verify that every reduced society has a non-empty r-core. In effect, we must verify that in any reduced society S, given any coalition structure $\mathcal{P}_{N\setminus S}$ for the complement, no single firm in S wants to deviate from coalition S if it thinks the complement $S\setminus\{i\}$ will stick together (as we have argued, this is the only belief consistent with the r-core for reduced society $S\setminus\{i\}$). Define $\hat{r}(n)$ as follows:

$$\hat{r}(1)=\hat{r}(2)=\hat{r}(3)=0,\qquad \hat{r}(4)=1.86,\qquad \hat{r}(5)=4.55,\qquad \hat{r}(6)=8.12,$$
$$\hat{r}(7)=12.56,\qquad \hat{r}(8)=19,\qquad \hat{r}(9)=43.75,\qquad \hat{r}(n)=+\infty\quad \text{if } n\geqslant 10.$$

Proposition 3. *The r-core exists and is non-empty in the n-firm Bertrand model with differentiated commodities if and only if $r\geqslant\hat{r}(n)$.*

We can express Proposition 3 in an alternative way: for given level of substitutability r, there exists a critical number $\hat{n}(r)$ (increasing in r, with $\lim_{r\to\infty}\hat{n}(r)=9$) such that the r-core exists if and only if $n\leqslant\hat{n}(r)$.

The proof of Proposition 3 consists of straightforward calculations. In a two-player society, free-riding cannot possibly pay. For $n=3$ we only need to make sure that a single firm will not defect from the grand coalition (assuming the remaining two firms stick together). This is true if and only if what the defector gets, $\pi_i^0(2)$, is no bigger than the per-capita monopoly profit $1/4$. Straightforward calculations show that this holds for all $r>0$. Thus, the r-core exists and is non-empty for any $r>0$ if $n\leqslant 3$. When $4\leqslant n\leqslant 7$, the most difficult constraint to satisfy is that a single firm should not have an incentive to defect from the grand coalition N. Calculations show that there is no such incentive if and only if $r\geqslant\hat{r}(n)$. When $8\leqslant n\leqslant 9$, the most difficult free rider constraint turns out to relate to a single firm in a reduced society with $n-1$ players (recall that for the r-core to exist, it must exist and be non-empty in every reduced society). Again, calculations show that there is no such incentive if and only if $r\geqslant\hat{r}(n)$.

Finally, suppose $n\geqslant 10$. Consider a reduced society S with $|S|=8$ players, and suppose the outside $n-8\geqslant 2$ players are all split into singleton coalitions. If the r-core exists then any firm $i\in S$ who defects from S must expect the complement $S\setminus\{i\}$ to stick together (Proposition 1 implies that it is the only possible belief that is consistent with the r-core). The free riding payoff would be $\pi_i^0(7)$ as given by (16). On the other hand, if the reduced society S forms a coalition of 8 players, given all outside $n-8$ players are singleton coalitions, the per capita payoff for coalition S is $\pi_i^c(8)$ as given by (15). Hence, firm i can block cooperation within the reduced society S if $\pi_i^0(7)>\pi_i^c(8)$, which, by (16) and (15), is equivalent to

$$\begin{aligned}
&\big(2n+r(2n-7)\big)^2\left[1+r\frac{n-1}{n}\right]\left(4n+2r(3n-9)+r^2\left(\frac{n-8}{n}\right)(2n+6)\right)^2\\
&\quad>\big(2n+r(2n-1)\big)^2\left[1+r\frac{n-8}{n}\right]\\
&\qquad\times\left(4n+2r(3n-8)+r^2\left(\frac{n-7}{n}\right)(2n+5)\right)^2. \qquad (17)
\end{aligned}$$

One can calculate that (17) is satisfied if and only if

$$\begin{aligned}&\big(28n^5+84n^4+27538n^2+94983n-7189n^3-103096\big)r^7\\&\quad+\big(420n^5+99498n^2+61271n-1792n^4-27489n^3\big)r^6\\&\quad+\big(1652n^5+69216n^2-7868n^4-28756n^3\big)r^5\\&\quad+\big(2716n^5-10024n^4-8456n^3\big)r^4+\big(2016n^5-4032n^4\big)r^3+\big(560n^5\big)r^2\\&>0.\end{aligned} \tag{18}$$

