THE CORE IN NORMAL FORM GAMES

LÁSZLÓ Á. KÓCZY

Abstract. Due to the externalities, in normal form games a deviation changes the payoff of all players inducing a retaliation by the remaining or residual players. The stability of an outcome depends on the expectations potential deviators have about this reaction, but so far no satisfactory theory has been provided. The present paper continues the work of Chander and Tulkens (1997) where deviators consider residual equilibria, but we allow coalitions to form, moreover introduce consistency between the residual solution and the solution of the original game. Optimistic and pessimistic considerations produce a pair of cores. These cores are compared to some existing cooperative concepts such as the γ- and r-cores and the equilibrium binding agreements. In our final section we discuss the predominance of the grand coalition and suggest a generalisation of the normal form where such a precedence can be removed.

1. Introduction

Its intuitive, straightforward definition makes the core one of the most popular solution concepts in coalition formation games. Further, Peleg (1992) claims that a solution is "acceptable" only if its axiomatisation is similar to that of the core. The original definition, however, does not account for externalities, but assumes that a deviating coalition gets the coalitional payoff given by a characteristic function.

In a more general setting, such as the normal form we consider here, the payoff is also a function of the others' behaviour. Since such externalities typically go in both directions, a deviation will influence (that is: change) others' payoffs prompting a response that may in turn lead to secondary reactions. Since in this paper we are not concerned with the detailed mechanisms of coalition formation, but take a cooperative approach, the deviating players can only realise payoffs once the reaction is known. The residual reaction and the profitability of the deviation are closely linked. Consequently an accurate modelling of the reaction of the residual players is essential to understand domination

Discussions with Johan Eyckmans, Luc Lauwers, Carmen Bevía, Prakash Chander and Henry Tulkens lead to considerable improvement of the paper. The paper further benefited from presentations of an earlier version at the 7th Coalition Theory Workshop in Venice, at PET02 in Paris, and at ICM2002GTA, Qingdao, where participation was made possible by a grant from the Royal Economic Society. The joint support (SOR-H/99/23, SOR-H/00/01 and COE/01/02) of the Katholieke Universiteit Leuven and the Soros Foundation is gratefully acknowledged.

and thence the core. The present paper contributes a model where residuals are treated consistently.

In the following we briefly overview the development of the deviators' expectations. The characteristic function form was defined by (von Neumann and Morgenstern 1944) from the normal form using the minimax rule. This definition equips deviators with excessive pessimism: they expect that residuals hurt them as much as possible even counter their interests. Much of the literature still uses such and similar models (such as extreme optimism). The recent surge of interest in games with externalities has led to a reconsideration of this definition.

While the noncooperative literature, such as Ichiishi (1981) in his *social coalitional equilibrium* ignores potential reactions citing simultaneity, in cooperative games, where agreements are made collectively a deviation cannot go unnoticed. Harsányi (1963) introduced a model where deviators and residuals (cooperatively) bargain with each other and so the interests of the latter are also represented. Here residuals act as one coalition. In the model of Chander and Tulkens (1994) and Chander and Tulkens (1997) the residuals play *individually rational* strategies, that is, a Nash equilibrium is reached against the deviating players, the residuals acting as singletons. The present paper generalises these approaches by permitting residual players to form (intermediate) coalition structures and thus pursue group-wise rational strategies as well.

What is common to most of these models is that they solve the normal form game essentially via a –however sophisticatedly defined– characteristic function form game. We do not aim to introduce a new way to define a characteristic function, but to apply the core *directly* to normal form games.

More recently Ray and Vohra (1997) defined the set of the coarsest equilibrium binding agreements and Huang and Sjöström (2001) the r-core. The set of equilibrium binding agreements is a farsighted concept, where only credible, refining deviations are permitted. The r-core is more general than that; its limitation is only that it solves games via a –though sophisticated– characteristic function that is seldom defined. Non-trivial games exist with non-empty cores where the r-core is undefined.

The concept we define allows for arbitrary deviations and includes some limited farsighted flavour while being a fundamentally myopic concept; it offers the same rational behaviour to residuals as to deviators.

The structure of the paper is as follows. The next part contains some fundamental definitions. Section 3 presents some criticisms of the existing theories as a motivation for the introduction of a new concept. The fourth section contains the actual definitions and some comparisons with other concepts on the market. Section 5 closes with the ambitious plan to generalise the normal game form allowing for games where the grand coalition has a more limited power and where our cores (after small modifications) can return some –mathematically– much more diverse results.

2. Preliminaries

2.1. **Normal form games.** Let N be the set of players. Subsets of N are called *coalitions*. A partition $\mathcal{P}$ of N is a breaking up of N into disjoint coalitions. $\Pi(S)$ is the set of partitions of any set S of the players. Let also X_i denote the strategy set for player i and $u_i : \prod_{i\in N} X_i \to \mathbb{R}$ be a *payoff function* for player i. We call $\Gamma = \big(N, (X_i, u_i)_{i\in N}\big)$ a *game in normal form* or simply a *game*.

For any coalition $S \subseteq N$ let X_S denote $\prod_{j\in S} X_j$ and let

$$X_{-S} = \prod_{j\in N\setminus S} X_j.$$

We will also use X to denote X_N. For any $x \equiv (x_i)_{i\in N} \in X$ and $S \subseteq N$ we use x_S to denote $(x_j)_{j\in S}$ and if $N \setminus S$ is nonempty x_{-S} to denote $(x_j)_{j\in N\setminus S}$. Also, for $x \in X$ and $S \subseteq N$ we denote $\sum_{i\in S} u_i(x)$ by $u_S(x)$. We will economise our notation by writing the element instead of a singleton set, such as i instead of $\{i\}$ or R instead of $\{R\}$, in particular $-i$ stands for $N \setminus \{i\}$.

The game is a *hybrid game* in the sense of Zhao (1992) and is played in two stages. First players form coalitions, then these coalitions, as superplayers play non-cooperatively against each other. We focus on the first, cooperative part and assume that cooperation within coalitions also includes transfers between players of the coalition. The concept we introduce can be extended to the more general non-transferable utility case as well.

For a given partition $\mathcal{P}$ the transfers can be described by a *transfer scheme* $(t_i)_{i\in N}$ where $t_i \in \mathbb{R}$ for all i and we restrict transfers to coalitions by requiring $\sum_{j\in S} t_j = 0$ for all $S \in \mathcal{P}$. The actual payment player i receives is its *profit* π_i, given by $\pi_i(x, t, \mathcal{P}) = u_i(x) + t_i$. It will, in general, be different from its payoff, although it is clear that for coalitions we must have $u_S(x) = \pi_S$, where $\pi_S = \sum_{i\in S} \pi_i$. A player's aim is to maximise its profit π_i.

Finally, an *outcome* is a state of the game summarised by a triple consisting of a profit and a strategy vector and the partition of the players, $a = (\pi, x, \mathcal{P})$.

