

University of Warwick institutional repository: <http://go.warwick.ac.uk/wrap>

This paper is made available online in accordance with publisher policies. Please scroll down to view the document itself. Please refer to the repository record for this item and our policy information available from the repository home page for further information.

To see the final version of this paper please visit the publisher's website. Access to the published version may require a subscription.

Author(s): Bhaskar Dutta with Francis Bloch

Article Title: Correlated Equilibria, Incomplete Information and Coalitional Deviations

Year of publication: 2006

Link to published version:

[http://www2.warwick.ac.uk/fac/soc/economics/research/papers/2006\\_publications/twerp\\_763.pdf](http://www2.warwick.ac.uk/fac/soc/economics/research/papers/2006_publications/twerp_763.pdf)

Correlated Equilibria, Incomplete Information and Coalitional Deviations

No 763

**WARWICK ECONOMIC RESEARCH PAPERS**

**DEPARTMENT OF ECONOMICS**

THE UNIVERSITY OF  
**WARWICK**

# Correlated Equilibria, Incomplete Information and Coalitional Deviations\*

Francis Bloch, Bhaskar Dutta †

June 21, 2006

## Abstract

This paper proposes new concepts of strong and coalition-proof correlated equilibria where agents form coalitions at the interim stage and share information about their recommendations in a credible way. When players deviate at the interim stage, coalition-proof correlated equilibria may fail to exist for two-player games. However, coalition-proof correlated equilibria always exist in dominance-solvable games and in games with positive externalities and binary actions. *JEL Classification Numbers: C72*

*Keywords: correlated equilibrium, coalitions, information sharing, games with positive externalities*

---

\*Dutta gratefully acknowledges support from ESRC Grant RES-000-22-0341.

†Bloch is at GREQAM, Universite d'Aix-Marseille, 2 rue de Charite, 13002 Marseille, France. He is also affiliated with the University of Warwick. Dutta is in the Department of Economics, University of Warwick, Coventry CV4 7AL, England.

# 1 Introduction

A game with *communication* arises when players have the opportunity to communicate with one other prior to the choice of actions in the actual game. The presence of a *mediator* is a particularly powerful device in such games because it allows players to use *correlated* strategies- the mediator (privately) recommends actions to each player according to the realization of an agreed upon correlation device over the set of strategy  $n$ -tuples. Of course, the use of correlated strategies can in principle allow players to achieve higher expected payoffs than those possible through independent randomisations of strategies.

A *correlated equilibrium* is a *self-enforcing* correlated strategy  $n$ -tuple because no individual has an incentive to deviate from the recommendation received by her, given the information at her disposal. This information is simply the recommendation received by her and the (prior) probability with which each strategy  $n$ -tuple was to be chosen. In particular, she has no information about the recommendations received by others- unless she can infer these from her own recommendation and the prior probability distribution over the set of strategy  $n$ -tuples. However, if players can communicate with each other, then it is natural to ask whether coalitions of players cannot exchange information about the recommendations received by them and plan mutually beneficial *joint* deviations. Although pre-play communication opens up the possibility of sharing information about the recommendations received by different players, there may still be constraints on the extent of information which can be shared by different players. The possibility of constraints on information sharing has been recognised in the literature on cooperative game theory with incomplete information ever since the classic paper of Wilson (1978). These constraints are clearly present in the present context since the recommendation received by each player  $i$  can be viewed as her "type".

Wilson (1978) defined two polar solution concepts - the *coarse* core and the *fine* core of an exchange economy with private information. The coarse core corresponds to the case where deviating coalitions cannot share any information about their types, so that blocking plans have to be drawn up on common knowledge events. The fine core corresponds to the case where there is no constraint on information sharing within a coalition. Dutta and Vohra(2005) argue that both notions are extreme in the sense that the typical situation is one where *some* but not *all* information can be shared. They

propose a notion of the *credible core*, which allows for information-sharing which is credible.<sup>1</sup>

In this paper too, we impose the requirement that members of a deviating coalition can only share "credible" information about the recommendations received by them. Given this basic premise, we define two refinements of correlated equilibria. The first concept is analogous to that of strong Nash equilibrium. A correlated strategy  $n$ -tuple is a *strong correlated equilibrium* if it is immune to deviations by coalitions of essentially myopic players who do not anticipate any further deviations after the coalition has implemented its blocking plan. The second concept is that of *coalition-proof correlated equilibrium*. According to this concept, coalitions take into account the possibility that sub-coalitions may enforce further deviations.