Given $n \geqslant 10$, all polynomials in the parentheses in (18) are positive. Therefore, for any $r > 0$, (18) holds, so $\pi_i^0(7) > \pi_i^c(8)$. Hence there is no outcome in the r-core for the 8-person reduced society S, so no characteristic function can be defined. Hence, if $n \geqslant 10$ then the r-core does not exist for any r.

The negative result for $n \geqslant 10$ may be surprising since when r is large the firms have strong incentives to merge into a monopoly, and any defection from the grand coalition triggers cut-throat competition. So one may have expected cooperation to be possible for sufficiently large r. But while it is true that a firm can never benefit from defecting from the grand coalition when r is sufficiently large, the key point is that for the r-core to exist and be non-empty it must exist and be non-empty for all coalition structures. When $n \geqslant 10$, then (as we just argued) the r-core is empty for an 8-player reduced society S given that the complement $N\backslash S$ is divided into singletons. This can be explained as follows. The 8-player coalition S maximizes its joint profit by sticking together. However, even if they stick together there will be at least three competing firms (at least two singleton coalitions in $N\backslash S$, plus the merged entity S) playing non-cooperatively, which is enough to drive profits down close to zero when r is large. Thus, there is not much surplus to be distributed among the members of S. A firm $i \in S$ that considers a defection from S has to compare the payoff it gets in two very competitive situations, one where it has defected (and $S\backslash\{i\}$ sticks together) and one where it has joined S in competition with the outside firms. In fact, (18) shows that the surplus available to S is not big enough to prevent firm i from defecting. Even though free riding on the grand coalition N does not pay when the market is very competitive, it *does* pay to free ride on large subcoalitions when most of the profits are anyway destroyed by competition with outsiders. For this reason, the r-core may not exist even when r is very large.

In contrast, for any n there exists $\bar{r}(n) < \infty$ such that an EBA exists for the grand coalition whenever $r > \bar{r}(n)$. Indeed, an equal share of the monopoly profit is worth $1/4$, regardless of r. If firm i defects from the grand coalition, the best it can hope for is that the complement $N\backslash\{i\}$ stays together, but even in this best of all possible cases firm i's profit is close to zero for r large enough. For even if there are only two firms, $\{i\}$ and the composite firm $N\backslash\{i\}$, profits are almost completely dissipated if the goods are very close substitutes. Thus, no leading perpetrator can possibly gain from a defection if r is large enough, so σ^m (together with zero transfer payments) is an EBA. Thus, there can exist an EBA for the grand coalition even though the r-core does not exist (in particular, this is true when $n \geqslant 10$). On the other hand, in this application any EBA payoff for the grand coalition belongs to the α-core, which is always non-empty (cf. (Ray and Vohra, 1997, Remark 6.1)).

Notice that, unlike the example in Section 4, the difference between the EBA and the r-core has nothing to do with the internal stability of blocking coalitions, since the argument here depends completely on blocking by singletons (which are by definition internally stable). Instead, it depends on our insistence that the r-core exists for every possible reduced society. Again, this insistence has a natural counterpart in the subgame perfect equilibrium of Perry and Reny's (1994) non-cooperative model, where a continuation equilibrium must exist for every possible subgame.