2.2. **Characteristic function form games.** In characteristic function form games players do not make their own decisions nor receive individual payoffs. Their freedom is limited to joining a coalition (the singleton coalition consisting of itself included). Then the coalition is assigned a payoff as given by the characteristic function and independently of the formation of other coalitions.

Definition 2.1 (Characteristic function)**.** The characteristic function assigns a real number to each coalition:

$$v:\ 2^N \longrightarrow \mathbb{R} \tag{2.1}$$
$$S \longmapsto v(S). \tag{2.2}$$

A game in the characteristic function form is a pair (N, v).

In the following we give the definition of the core and the coalition structure core.

Imputations are profit vectors that satisfy $\pi_i \geq v(i)$ (individual rationality) and $x = v(N)$ (efficiency).

An imputation π is dominated via coalition S if $v(S) > \pi_S$. Members of the dominating coalition S benefit from forming S and leaving the grand coalition, which is often regarded as the socially desirable outcome (especially if superadditivity is assumed). Imputation π is dominated if it is dominated via some coalition $S \subseteq N$. The *core* collects undominated imputations.

In characteristic function games an *outcome* is a pair $(\pi, \mathcal{P})$ that satisfies $\pi_i \geq v(i)$ (individual rationality) and $\pi_S = v(S)$ for all $S \in \mathcal{P}$ combining *feasibility* and *efficiency*. Outcomes generalise imputations as the formation of the grand coalition is not presumed.

An outcome $(\pi, \mathcal{P})$ is dominated via coalition S if $v(S) > \pi_S$. Observe that $S \notin \mathcal{P}$. The outcome $(\pi, \mathcal{P})$ is dominated if it is dominated via some coalition $S \subseteq N$. The *coalition structure core* collects the undominated outcomes (Greenberg 1994).

3. Solving the game

The game consists of two stages and the solution of the two stages is presented independently. We require the solutions to be subgame perfect: we may solve the second stage independently of the first, but in solving the first we must look into the consequences in the second stage. That is, first we must establish the solution of the non-cooperative stage and only then are we able to solve the cooperative stage.

3.1. **Noncooperative stage.** Let us assume that a coalition structure $\mathcal{P}$ has already formed. To solve this stage we use the *best-response property* Ray and Vohra (1997) have introduced, essentially a generalisation of the Nash-equilibrium for coalitions that act as players. Although Ray and Vohra (1997) provide a direct definition, here it is more suitable to us to present it by first considering individual best responses. Thus our definition follows the logic of the best-response equilibrium for extensive form games (Kuhn 1953).

Definition 3.1 (Best response of a coalition)**.** The strategy $x^*_S \in X_S$ is a best response for coalition S if

$$u_S(x^*_S, x_{-S}) = \max_{x_S \in X_S} u_S(x_S, x_{-S}),$$

where $x_{-S} \in X_{-S}$.

We assume that the payoff function satisfies a non-levelness property, that is, the best responses are unique. While this assumption is by no means essential to our results, it simplifies the arguments and removing it poses no difficulty.

The above definition is generalised directly to more then one coalition.

Definition 3.2 (Best response of a set of coalitions)**.** The strategy $x^*_{\mathcal{S}} \in X_{\mathcal{S}}$ is a best response for the partition $\mathcal{S} \in \Pi(S)$ for $S \subseteq N$ if it is a best response for all coalition $C \in \mathcal{S}$.[1]

The case when $S = N$ deserves special attention.

Definition 3.3 (Best response property)**.** The strategy vector $x \in X$ satisfies the best response property relative to $\mathcal{P}$ if it is a best response to itself for all coalitions S in $\mathcal{P}$.

This definition is equivalent to the following: The strategy vector $x \in X$ satisfies the best response property relative to $\mathcal{P}$ if for each coalition $S \in \mathcal{P}$ and $x'_S \in X_S$ we have $u_S(x'_S, x_{-S}) \leq u_S(x)$. Let $\beta(\mathcal{P})$ denote the set of *best response strategy profiles*.

The existence of such equilibria is well known under certain conditions that mostly hold.

3.2. **Cooperative stage.** Knowing what coalition formation yields we solve the first, cooperative stage. Since we look for a core-like concept collecting undominated outcomes we must know what a coalition can expect by deviating. Normal form games are more complex than characteristic form games in two ways: We may have multiple best-response strategy profiles in the second stage even for a given partition, and secondly a deviation is followed by a retaliation by the residuals due to the widespread externalities present. This retaliation may then prompt secondary and even further reactions.

Unlike Ray and Vohra (1997) we consider myopic players. How can deviators predict the end outcome, in particular, their terminal payoff? The mathematically appealing simplest approach is conservatism: indeed the characteristic function of a coalition has originally been defined by von Neumann and Morgenstern (1944) by the *minimax representation* as

$$v(S) = \min_{x_{-S} \in X_{-S}} \max_{x_S \in X_S} u_S\left(x_S, x_{-S}\right)$$

[1]In particular the responses must be the best to each other.

Thus deviating players are assuming the worst: A coalition deviates only if in all cases, with all possible reactions it will be better off. However "all cases" may very unrealistic ones, too. Using the words of Ray and Vohra, why should we expect that residual players act in such a bloodthirsty fashion as to hurt deviators to the maximum extent? Excessive pessimism, or, for that matter, excessive optimism can be very misleading and are subject to significant improvements.

3.3. **The $\gamma-$core.** Chander and Tulkens (1997) took a major step in refining these approaches. Their key idea is that non-equilibrium residual strategy-profiles are subject to further deviations, and hence will necessarily be abandoned. Therefore should the chain of reactions stop it cannot stop at such instable outcomes. Therefore the deviating players should only consider residual equilibria as possible reactions. Chander and Tulkens (1994) picked Nash equilibria, or in their terminology, individually rational behaviour.

Definition 3.4 (Dominance with individually rational residual reaction)**.** An outcome $a = (\pi, x, \mathcal{P})$ is dominated via coalition S if in the partition $\mathcal{P}_S$ consisting of S and singletons there exists a strategy vector $x' \in X$ that satisfies

$$u_S(x') > \pi_S, \text{ moreover} \tag{3.1}$$

$$u_S(x') \geq u_S(x''_S, x'_{-S}) \text{ for all } x''_S \in X_S \text{ and} \tag{3.2}$$

$$u_i(x') \geq u_i(x''_i, x'_{-i}) \text{ for all } i \text{ in } N \setminus S \text{ and } x''_i \in X_i. \tag{3.3}$$

If a is dominated via coalition S, then there exists a profit vector $\pi' \in \mathbb{R}^N$ such that $\pi'_S = u_S(x')$ and $\pi'_i \geq \pi_i$ for all $i \in S$ moreover there exists a player $i \in S$ such that $\pi'_i > \pi_i$. Then we can say that outcome $b = (\pi', x', \mathcal{P}_S)$ dominates a via S.

This definition then enables us to define a characteristic function and the core of the corresponding characteristic function form game, broadly known as the $\gamma-$core (Chander and Tulkens 1997).