Notice that in our framework, coalitions plan deviations at the *interim* stage - that is, after the mediator has communicated his recommendation to each player. Of course, coalitions could also form at the *ex ante* stage, that is before the mediator has communicated his recommendations to the players. Moreno and Wooders (1996), Milgrom and Roberts (1996) focus on these *ex ante* concepts. We comment on the relationship between their solution concepts and ours in section 3. In particular, we construct a 2-person game which has no interim coalition-proof equilibrium, although coalition proof equilibria in the *ex ante* sense always exist in 2-person games. In section 4, we focus on strong correlated equilibrium. We construct examples to show that there is no connection between the existence of strong Nash equilibria and strong correlated equilibrium. We also show that a class of positive externality games studied by Konishi et al (1997a) ensure the existence of strong correlated equilibrium. Finally, in section 5, we switch attention to (interim) coalition proof correlated equilibrium. We show that if a pure strategy action profile Pareto dominates all other pure strategy profiles which survive the iterated elimination of strictly dominated strategies, then it must be an interim coalition-proof correlated equilibrium. We also show that if the action sets of all individuals is restricted to *two* identical actions, then all positive externality games have interim correlated coalition proof equilibria.

---

<sup>1</sup>For instance, in the lemons problem, the seller cannot credibly claim to be the high-quality seller.

## 2 Notions of Interim Coalitional Equilibrium

Let  $N$  denote the set of players, indexed by  $i = 1, 2, \dots, n$ . Each player has a finite set of pure strategies,  $A_i$  with generic element  $a_i$ .  $A$  denotes the Cartesian product,  $A = \prod_{i \in N} A_i$ . The utility of player  $i$  is given by  $u_i : A \rightarrow \mathfrak{R}$

A *correlated strategy*  $\mu$  is a probability distribution over  $A$ .

We consider an extended game with a mediator. Before the game is played, a mediator privately sends recommendations to the players,  $a$ , chosen according to the probability distribution  $\mu$ . Each player observes his recommendation and then proceeds to playing the game. As opposed to Moreno and Wooders (1996), Ray (1996) and Milgrom and Roberts (1996), we analyze deviations at the *interim* stage, when every player has received his recommendation.

Consider a correlated strategy  $\mu$  and coalition  $S$ . Suppose members of  $S$  have privately received the recommendations  $a_{-S}$ . How can they plan mutually beneficial deviations from  $\mu$ ? Any plans to “block”  $\mu$  must depend crucially on their beliefs about the realization of  $\mu$ . Moreover, each individual  $i$ 's belief about the realization of  $\mu$  depends upon the recommendation received by  $i$  himself as well as the information about  $a_{-S}$  which can be *credibly* transmitted by members of  $S$  to each other. In what follows, we adapt the notion of the *credible core* of Dutta and Vohra (2005) to this setting.

Suppose all members of  $S$  *believe* that the recommendations received by the players lie in some subset  $E$  of  $A$ . We will call such a set  $E$  an *admissible event*, and describe some restrictions which must be satisfied by such an event. First, an element  $a'_{-S}$  can be ruled out only by using the private information of members of  $S$ . Since we will use conditional expected utilities to evaluate action profiles, we can without loss of generality express this requirement as  $E = E_S \times A_{-S}$ . Second, if  $i \in S$ , then her claim that she has not received recommendation  $a'_i$  cannot depend on the claims made by other members of  $S$ . Hence,  $E_S$  must be the cartesian product of some set  $\{E_i\}_{i \in S}$ . Third, no agent can, after receiving her own recommendation, rule out the possibility that the “true” profile of recommendations lies in the set  $E$ . Hence, an admissible set for the coalition  $S$ , must satisfy the following.