If the r-core does exist and is non-empty in this application, then there exists an EBA for the grand coalition. For if the r-core is not empty then clearly it contains σ^m together with a zero net transfer profile (by symmetry). By definition of the r-core, no coalition $S \subset N$ can block this outcome. That is, each S gets at least $V(S)$, where $V(S)$ is calculated under the expectations that the complement $N \backslash S$ sticks together (this is the only possible belief about the r-core outcome in reduced society N/S once S has left, in view of Proposition 1). But then it is clear that σ^m together with a zero net transfer profile is also an EBA. Indeed, the best any leading perpetrator S can hope for is for the complement to stick together (its payoff decreases when the complement breaks up), but not even this situation makes S better off if the r-core is non-empty. Hence, we have:

Proposition 4. *In the differentiated Bertrand model, if the r-core exists and is non-empty then σ^m together with a zero net transfer profile is an equilibrium binding agreement for the grand coalition.*

This result does not say anything about the complete set of r-core payoff vectors. In fact, when the r-core exists the most difficult blocking constraint pertains to singleton coalitions, and both the r-core and the EBA payoffs for the grand coalition have to satisfy this constraint. Thus, in this application, any r-core payoff is an EBA payoff for the grand coalition, while on the other hand (as argued above) an EBA can exist even if the r-core does not exist.

If we use the Zhou bargaining set as the presolution instead of the core, we get exactly the same result: the cutoff threshold $\hat{r}(n)$ is also the threshold for the r-bargaining set to be non-empty. This is so because the binding constraints relate to defections by singleton coalitions for which there is no difference between the core and bargaining set.

6. Non-cooperative implementation

Perry and Reny (1994) studied the non-cooperative implementation of the core for a given characteristic function V. In their model, proposals to form a coalition are made anonymously and consist of a coalition S and a set of payoffs $(x_j)_{j\in S}$ such that $\sum_{j\in S} x_j \leqslant V(S)$. Time is continuous. At any time $\tau \geqslant 0$ a player can make a proposal, accept a current proposal, be quiet, or leave. If a current proposal $((x_j)_{j\in S}, S)$ is accepted by all members of S, then it becomes *binding* and coalition S is said to have formed. In this case, any player $i \in S$ can choose to "leave" and consume x_i, in which case every other member $j \in S$ must also "leave" and consume x_j. If S has formed, then a new proposal cannot be directed to a non-empty strict subset of S, but (as long as the members of S have not "left") a proposal can be directed to a *superset* $T \supseteq S$. If a new proposal is made *before* the current proposal has become binding, then the current proposal is void. The payoff to a player who never leaves is low enough that, following any history, any continuation equilibrium will be such that all players will leave in finite time. (Player i cannot leave before accepting a proposal, but he can always propose $(x_i, \{i\})$ where $x_i = V(\{i\})$, accept, and then leave and consume $V(\{i\})$.) Perry and Reny make the technical assumption that for any time $\tau > 0$ and for any history up to τ, there is $\varepsilon > 0$ such that each player i is quiet in the open intervals $(\tau - \varepsilon, \tau)$ and $(\tau, \tau + \varepsilon)$. They consider stationary subgame perfect equilibria (SSPE), where a player's action can only depend on the set of players remaining, the existing set of binding proposals, the current proposal, and which players have accepted it. They show that only core payoff vectors can be SSPE payoff vectors. We will consider a slightly modified version of Perry and Reny's game which is appropriate in the presence of externalities.

The modification we need is that a proposal to form coalition S cannot specify payoffs directly. Instead, a proposal $((\alpha_i)_{i\in S}, S)$ specifies a vector of non-negative *shares* $(\alpha_i)_{i\in S}$ that sum to one:

$$\sum_{i\in S} \alpha_i = 1. \tag{19}$$

If S forms and player $i \in S$ decides to "leave," then the members of S must "sit on the side line" until all players in N have "left." Suppose the last player leaves at time τ^* (we will have $\tau^* < \infty$ in equilibrium since never leaving gives a very low payoff by assumption). Then at time τ^*, each player is a member of a unique coalition, i.e., there is a partition $\mathcal{P}_N = \{S_1, \ldots, S_K\}$ of N. Then the game $\Gamma(\mathcal{P}_N)$ is played. For each $S \in \mathcal{P}_N$, a member of S whose share is strictly positive chooses $\sigma_S \in \Delta(A_S)$ on behalf of his coalition. If the resulting strategy profile is σ, then each player i's final payoff is $\alpha_i \sum_{j\in S} \pi_j(\sigma)$, where $\alpha_i \geqslant 0$ is player i's share of the payoff of the coalition of which he is a member ($i \in S$). Notice that any player with $\alpha_i > 0$ is motivated to maximize the joint payoff $\sum_{j\in S} \pi_j(a)$. Thus, in subgame perfect equilibrium, $\sigma = (\sigma_{S_1}, \ldots, \sigma_{S_K}) \in E(\mathcal{P}_N)$.