Note the hidden optimism in this definition: the coalition is willing to abandon a if there exists a better outcome b even though the formation of b is not guaranteed: a different, less advantageous Nash equilibrium may equally likely occur.

While in this model deviators are more realistic about the residual reaction, they presume that no other coalitions form. In the following concept we relax this assumption.

4. The core of a normal form game

4.1. **Need for a more general solution.** A new concept should satisfy the following:

- Solution in the original game form.
- Domination/deviation is for once and all without renegotiation.

- Following Chander and Tulkens (1994) only look at residual equilibrium outcomes.
- Consistency of global and residual equilibrium concepts including equal treatment of deviating and residual players.
- Pareto-efficient.

The last two points need elaboration.

4.1.1. *Pareto efficiency.* We insist on the Pareto-efficiency of core outcomes. For that we need to allow multi coalition deviations.

Most existing concepts exclude the possibility that more then one coalition deviates at the same time. In games without externalities multi-coalition deviations reproduce as a sequence of single-coalition deviations and as long as "undominance" is concerned, a deviation by a set of coalitions or by only one of them makes no difference.

As we introduce externalities this feature changes. The appendix shows a 4-player game where for the partitions $\mathcal{P}_0 = \{1, 2, 3, 4\}$, $\mathcal{P}_1 = \{\{1, 2\}, 3, 4\}$, $\mathcal{P}_2 = \{1, 2, \{3, 4\}\}$ and $\mathcal{P}_3 = \{\{1, 2\}, \{3, 4\}\}$ have both unique Nash equilibria giving respectively 2, 4, 4 and 6 to each coalition. While $\mathcal{P}_3$ Pareto-dominates $\mathcal{P}_3$ it cannot be attained via a sequence of profitable deviations as the formation of $\{1, 2\}$ or $\{3, 4\}$ is not profitable, moreover at the same time the remaining players (still singletons) enjoy 4 each, so for them forming the other pair is disadvantageous. In this paper we require Pareto domination to imply domination, and for this reason we permit multi-coalition deviations.

4.1.2. *Residual behaviour.* The number of different approaches to model the residual behaviour indicates how difficult this task actually is. In order to make progress one often had to make compromises and simplify the model by additional –often arbitrary– assumptions. While these assumptions may hold in specific situations, they limit the overall applicability of the model. In designing the following model the idea is exactly not to introduce new ideas, not to make new assumptions, but to build from the bricks we already have: deviations induce residual subgames of a more limited strategy space and these are solved using the same concept as the large game. This results in a recursive definition.

We begin by defining this residual game, a function of the deviating partition.

Definition 4.1 (Residual game)**.**

$$\Gamma^s = \left(R, \left(X_i, u_i^{\mathcal{S}} \right)_{i \in R} \right), \tag{4.1}$$

where R is the residual set, $\mathcal{S} \in \Pi N \setminus R$ is the partition of the deviating players, and $u_i^{\mathcal{S}}$ is the *residual payoff function* that we define as follows for all $x \in X_R$.

$$u_i^{\mathcal{S}}(x) = u_i(x, x^*), \tag{4.2}$$

where $x^* \in X_{N\setminus R}$ is a best response strategy profile for all coalitions in $\mathcal{S}$.

The residual game is a normal form game like the initial game, with a trivial, but crucial difference that $|R| < |N|$. This enables us to define the core in a recursive way.

Our definition makes use of a simplification: the deviating coalitions play the best response to the residuals' strategies. This appears to give them an advantage over the residual players. However, a point we make is that the residual game is played as a fully independent game, that is, with both the first and the second stage played. When defining best response equilibria in the second stage, we will however already consider the externalities of $\mathcal{S}$ playing the best response all the time, so that any best response in the residual game is also a best response to that. This removes for the aforementioned advantage. Consequently residual best response equilibria embed into best response equilibria in the original game, when complemented with the best responses of the deviating players.

4.1.3. *Behavioural assumptions.* Deviating players consider a set of residual equilibria that have different implications on their payoffs. Without introducing further considerations, such as efficiency or a preference for a certain coalition structure we cannot make a selection of these equilibria. We use optimism and pessimism on the side of the deviating players to consider the extreme cases, but here the core shows a reduced sensitivity to behavioural assumptions.

4.2. **Definitions.** The definitions are inductive and are done in four steps each. For a trivial single-player game we can give the core explicitly. Given the definition for all at most $k-1$ player games we can give our definition of dominance for k player games. Once dominance is defined, we may define the core. First we give the definition for the pessimistic case, and then a slightly modified version comes for the optimistic core.

Definition 4.2 (Core - pessimistic case)**.** The definition consists of four steps:

Step 1. The core of a trivial game.

Let $\left(N, (X_i, u_i)_{i\in N}\right)$ be a game. The core of a game Γ_1 with $N = \{1\}$ is given by the payoff-dominant strategy of this single player:

$$C_-(\Gamma_1) = \left\{(\pi, x, \mathcal{P}) \,\middle|\, \mathcal{P} = \{1\},\ x \in \beta(\mathcal{P}),\ \pi = \pi_1 = u_1(x)\right\} \tag{4.3}$$

Step 2. Inductive assumption.

We assume that the core has been defined for all games with at most $k-1$ players. Now we give the definition for a game with k players.

Step 3. Dominance

Consider the outcome $(\pi, x, \mathcal{P})$ and a deviation $\mathcal{S} \in \Pi(S)$ by $S \subseteq$

N and the corresponding residual game $\Gamma^{\mathcal{S}}$ over the residual set R. Let $r^* \in X_S$ denote the best response of the deviating coalitions to a residual strategy vector $r \in X_R$

The outcome $(\pi, x, \mathcal{P})$ is dominated via $\mathcal{S} \in \Pi(S)$ if in the corresponding residual game Γ^s over the residual set R either:

(1) the residual core $C_- (\Gamma^s)$ is empty and for all strategy profiles $r \in X_R \cap \beta(\mathcal{R})$ for some $\mathcal{R} \in \Pi(R)$ we have $u_T(r^*, r) > \pi_T$ for all $T \in \mathcal{S}$, or
(2) the residual core $C_- (\Gamma^s)$ is not empty and for all $(\pi', r, \mathcal{Q}) \in C_- (\Gamma^s)$ we have that r satisfies $u_T(r^*, r) > \pi_T$ for all $T \in \mathcal{S}$.

An outcome $(\pi, x, \mathcal{P})$ is dominated if it is dominated via a set of coalitions.

Step 4. Core

The core of a game of k players is the set of undominated outcomes and we denote it by $C_- \left(N, (X_i, u_i)_{i\in N}\right)$.

Definition 4.3 (Core - optimistic case)**.** The definition consists of four steps:

Step 1. The core of a trivial game.