**Definition 1** *Given  $\mu$  and  $a_S$ , an event  $E$  is admissible for  $S$  if and only if*

$$E = \prod_{i \in S} E_i \prod A_{-S}, \text{ and } \sum_{\widetilde{a}_{-i} \in E_{-i}} \mu(a_i, \widetilde{a}_{-i}) > 0 \text{ for all } i \in S.$$

where

$$E_{-i} = \prod_{j \in S \setminus i} E_j \prod A_{-S}$$

Given an admissible event  $E$ , we define player  $i$ 's conditional probability of  $a_{-i}$  given  $a_i$  and  $E$  as

$$\tilde{\mu}(a_{-i}|a_i, E_{-i}) = \frac{\mu(a_{-i}, a_i)}{\sum_{\tilde{a}_{-i} \in E_{-i}} \mu(a_i, \tilde{a}_{-i})}.$$

We also define the marginal probability over  $a_{-S}$  given  $a_i$  and  $E$  as:

$$\bar{\mu}(a_{-S}|a_i, E_{-i}) = \sum_{\alpha_{S \setminus i} \in \prod_{j \in S \setminus i} E_j} \tilde{\mu}(a_S, \alpha_{S \setminus i}|a_i, E_{-i}).$$

A *blocking plan* for coalition  $S$ ,  $\eta_S$ , is a correlated strategy over  $A_S$ .

Once the blocking plan is implemented, a player  $i$  in  $S$  has the following posterior belief over the actions in the game:

$$\gamma_i(a) = \bar{\mu}(a_{-S}|a_i, E_{-i})\eta_S(a_S).$$

Given  $a_i$  and  $E$ , player  $i$  evaluates the correlated strategy  $\mu$  according to:

$$U_i(\mu|a_i, E_{-i}) = \sum_{a_{-i} \in E_{-i}} \tilde{\mu}(a_{-i}|a_i, E_{-i})u_i(a_i, a_{-i}).$$

Player  $i$  evaluates the blocking plan according to:

$$U_i(\eta_S|a_i, E_{-i}) = \sum_{\tilde{a} \in A} \gamma_i(\tilde{a})u_i(\tilde{a})$$

Definition 1 ensures that if members of  $S$  each claim to have received recommendations in the set  $E_i$ , then no individual in  $S$  can conclude that some individual has lied given knowledge of his own recommendation. Notice that this condition places some restriction on how *large* an admissible event can be. However, this condition by itself does not guarantee that each individual in  $S$  will believe the claims of other members of  $S$ . We explain below why there should be some restriction on how *small* an admissible event must be before individuals can agree on a plan to block a correlated strategy  $\mu$ .

Suppose  $E$  is an admissible event for coalition  $S$ , and  $i \in S$ . We want to rule out the possibility that  $i$  after receiving a recommendation  $a'_i \notin E_i$  actually claims to have received a recommendation in  $E_i$ .

Let  $V_i(E) = \{a'_i \in A_i \setminus E_i \mid \text{there is } a_{-i} \in E_{-i} \text{ such that } \mu(a_{-i}, a'_i) > 0\}$ . So, if the other individuals in  $S$  independently claimed to have received recommendations in  $E_{-i}$ , then  $i$  believes that she can declare to have received a recommendation in  $V_i(E)$ .

For any coalition  $S$ , a blocking plan  $\eta_S$  on an admissible event  $E$  satisfies *self selection* if for all  $i \in S$  and all  $a'_i \in V_i(E)$ ,

$$U_i(\mu|a'_i, E_{-i}) > U_i(\eta_S|a'_i, E_{-i}) \quad (1)$$

Self-selection guarantees that if  $i$  has agreed to the blocking plan  $\eta_S$ , then  $i$  will not falsely claim to have received a recommendation in  $E_i$ .

**Definition 2** *A coalition  $S$  blocks the correlated strategy  $\mu$  if there exists a blocking plan  $\eta_S$  and admissible event  $E$  such that*

(i)  $\eta_S$  satisfies self-selection on  $E$ .

(ii) For all  $i \in S$ ,  $U_i(\mu|a_i, E_{-i}) < U_i(\eta_S|a_i, E_{-i})$  at some  $a_S \in E_S$ .