If a proposal $((\alpha_i)_{i\in S}, S)$ becomes binding, then either S will be a part of the final coalition structure, or S will by unanimous consent be absorbed into some larger coalition $T \supset S$. The rules do not allow defections from the binding proposal by any strict subcoalition of S. If the proposal $((\alpha_i)_{i\in S}, S)$ is binding and player $i \in S$ "leaves," then he cannot consume immediately because his final payoff $\alpha_i \sum_{j\in S} \pi_j(\sigma)$ will depend on the final partition $\mathcal{P}_N$ (that is, $\sigma \in E(\mathcal{P}_N)$). What he can guarantee by leaving is that coalition S will be one element of the final partition $\mathcal{P}_N$, and that he will get a share α_i of the coalition's payoff. Although player $i \in S$ does not know his final payoff until he knows which *other* coalitions form, he does not fear that some partners in S will "double-cross" him by defecting from S, because the rules do not allow such destabilizing deviations from a binding proposal.

For this modified game we can reprove Theorem 1 of Perry and Reny (1994), with r-core replacing core, essentially using their arguments. For the sake of exposition we restrict attention to the case $n = 3$. Notice that in any three person game, $V(S)$ is well-defined for all $S \subseteq N$, because $C(N\backslash S \mid N\backslash S, S) \neq \emptyset$. This is clearly true for the grand coalition or any two-person coalition S, and for any singleton coalition $S = \{i\}$ it is also true because the remaining two players in $N\backslash\{i\}$ either stay together or break up. Hence, we can always define $V(S)$ when $n = 3$.

Proposition 5. *Suppose* $n = 3$. *In the modified Perry–Reny game, only r-core payoff vectors can be SSPE payoff vector.*

Proof. Fix an SSPE of the modified Perry–Reny game and $x = (x_1, x_2, x_3)$ be the corresponding payoff vector. We will show that x must be in the r-core. If it is not, then there is $S \subseteq N$ such that $\sum_{i \in S} x_i < V(S)$, where $V(S)$ is the worth of coalition $S \subseteq N$ calculated according to the r-core theory. There are three possibilities for S, and each will yield a contradiction.

Case 1. $|S| = 2$, say $S = \{1, 2\}$. Suppose player 1 deviates from the SSPE by proposing $((\alpha_1, \alpha_2), \{1, 2\})$ such that $\alpha_1 V(\{1, 2\}) > x_1$ and $\alpha_2 V(\{1, 2\}) > x_2$ and $\alpha_1 + \alpha_2 = 1$. Such a proposal is possible since $x_1 + x_2 < V(\{1, 2\})$ by hypothesis. Moreover, by assumption there is $\varepsilon > 0$ such that no proposal is accepted in the interval $[0, \varepsilon]$, and any player can deviate by making a proposal before time ε. Let player 1 immediately accept his own proposal. If player 2 also accepts, then player 1 can guarantee that the final coalition structure is $(\{1, 2\}, \{3\})$ in which case his payoff is at least $\alpha_1 V(\{1, 2\}) > x_1$. This is so because $V(\{1, 2\})$ is by definition the minimum that coalition $\{1, 2\}$ gets in coalition structure $(\{1, 2\}, \{3\})$, and α_1 is player 1's share of it. So to support the SSPE, player 2 must not accept this proposal, for if he accepts it then the deviation makes player 1 better off. By accepting, however, player 2 can guarantee that the final coalition structure is $(\{1, 2\}, \{3\})$ which gives him at least $\alpha_2 V(\{1, 2\}) > x_2$. So, if player 2 does not accept then he must expect that some later proposal will give him even more than $\alpha_2 V(\{1, 2\})$. But then he could *himself* have deviated from the original SSPE by making that very proposal and obtained a payoff greater than x_2, which is a contradiction. This argument, which follows Perry and Reny (1994), relies on the assumption that proposals are anonymous and strategies stationary so it does not matter who makes a particular proposal. Thus, we have shown that if $x_1 + x_2 < V(\{1, 2\})$ then x is not an SSPE payoff vector.