Let $\left(N, (X_i, u_i)_{i\in N}\right)$ be a game. The core of a game Γ_1 with $N = \{1\}$ is given by the payoff-dominant strategy of this single player:

$$C_+ (\Gamma_1) = \left\{(\pi, x, \mathcal{P}) \;\middle|\; \mathcal{P} = \{1\},\ x \in \beta(\mathcal{P}),\ \pi = \pi_1 = u_1(x)\right\} \tag{4.4}$$

Step 2. Inductive assumption.

We assume that the core has been defined for all games with at most $k-1$ players. Now we give the definition for a game with k players.

Step 3. Dominance

Consider the outcome $(\pi, x, \mathcal{P})$ and a deviation $\mathcal{S} \in \Pi(S)$ by $S \subseteq N$ and the corresponding residual game $\Gamma^{\mathcal{S}}$ over the residual set R. Let $r^* \in X_S$ denote the best response of the deviating coalitions to a residual strategy vector $r \in X_R$

The outcome $(\pi, x, \mathcal{P})$ is dominated via $\mathcal{S} \in \Pi(S)$ if in the corresponding residual game Γ^s over the residual set R either:

(1) the residual core $C_- (\Gamma^s)$ is empty and for all strategy profiles $r \in X_R \cap \beta(\mathcal{R})$ for some $\mathcal{R} \in \Pi(R)$ we have $u_T(r^*, r) > \pi_T$ for all $T \in \mathcal{S}$, or
(2) the residual core $C_- (\Gamma^s)$ is not empty and for all $(\pi', r, \mathcal{Q}) \in C_- (\Gamma^s)$ we have that r satisfies $u_T(r^*, r) > \pi_T$ for all $T \in \mathcal{S}$.

(1) the residual core $C_- (\Gamma^s)$ is empty and there exists a strategy profile $r \in X_R \cap \beta(\mathcal{R})$ for some $\mathcal{R} \in \Pi(R)$ such that $u_T(r^*, r) > \pi_T$ for all $T \in \mathcal{S}$, or
(2) the residual core $C_- (\Gamma^s)$ is not empty and there exists an outcome $(\pi', r, \mathcal{Q}) \in C_- (\Gamma^s)$ such that r satisfies $u_T(r^*, r) > \pi_T$ for all $T \in \mathcal{S}$.

An outcome $(\pi, x, \mathcal{P})$ is dominated if it is dominated via a set of coalitions.

Step 4. Core

The core of a game of k players is the set of undominated outcomes and we denote it by $C_+\left(N, (X_i, u_i)_{i\in N}\right)$.

4.3. **Interpretation.** The induction relies on two simple facts: the obvious definition for a trivial one-person game, and on the aforementioned reduction of residual games. To test whether a given outcome is in the core we must test it against deviations by all possible combinations of disjoint coalitions and for all of their strategies. A deviation is successful if it represents an improvement and deviating players deviate expecting rational residual behaviour.

Two cases are considered: a non-empty residual core implies a non-transient residual outcome. Moreover for normal form games the non-emptiness of the core implies that the grand coalition can form where all claims are satisfied, that is, in contrast to characteristic function form games the core outcomes represent social optima. Should the core be empty such outcomes do not exist, and based on our theories there is no way to predict which outcomes arise. Therefore we only require that the strategy profile chosen satisfies the best-response property for the coalition structure that formed, but make no restriction on the coalition structure. This is a direct generalisation of the concept of Chander and Tulkens (1994).

Observe the different attitude in the optimistic and the pessimistic approaches. In the optimistic case a deviation occurs if there is a way to achieve improvement for all deviating coalitions, while in the pessimistic case such an improvement must be guaranteed in all cases.

4.4. **Features and properties.**

Lemma 4.4. *Let* $\Gamma = \left(N, (X_i, u_i)_{i\in N}\right)$ *be a game and let* $C(\Gamma)$ *be one of its cores (optimistic or pessimistic). Then if* $(\pi, x, \mathcal{P}) \in C(\Gamma)$ *such that* $\mathcal{P} \neq \{N\}$*, then* $(\pi, x, \{N\}) \in C(\Gamma)$.

Proof. First we show that $x \in \beta(\{N\})$. Should this not hold then the grand coalition would upset the outcome, and play something else. Since the grand coalition has a complete control over the strategy selection it deviates if there exists a strategy profile $x' \in X$ such that $u_N(x') > u_N(x)$. However, should such an x' exist the outcome $(\pi, x, \mathcal{P})$ would be dominated via the grand coalition contradicting $(\pi, x, \mathcal{P}) \in C$.

Now given that $x \in \beta(\{N\})$, the outcome $(\pi, x, \{N\}) \in C$ is only subject to coalitional deviations. Since $(\pi, x, \mathcal{P})$ has the same payoff for all players and by our assumption is immune to all deviations, we get the required result. $\square$

Now consider an outcome and a coalitional deviation. The residual core –provided non-empty– consists of outcomes with strategies that are optimal under the grand coalition. From the point of view of externalities, transfers between players do not matter, so we focus on the set of core-strategy profiles. While theoretically it is possible that they all produce the same total payoff for the grand coalition, in applications we will often find either a unique core-strategy vector or strategy vectors with symmetries across players that give the same to the deviating players. Then the expectations and hence the cores under the optimistic and the pessimistic scenarios coincide. Note, however that this property holds only for 'games with nonempty residual cores and only for "most" of these.[2]

4.5. **Comparison to existing concepts.**

4.5.1. *Classical approaches.* Since the seminal work of von Neumann and Morgenstern (1944) one typically constructs the characteristic function by taking the worst case. Cornet (1998) considers variations of this approach as well as its optimistic pair. While the simplicity of these definitions is appealing it is clear that if we insist on externalities or if externalities play a significant role such considerations are misleading.

It requires no proof to see that when we apply a smaller residual strategy set the behavioural assumptions (optimism/pessimism) play a smaller role and as a result our concept, although contains an element of optimism or pessimism, is less sensitive to them. An illustration for this can be found at (Kóczy 2002) for a game in the partition function form, a generalisation of the characteristic function form that includes externalities.

Another point we insist on is to take the two cores together, as the upper and lower ends of an interval of sets (for inclusion). This interval is always confined in the interval of classical optimistic and pessimistic cores and is typically narrower. The two cores enable us to make different types of predictions, or to avoid different types of errors. The optimistic core is smaller; if an outcome belongs to this core it is surely in the core. If an outcome is outside the –typically larger– pessimistic core then it is definitely dominated. For the rest these theories do not give a decisive conclusion.

4.5.2. *The γ-core.* In Section 3.3 we have already discussed the γ-core of Chander and Tulkens (1997) in detail. The core-pair we have introduced generalises the notion of γ-core. On the other hand there is no exclusion-inclusion relation between the γ-core and the above core-pair.

[2] Our remark follows the same philosophy as with the emptiness of the core: Although we know that the core can be empty, we keep on using it, and hope that in the given application it will not be. Here we can also hope to have coinciding optimistic and pessimistic cores giving very robust results.