The definition of self-selection is similar to that used by Dutta and Vohra (2005) in the definition of the credible core of an exchange economy with incomplete information. The underlying idea is the same: if members of a coalition agree to a blocking plan, this information should be used to update players' information over the recommendations received by other players in the coalition. In other words,  $E$  defines the event for which all players in  $S$  have an incentive to accept the blocking plan  $\eta_S$ . Every player in  $S$  thus updates his beliefs by assuming that players in  $S \setminus i$  have received recommendations in  $\prod_{j \in S \setminus i} E_j$ . If given these updated beliefs, all players in  $S$  have an incentive to accept the blocking plan  $\eta_S$ , then the coalition  $S$  blocks the correlated strategy at  $a_S$ .

**Definition 3** *A correlated strategy  $\mu$  is an interim strong correlated equilibrium (ISCE) if there exists no coalition  $S$  that blocks  $\mu$ .*



If the coalition  $S$  is a singleton,  $E_{-i} = A_{-i}$  and the self-selection constraint is vacuous. A singleton coalition  $\{i\}$  thus blocks the correlated strategy  $\mu$  if there exists a mixed strategy  $\sigma_i$  such that  $U_i(\mu|a_i, A_{-i}) < U_i(\sigma_i|a_i, A_{-i})$ . Hence, for singleton coalitions, our definition corresponds to the usual definition of correlated equilibrium.

As with the concept of strong Nash equilibrium, the concept of interim strong correlation equilibrium implicitly assumes that players are *myopic* when they plan deviations. In particular, the deviating coalition does not take into account the possibility that there may be further deviations. Several papers define different notions of coalitional stability when players are farsighted in the context of games with complete information.<sup>2</sup> Moreno and Wooders (1996) define a notion of (ex ante) *coalition proof correlated equilibrium* when coalitions form *before* players receive recommendations from the mediator. As in the original definition of coalition proof equilibrium for complete information games, their definition explicitly takes into account the possibility that subcoalitions may carry out further deviations.

We now define a notion of coalition proof equilibrium when coalitions form *after* players have received recommendations from the mediator. Notice that if a nested sequence of coalitions each form blocking plans, then the posterior beliefs of players “later on” in the sequence keep changing. For suppose the original correlated strategy is  $\mu$ , and coalition  $S^1$  considers a blocking plan  $\eta_{S^1}$  on the admissible event  $E^1$ . Then, players in  $S^1$  believe that the recommendations sent by the mediator lie in the set  $E^1$ . Moreover, the posterior beliefs of players in  $S^1$  are different from their prior beliefs. Now, consider “stage 2” when the coalition  $S^2 \subset S^1$  contemplates a blocking plan  $\eta_{S^2}$  on the admissible event  $E^2$ . First, their prior beliefs coincide with the posterior beliefs formed at the end of stage 1. Second, players in  $S^2$  now believe that the mediator has recommended an action profile in  $E^2$ . Implementation of the blocking plan  $\eta_{S^2}$  will result in a new set of posterior beliefs for players in  $S^2$ , and this change in posterior beliefs will also change the way in which players evaluate blocking plans. This needs to be kept in mind when defining an interim notion of coalition proofness, and also provides the motivation for the following definitions.

Consider a coalition  $S \subseteq N$ , and a *blocking sequence*  $\mathbf{B} = \{(S^k, \eta_{S^k}, E^k)\}_{k=1}^K$  to the correlated strategy  $\mu$ , where

---

<sup>2</sup>See, amongst others, Chwe (1994), Bernheim, Whinston and Peleg (1987), Ray and Vohra (1997).

- (i)  $S^1 \equiv S$ , and for each  $k = 2, \dots, K$ ,  $S^k \subset S^{k-1}$ .
- (ii)  $E^1$  is an admissible event for  $S^1$ , and for each  $k > 1$   $E_i^k \subset E_i^{k-1}$  for  $i \in S^k$ , and  $E_i^k = E_i^{k-1}$  for  $i \notin S^k$ .
- (iii) Each  $\eta_{S^k}$  satisfies self-selection on  $E^k$ .
- (iv) Each  $\eta_{S^k}$  is a correlated strategy over  $A_{S^k}$ .

We can now define the posterior beliefs of each coalition figuring in the blocking sequence as well as how members of these coalitions evaluate the blocking plans.