Case 2. $S = N$. The argument is similar to Case 1.

Case 3. $S = \{i\}$, say $S = \{1\}$. Then $x_1 < V(\{1\})$. Suppose player 1 deviates from the SSPE by proposing and accepting the coalition $\{1\}$ and then "leaving" at time τ very close to zero. By doing so he irrevocably commits not to cooperate with players 2 and 3. Players 2 and 3 remain at time τ without any current proposal on the table. Now, the subgame that starts at time τ must either end with the formation of $\{2, 3\}$, or by the formation of two singletons, $\{2\}$ and $\{3\}$. By the same argument as in Case 1, the continuation equilibrium must agree with the r-core for the reduced society $\{1, 2\}$, i.e., the outcome must be in $C(\{2, 3\} \mid \{2, 3\}, \{1\})$. However, since $V(\{1\})$ is player 1's minimal payoff among all the strategy vectors in $C(\{2, 3\} \mid \{2, 3\}, \{1\})$, the deviation by player 1 will give him at least $V(\{1\}) > x_1$, which is a contradiction. Thus, we have shown that if $x_1 < V(\{1\})$ then x is not an SSPE payoff vector. □

Acknowledgments

We are grateful to Kalyan Chatterjee, C.Y. Cyrus Chu, Jerry Green, Eric Maskin, Debraj Ray, and two anonymous referees for helpful comments.

References

Aumann, R., 1989. Game theory. In: Eatwell, J., Milgate, M., Newman, P. (Eds.), The New Palgrave: A Dictionary of Economics. Macmillan, New York.
Bloch, F., 1995. Endogenous structures of association in oligopolies. RAND J. Econ. 26, 537–556.
Deneckere, R., Davidson, C., 1985. Incentives to form coalitions with Bertrand competition. RAND J. Econ. 16, 473–486.
Huang, C.-Y., Sjöström, T., 1998. The p-core. Working Paper, Penn State University.
Ichiishi, T., 1981. A social coalitional equilibrium existence lemma. Econometrica 49, 369–377.
Luce, R.D., Raiffa, H., 1957. Games and Decisions. Wiley, New York.
von Neumann, J., Morgenstern, O., 1944. Theory of Games and Economic Behavior. Princeton Univ. Press, Princeton.
Perry, M., Reny, P., 1994. A noncooperative view of coalition formation and the core. Econometrica 62, 795–817.
Ray, D., 1989. Credible coalitions and the core. Int. J. Game Theory 18, 185–187.
Ray, D., Vohra, R., 1997. Equilibrium binding agreements. J. Econ. Theory 73, 30–78.
Ray, D., Vohra, R., 1999. A theory of endogenous coalition structures. Games Econ. Behav. 26, 286–336.
Shapiro, C., 1989. Theories of Oligopoly Behavior. In: Handbook of Industrial Organization, Vol. I. North-Holland, Amsterdam.
Shubik, M., 1983. Game Theory in the Social Sciences. Concepts and Solutions. MIT Press, Cambridge, MA.
Vohra, R., 1997. Coalitional non-cooperative approaches to cooperation. In: Hart, S., Mas-Colell, A. (Eds.), Cooperation: Game Theoretic Approaches. Springer-Verlag, Berlin.
Yi, S.-S., 1997. Stable coalition structures with externalities. Games Econ. Behav. 20, 201–237.
Zhou, L., 1994. Bargaining set and coalition formation. Games Econ. Behav. 6, 512–526.