4.5.3. *Equilibrium binding agreements.* [3] Ray and Vohra (1997) defined equilibrium binding agreements (EBA's) and the concept became instantaneously popular for games with externalities. The EBA's are outcomes with best-response strategy profiles that are also immune to "credible" defections by a subcoalition. While the original definition did not allow transfers here we consider a TU-version to allow a better comparison.

We need some extra notation. For a partition $\mathcal{P}$, let $\mathcal{R}(\mathcal{P})$ denote its refinements. The coalitions of a refinement $\hat{\mathcal{P}}$ of $\mathcal{P}$ are subsets of coalitions in $\mathcal{P}$. The deviating coalitions in $\hat{\mathcal{P}}$ that enforced $\hat{\mathcal{P}}$ from $\mathcal{P}$ are called *perpetrators* while the rest are the residuals. As a coalition may break into several, say k subcoalitions, $k-1$ of these have to be labelled as perpetrators. A *re-merging* is a coalition structure formed by the merger of perpetrators with their respective residuals.

Definition 4.5 (Equilibrium binding agreements for a given partition)**.** The definition is recursive. Let $\mathcal{B}(\mathcal{P})$ denote the set of EBA's for a given partition $\mathcal{P}$.

1. For the trivial partition, $\mathcal{P}_0$ of singleton coalitions as no further deviations are allowed, $\mathcal{B}(\mathcal{P}_0) = \beta(\mathcal{P}_0)$.

2. Now consider partitions $\mathcal{P}$ with $\mathcal{P}_0$ as the only possible refinement. For any $x \in \beta(\mathcal{P})$ we say that $(u(x_0), x_0, \mathcal{P}_0)$ blocks $(\pi, x, \mathcal{P})$ if $x_0 \in \mathcal{B}(\mathcal{P})$, and there exists a perpetrator S such that $u_S(x_0) > \pi_S(x)$.

3. Assume that for some $\mathcal{P}$ the set $\mathcal{B}(\mathcal{P}')$ has been defined for all $\mathcal{P}' \in \mathcal{R}(\mathcal{P})$ and that for each $x' \in \beta(\mathcal{P}')$ all outcomes $(\pi'', x'', \mathcal{P}'')$ blocking outcome $(\pi', x', \mathcal{P}')$ have been found.

4. Let $x \in \beta(\mathcal{P})$. Then $(\pi, x, \mathcal{P})$ is blocked by $(\pi', x', \mathcal{P}')$ if $\mathcal{P}' \in \mathcal{R}(\mathcal{P})$ and there exists a collection of perpetrators and residuals in the move from $\mathcal{P}$ to $\mathcal{P}'$ such that

   (1) x' is a binding agreement for $\mathcal{P}'$,
   (2) x' satisfies that $x'_S = u_S(x')$ for all S in $\mathcal{P}$,
   (3) there is a *leading perpetrator* S, which gains from the move, that is, $x'_S(x') > u_S(x)$,
   (4) any re-merging, $\hat{\mathcal{P}}$ of the *other* perpetrators is blocked by outcome $(\pi', x', \mathcal{P}')$ as well, with one of these perpetrators as a leading perpetrator. Formally, let $\mathcal{S}$ be the set of perpetrators other than S in the move from $\mathcal{P}$ to $\mathcal{P}'$. Then $\mathcal{B}(\hat{\mathcal{P}}) = \emptyset$ and there exists a strategy profile $\hat{x} \in \beta(\hat{\mathcal{P}})$, a profit vector $\hat{\pi}$ and

[3] Due to the formal similarities our core concept is often compared to the Equilibrium Binding Agreements of Ray and Vohra (1997). We have therefore felt it important to compare the two concepts in some detail, although we have some reservations about the implications of such a comparison. These concerns were articulated by Murat Sertel at the 8th Coalition Theory Workshop in Aix-en-Provance, saying that farsighted concepts deal with fundamentally different games and as such cannot be used to improve or correct myopic concepts. Thus a comparison of Equilibrium Binding Agreements and the core is also not really appropriate.

$S' \in \mathcal{S}$, such that $(\hat{\pi}, \hat{x}, \hat{\mathcal{P}})$ is blocked by $(\pi', x', \mathcal{P}')$ with S' as the leading perpetrator.

5. A strategy profile a is an *equilibrium binding agreement for* $\mathcal{P}$ if $x \in \beta(\mathcal{P})$ and there is no $(\pi', x', \mathcal{P}')$ that blocks $(\pi, x, \mathcal{P})$.

Several different partitions may admit EBA's. Since the grand coalition is taken as the starting point the focus is on the the EBA's with the grand coalition itself, or if such agreements do not exists the EBA's admitted by the coarsest partitions.

In the following we list some similarities and differences between the cores and the EBA's and produce some simple results about their relation.

One notices immediately the otherwise uncommon recursive definition. The leading perpetrators "leave the room" and the negotiations continue with the residuals only, just as in the theory we proposed. However, equilibrium binding agreements are only safe against refining deviations: Ray and Vohra further assumes that "leaving the room" breaks all contacts and further communication is impossible. We allow arbitrary deviations, as we think that the purpose of a deviation is exactly to establish a new alliance where the individuals interests are better respected. This aforementioned limitation is the reason for the arbitrary starting point: the grand coalition, while in our case the bargaining can start anywhere.

EBA's are only safe against *credible* deviations that are not subject to subsequent deviations. Clearly our concept cannot satisfy such a property. Most outcomes are subject to a deviation via the grand coalition. In particular, when we test the stability of the grand coalition we find that (almost) no deviations are credible as renegotiation is always possible. On the other hand, while we call our concept myopic the stability of the resulting outcome is safe against most possible subsequent deviations, and so stability is concerned our concept contains also some farsighted flavour. Before we discuss this further, note that we do not explicitly determine the post-deviation outcome, but consider a set of outcomes that *can* form. Recall also that the deviation occurred not to attain certain outcomes, but based on beliefs about what the resulting outcome *can* be. Therefore the credibility of a deviation does not depend on what the deviating players get in a given post-deviation outcome, but what *could* it get considering all cases. A post-deviation outcome must be stable against three types of deviations to have credibility.

(1) The deviating coalitions contain both deviating and residual players. If we measure farsightedness against EBA's where such regrouping is not permitted we can exclude this case.
(2) The deviating players are all original deviators: This subsequent deviation is also profitable from the starting outcome so this

way we replace a non-credible deviation by another profitable deviation. Moreover, if we assume that a deviating set of players always chooses the best deviating partition the new deviation must contain only a subset of players eventually leading to a deviation that is immune to subsequent ones.

(3) The deviating players are all residuals. Here we discuss two subcases.
   (a) The residual core was empty. Deviating players already expected this instability of the residual game hence the deviation results in no change.
   (b) The residual core was not empty. A subsequent deviation can only be profitable if the leading deviators change their strategies as well. But since the post-deviation outcome is now taken as the starting point there seems to be no problem with changing roles, and indeed this type would break the credibility of some deviations.

Farsightedness makes stability easier, restricting deviations to refining ones makes it harder to maintain, and so the relation between the cores and the equilibrium binding agreements is somewhat unclear.

In the following we give a simple lemma about the relation of the core and the EBA's. Since EBA's are defined in an optimistic form, we also use the optimistic core for a better comparison.

Lemma 4.6. *If all equilibrium binding agreements are inefficient the (optimistic) core of the game is empty.*

Proof. Inefficiencies arise if all outcomes $(\pi, x, \mathcal{P})$ with $\mathcal{P} = \{N\}$ are blocked, where $x \in \beta(\{N\})$. we show that blocking correspond to deviations. Our proof is by induction.

1. The result is trivial for single-player games.
2. Assume that the lemma holds for all games with at most $n-1$ players.
3. Prove the lemma for a game with n-players.

If an outcome is blocked, then there exists a perpetrator that gets better off *given* the EBA's in the residual game.

(1) If the residual core of the corresponding deviation is empty, under the optimistic beliefs the deviating coalition expects at least this much, thus it will also deviate.
(2) If the residual core is not empty, by Lemma 4.4 it contains outcomes with strategies supporting the grand coalition, then –using our inductive assumption– the grand coalition is the coarsest residual EBA, hence the perpetrators of the ongoing blocking have the same beliefs as the deviating coalition(s), and so deviation and blocking happens at the same time.

We have checked all possibilities, and found that deviations and blockings correspond to each other. Inefficiencies imply blockings for the best response strategies under the grand coalition. By our proof then there are profitable deviations from all core outcomes supporting the grand coalition. But by the same Lemma 4.4 this implies that the core is empty. This proves our lemma. □

This result, however, does not imply that the set of the coarsest EBA's would include the core. If the core is not empty, by Lemma 4.6 there exist efficient EBA's that is, there exists an EBA for the grand coalition, too. Thus if the core contains outcomes with different partitions they will not be in the solution of Ray and Vohra.

4.5.4. *The r-core.* Huang and Sjöström (2001) define the r-theory and among others the r-core. In this parallel work to ours they consider a similar game and depart from a rather similar idea, but on the end their implementation is somewhat different and these differences become magnified as we look at the final product. Their aim is also to introduce consistency in dealing with the residual players, but they keep the characteristic function as an intermediate tool to find a solution. For completeness we present their definition. In order to do this we need to introduce some extra notation.

For any set $S \subseteq N$ let $\mathcal{S}$ be a partition of S and let R be its complement.Then let $X(\mathcal{S})$ denote the set of possible strategies given $\mathcal{S}$. The worth of each subcoalition $T \subseteq R$ is denoted by $V^{\mathcal{S}}(T)$.[4]

Definition 4.7 (r-core (Huang and Sjöström 2001))**.** The definition consists of five steps.
Step 1. Trivial case
If $|R| = 1$ then for any $\mathcal{S}$ partition of $N \setminus R$,

$$X(\mathcal{S}) \equiv \beta(\mathcal{S} \cup \{R\}).$$

Step 2. Induction hypothesis
Assume that $X(\mathcal{S})$ has been defined for all $1 \leq |R| \leq k-1$ and all partitions $\mathcal{S}$ of $N \setminus R$.

Step 3. Inductive step
Consider R, such that $|R| = k$. Then let

$$V^{\mathcal{S}}(R) = \min\left\{u_R(x) \;\middle|\; x \in \beta(\mathcal{S} \cup \{R\})\right\},$$

and for any proper subset $T \subsetneq R$ define

$$V^{\mathcal{S}}(T) = \min\left\{u_T(x) \;\middle|\; x \in X(\mathcal{S} \cup \{R \setminus T\})\right\}$$

The right hand side of the expression is well-defined since $R \setminus T$ has strictly fewer players and is nonempty.

[4]Our notation simplifies that of Huang and Sjöström (2001). In their work $V^{\mathcal{S}}(T)$ is written as $V(T\,|R,\mathcal{S}) = V\left(T\,\middle|N \setminus (\bigcup_{C\in\mathcal{S}} C), \mathcal{S}\right)$, and $X(\mathcal{S})$ as $C(R\,|R,\mathcal{S}) = C\left(N \setminus (\bigcup_{C\in\mathcal{S}} C)\,\middle|N \setminus (\bigcup_{C\in\mathcal{S}} C), \mathcal{S}\right)$.

Step 4. The definition of $X(\mathcal{S})$
Let $C(R, V^{\mathcal{S}})$ denote the core of the (characteristic function form) game $(R, V^{\mathcal{S}})$. Then

$$X(\mathcal{S}) = \left\{ x \mid x \in \beta(\mathcal{R} \cup \mathcal{S}),\ \mathcal{R} \in \Pi(R),\ \exists u(x) \in C(R, V^{\mathcal{S}}) \right\}. \tag{4.5}$$

Step 5. The largest game
As the induction continues we reach larger and larger residual games. Huang and Sjöström (2001) consider the extreme case a deviation by a grand coalition at the end as if it happened in a residual game that coincides with the entire game. The above definitions can be reproduced here without a problem bearing in mind that $\mathcal{S} = \emptyset$. This concludes the definition.

While reducing the game to the characteristic function form saves us from part of the complicity we encountered in our definitions, it becomes a limiting feature in two aspects. Firstly the characteristic function cannot directly give payoffs for multi-coalition deviations; so Huang and Sjöström allows deviators to optimise their partition (Equation 4.5), but this definition seems to allow transfers *across* deviating coalitions. This feature may lead to deviations that are profitable for all deviators in the characteristic function form game but the same deviation in the normal form could even result in a loss for players in some of the coalitions involved. For this reason we think it is unappropriate to use the characteristic function as an intermediate tool to solve normal form games.

Secondly, Huang and Sjöström insist on consistency to the extent that their characteristic function is only defined if all residual cores are nonempty. As the size of games increase, the number of residual games increase tremendously and this requirement becomes very demanding. Huang and Sjöström realise this problem and suggest it as a topic for further research. The present paper offers a possible answer.

This restriction raises a more fundamental issue that we illustrate with a 4-player game having an empty residual core (Table 1).

		L			C			R		
		l	c	r	l	c	r	l	c	r
	u	5555	6005	0555	0065	0105	1845	5505	4815	0005
U	m	0605	0015	1485	1005	0005	0005	8415	0005	0005
	d	5055	4185	0005	8145	0005	0005	0005	0005	1115
	u	8880	0000	0000	0000	0000	0000	0000	0000	0000
D	m	0000	0000	0000	0000	0000	0000	0000	0000	0000
	d	0000	0000	0000	0000	0000	0000	0000	0000	0000

TABLE 1. A 4-player example with an empty residual core

Players 1, 2, 3 and 4 have strategies {l, c, r}, {u, m, d}, {L, C, R} and {U, D} respectively. Payoffs are represented as 4-digit numbers to save space, where the respective digits give the payoff of the individual players, so for instance at (c, d, L, U) we read 4185, therefore player 1 gets 4, player 2 gets 1, player 3 gets 8 and player 4 gets 5. The game is symmetric in the first three players. After an inspection of the payoffs we find that the grand coalition maximizes its payoff by playing (l, u, L, D), where it collects 24. Consider a deviation by player 4. The residual core is empty. This makes the r-characteristic function and hence the r-core undefined, even though the payoff of player 4 is independent of the residual strategies: 5. Our concepts find coincidental nonempty optimistic and pessimistic cores. An intuitive argument yields the same: All residual reactions satisfying the best-response property yield the same payoff for the deviating coalition. Moreover the payoff of a deviating coalition is the same regardless of the coalitions it coordinated its deviation. Hence in this case the characteristic function is well defined and yields the same nonempty core. Put it differently: the above game is a normal-form representation of a characteristic function game with a nonempty core. Yet, the r-core is undefined and hence it is not a generalisation of the core.

5. Cohesiveness and the grand coalition

5.1. **Criticisms of cohesiveness.** We have defined the new core concept for games in the classical and very intuitive normal or strategic form. A strategic form game is defined by a set of players, their possible strategies and a payoff function that describes the proceeds from a given strategy profile, or in other words, what a single play of the game yields. The definition is indeed very intuitive and flexible, we can easily model everyday situations by normal form games.

Problems begin when we allow coalitions to form. What is actually the role of a coalition? Players form coalitions to coordinate their strategy selection and to allow transfers, or compensations should the overall optimal strategy be suboptimal for individual members. Is it realistic to assume that a coalition's strategies are the combinations of its members' strategies? Can it follow all of these strategies, exactly these, and only these? In particular: is it realistic to assume that the grand coalition can do anything its members could, and, moreover, achieve the any payoff its members would? Cohesiveness as this feature is called seems to be rather limiting. Indeed, much of the criticism of superadditivity applies also to this weaker form.

The grand coalition is often the starting point and the purpose of the game is often to find a "fair" or "stable" distribution of the payoff of the grand coalition. For some games this idea might work well: "superadditivity is intuitively rather compelling; why should not disjoint coalitions when acting together get at least as much as they can when acting separately?" (Aumann and Drèze 1975, p233). However there might be situations when there are objective reasons that disallow the formation of the grand coalitions, such as legal objections against monopolies. Aumann and Drèze (1975), Guesnerie and Oddou (1979) and more recently Carraro and Siniscalco (1988) discuss circumstances where subadditivity occurs and/or the grand coalition does not form. The fact that here we always get the grand coalition (Lemma 4.4) suggest that there is either a problem with the solution applied or the specification of the game.

Also: Zhou (1994) lists 3 conditions that a solution should satisfy. The core already fails, because it can be empty, but this feature has long been accepted as a price for simplicity. Now here, although we started from the coalition structure core, we always produce outcomes with the grand coalition that makes the core fail for the second time. This again suggests that we should look for a change. We believe the improvement should be made at the side of the game and not of the solution.

The purpose of this section is to suggest a generalisation of the normal form that allows other coalition structures to emerge as optimal.

5.2. **Generalisation of the normal form.** Our generalisation is analogous to the step from games in the characteristic function form to games in partition function form although our motivation is different: The partition function form was defined by Thrall and Lucas (1963) to introduce externalities in coalition formation games. Externalities are already present here, we want to introduce subadditivity that is a matter of course in characteristic function games. In our definition we will essentially define a partition function that in turn defines a payoff function corresponding to a different normal form game for each partition as follows.

Definition 5.1 (Generalised normal form game)**.** Let N be the set of players, $\mathcal{P}$ a partition of the players and $S \subseteq N$ a coalition. Then we define a *strategic partition function*:

$$\begin{array}{rcccl} U : \Pi(N) & \longrightarrow & & \Phi & \\ \mathcal{P} & \longmapsto & \big(\prod_{S\in\mathcal{P}} \tilde{X}_S(\mathcal{P}) & \longmapsto & \mathbb{R}^{|\mathcal{P}|}\big) \\ & & \tilde{x}(\mathcal{P}) & \longmapsto & u_{\mathcal{P}}(\tilde{x}(\mathcal{P})), \end{array} \tag{5.1}$$

where Φ is the set of possible payoff functions.

In this game coalitions have their "own" strategies, such as $\tilde{x}_S$ and we tried to reflect it in the notation as well: here the coalitions' strategies are no more products of their members strategies (x_S), but act as players on their own ($\tilde{x}_S$). At the same time it is very easy to produce the generalised normal form of a normal form game starting from the payoff function for the all-singletons partition $\mathcal{P}_0$ and setting $\tilde{X}_S(\mathcal{P}) = X_S(\mathcal{P}_0) = \prod_{i\in S} X_i(\mathcal{P}_0)$.

Difficult as it seems the generalised normal form is not at all more complicated than the normal form while it offers the desired generality. It is also clear that the core concepts extend to this form after small changes in the notation and that once the grand coalition becomes less attractive (mathematically) more colourful cores may arise.

6. Conclusion

The model we presented here is rather general: although we have made certain assumptions to facilitate our arguments many of these can be removed without significantly affecting our results. In contrast to other concepts we have made every effort to minimise the advantage the deviating players over the residuals. We have obtained this both by permitting much more general and consequently more realistic residual reaction, and by restricting the deviating players' possibilities to impose an arbitrary situation on the residuals. While this formulation is more appropriate in many, symmetric situations, in others, such as a Stackelberg leader-follower model, or in a Bertrand competition we may actually want to give certain advantage to the deviating players.

In such a case the residual game is a function of the deviators' strategies, while it does not depend on their partition, which only influences the transfers they can make among themselves.

Definition 6.1 (Residual game)**.**

$$\Gamma^s = \left(R, (X_i, u_i^s)_{i \in R}\right), \tag{6.1}$$

where R is the residual set, $s \in X_{N \setminus R}$ is the strategy of the deviating coalitions, and u_i^s is the *residual payoff function* that we define as follows for all $x \in X_R$.

$$u_i^s(x) = u_i(x, s) \tag{6.2}$$

When we use this definition the two stages cannot be separated so clearly, as a deviation in the cooperative stage already announces a strategy for the non-cooperative stage. In addition to that the residual players will play knowing this strategy and hence the noncooperative play can be sequential, not necessarily simultaneous. Of course all this refers to deviations, while for core outcomes deviations do not arise and all these arguments are purely theoretical.

References

Aumann RJ, Drèze JH (1975) Cooperative games with coalition structures. International Journal of Game Theory 3:217–237. Also, as UCL-CORE Reprint 217.

Carraro C, Siniscalco D (1988) International environmental agreements: Incentives and political economy. European Economic Review 42:561–572

Chander P, Tulkens H (1994) The core of and economy with multilateral environmental externalities. Nota di Lavoro 69.94, Fondazione Eni Enrico Mattei, Milano

Chander P, Tulkens H (1997) The core of and economy with multilateral environmental externalities. International Journal of Game Theory 26(3):379–401

Cornet MF (1998) Game Theoretic Models of Bargaining and Externalities. Tinbergen Institute Research Series 176. Thesis Publishers, Amsterdam

Greenberg J (1994) Coalition structures. In: Aumann RJ, Hart S (eds.) Handbook of game theory with economic applications, Elsevier, Amsterdam chapter 37, pp. 1305–1337

Guesnerie R, Oddou C (1979) On economic games which are not necessarily superadditive: Solutions, concepts and application to a local public good game with 'few' agents. Economic Letters 3(4):301–306

Harsányi JC (1963) A simplified bargaining model of the n-person cooperative game. International Economic Review 4(2):194–220

Huang CY, Sjöström JT (2001) Consistent solutions for cooperative games with externalities. Based on a Penn State University working paper with title "The p-core". Available from the authors.

Ichiishi T (1981) A social coalitional equilibrium existence lemma. Econometrica 49(2):369–377

Kóczy LÁ (2002) A note on Funaki and Yamato's tragedy of the commons. Discussion Paper DPS-02.17, Centre for Economic Studies, Leuven

Kuhn HW (1953) Extensive games and the problem if information. In: Kuhn HW, Tucker AW (eds.) Contributions to the theory of games II, Annals of Mathematics Studies 28. Princeton University Press, Princeton pp. 193–216. Reprinted in Kuhn, ed (1997, pp. 46-68.)

Kuhn HW (ed.) (1997) Classics in Game Theory. Frontiers of Economic Research Princeton University Press, Princeton, New Jersey

Peleg B (1992) Axiomatizations of the core. In: Aumann RJ, Hart S (eds.) Handbook of game theory with economic applications, Elsevier, Amsterdam chapter 13, pp. 397–412

Ray D, Vohra R (1997) Equilibrium binding agreements. Journal of Economic Theory 73(1):30–78

Thrall RM, Lucas WF (1963) n-person games in partition function form. Naval Research Logistics Quarterly 10(4):281–298

von Neumann J, Morgenstern O (1944) Theory of Games and Economic Behavior Princeton University Press, Princeton

Zhao J (1992) The hybrid solutions of an n-person game. Games and Economic Behavior 4(1):145–160

Zhou L (1994) A new bargaining set of an n-person game and endogenous coalition formation. Games and Economic Behavior 6(3):512–526

Appendix A. Pareto efficiency requires multi-coalition deviations

Table 2 defines a 4-player game with the property that the partition $\mathcal{P} = \{\{12\},\{34\}\}$ Pareto-dominates $\mathcal{P}_0\{\{1\},\{2\},\{3\},\{4\}\}$, while only dominates it via the entire partition $\mathcal{P}$ and not via either of the coalitions. This observation supports the use of multi-coalition deviations. In this game player 1 has strategy set {l, c, r}, player 2 has {u, m, d}, player 3 has strategies {L, C, R}, player 4 {U, M, D}. In the tables payoff vectors are written as multi-digit numbers to save space, but this should lead to no confusion. The (up to symmetries) unique Nash-equilibria is set in boldface.

		L			C			R		
		l	c	r	l	c	r	l	c	r
	u	**2222**	0522	0522	1105	0050	0050	1105	0050	0050
U	m	5022	0000	0000	0050	0000	0000	0050	0000	0000
	d	5022	0000	0000	0050	0000	0000	0050	0000	0000
	u	1150	0005	0005	1111	0211	0011	0011	0011	0011
M	m	0005	0000	0000	2011	1111	0333	0011	3303	2424
	d	0005	0000	0000	0011	3033	2244	0011	4224	2209
	u	1150	0005	0005	0011	0011	0011	0033	1011	1011
D	m	0005	0000	0000	0011	3330	2442	0111	4422	0922
	d	0005	0000	0000	0011	4242	2290	0111	9022	1111

TABLE 2. The payoff function in case of pure noncooperative behaviour.

	ul			uc			ur		
	L	C	R	L	C	R	L	C	R
U	422	205	205	522	050	050	522	050	050
M	250	211	011	005	211	011	005	211	011
D	250	011	033	005	011	111	005	011	111

	mc			mr			dr		
	L	C	R	L	C	R	L	C	R
U	000	000	000	000	000	000	000	000	000
M	000	211	603	000	333	624	000	**444**	409
D	000	630	822	000	642	922	000	490	211

TABLE 3. Payoffs in the noncooperative stage with $\mathcal{P}_1 = \{\{1,2\},\{3\},\{4\}\}$.

Now consider the game where players 1 and 2 formed a coalition and so instead of maximising their individual payoff they maximise their joint payoff. The game is symmetric between 1 and 2 and between 3 and 4. Payoff-identical cases are only listed once. Table 3 gives the coalitional payoffs when partition $\mathcal{P}_1 = \{\{1,2\},\{3\},\{4\}\}$ is formed. In this representation the first coalition chooses matrices, the second chooses columns and the third chooses rows. The unique strategy profile that satisfies the best-response property is marked in boldface. In a similar fashion we can create a table of coalitional payoffs when $\mathcal{P}_2 = \{\{1\},\{2\},\{3,4\}\}$ is formed. The payoffs are presented in Table 4.

Finally we present the coalitional payoffs for the partition when both coalitions have formed, that is, when we have partition $\mathcal{P}_3 = \{\{1,2\},\{3,4\}\}$. In this case the strategies of coalition $\{1,2\}$ are {ul,

	UL			UC			UR		
	l	c	r	l	c	r	l	c	r
u	224	054	054	115	005	005	115	005	005
m	504	000	000	005	000	000	005	000	000
d	504	000	000	005	000	000	005	000	000

	MC			MR			DR		
	l	c	r	l	c	r	l	c	r
u	112	022	002	002	002	002	006	102	102
m	202	112	036	002	333	246	012	**444**	094
d	002	306	228	002	426	229	012	904	112

TABLE 4. Payoffs in the noncooperative stage with $\mathcal{P}_2 = \{\{1\},\{2\},\{3,4\}\}$.

	ul	uc	ur	mc	mr	dr
UL	44	54	54	00	00	00
UC	25	05	05	00	00	00
UR	25	05	05	00	00	00
MC	22	22	02	22	36	48
MR	02	02	02	63	**66**	49
DR	06	12	12	84	94	22

TABLE 5. Payoffs in the noncooperative stage with $\mathcal{P}_3 = \{\{1,2\},\{3,4\}\}$.

uc=ml, ur=dl, mc, mr=dc, dr}, the strategies for $\{3,4\}$ are {UL, UC=ML, UR=DL, MC, MR=DC, DR}. Table 5 contains the coalitional payoffs. This game has again Nash equilibria that are unique up to the coalitional payoffs.

The equilibria and the resulting payoffs are therefore exactly as described in Section 4.1.1.