Let  $\gamma^0(a) \equiv \mu(a)$  for all  $a \in A$ . Choose any  $k \in \{1, \dots, K\}$ , and  $i \in S^k$ . Then,

$$\tilde{\mu}^k(a_{-i}|a_i, E_{-i}^k) = \frac{\gamma^{k-1}(a_{-i}, a_i)}{\sum_{\tilde{a}_{-i} \in E_{-i}^k} \gamma^{k-1}(a_i, \tilde{a}_{-i})}.$$

Similarly, the marginal probability over  $a_{-S^k}$  given  $a_i$  and  $E$  is:

$$\bar{\mu}^k(a_{-S^k}|a_i, E_{-i}^k) = \sum_{\alpha_{S^k \setminus i} \in \Pi_{j \in S^k \setminus i} E_j^k} \tilde{\mu}^k(a_{S^k}, \alpha_{S^k \setminus i}|a_i, E_{-i}^k).$$

Once the blocking plan  $\eta^k$  is implemented, a player  $i$  in  $S^k$  has the following posterior belief over the actions in the game:

$$\gamma_i^k(a) = \bar{\mu}^k(a_{-S^k}|a_i, E_{-i}^k) \eta_{S^k}(a_{S^k}).$$

In order to define the concept of interim coalition proof correlated equilibrium (ICPCE), we first define the notion of self-enforcing blocking plans.

**Definition 4** *Let  $T$  be any coalition.*

(i) *If  $|T^1| = 1$ , say  $i \in T^1$ , then any mixed strategy  $\sigma_i$  is a self-enforcing blocking plan against any correlated strategy  $\mu$ .*

(ii) *Recursively, suppose self-enforcing blocking plans have been defined for all coalitions of size  $(|T^1| - 1)$  or smaller against any correlated strategy. Then,  $T^1$  has a self-enforcing blocking plan  $\eta_{T^1}$  against the correlated strategy  $\mu$  if*

(a) *There is an admissible event  $E^1$  such that  $\eta_{T^1}$  satisfies self-selection on  $E^1$ , and*

(b) There is no coalition  $T^2 \subset T^1$  with a self-enforcing blocking plan  $\eta_{T^2}$  and admissible event  $E^2$  such that for the blocking sequence  $\{(T^1, \eta_{T^1}, E^1), (T^2, \eta_{T^2}, E^2)\}$  and for some  $a_{T^2} \in E_{T^2}^2$

$$\sum_{a_{-i} \in E^2} \tilde{\mu}^2(a_{-i}|a_i, E_{-i}^2) u_i(a_i, a_{-i}) < \sum_{a \in A} \gamma_i^2(a) u_i(a) \text{ for all } i \in T^2$$

An interim coalition proof correlated equilibrium (ICPCE) is a correlated strategy against which no coalition has a self-enforcing blocking plan which makes everyone in the deviating coalition strictly better off.

**Definition 5** A correlated strategy  $\mu$  is ICPCE if there is no coalition  $S$  with a self-enforcing blocking plan  $\eta_S$  against  $\mu$  such that

(i)  $\eta_S$  satisfies self-selection on some admissible event  $E$ .

(ii) For all  $i \in S$ ,  $U_i(\mu|a_i, E_{-i}) < U_i(\eta_S|a_i, E_{-i})$  at some  $a_S \in E_S$ .

Some remarks are in order. First, any strong correlated equilibrium is a coalition proof correlated equilibrium, as any self-enforcing blocking plan is a blocking plan. Second, because any blocking plan by a single player coalition is self-enforcing, any coalition proof correlated equilibrium is a correlated equilibrium. Finally, as opposed to the classical notion of coalition proof equilibria, the set of coalition proof correlated equilibria in two-player games is *not equal* to the set of undominated correlated equilibria. In our model, deviations occur at the interim stage, and players may have a joint incentive to deviate after a realization in the support of an undominated correlated equilibrium.

### 3 Related definitions of strong and coalition proof correlated equilibrium

Different definitions of strong and coalition proof correlated equilibria have already been proposed in the literature. Moreno and Wooders (1996) and Milgrom and Roberts (1996) consider coalitional deviations at the *ex ante*

stage, before agents have received their recommendations.<sup>3</sup> Formally, in their setting, a blocking plan is a mapping  $\eta_S$  from  $A_S$  to  $\Delta A_S$ , assigning a correlated strategy over  $A_S$  to any possible recommendation  $a_S$ . All computations are made ex ante. Players evaluate the correlated strategy  $\mu$  according to the expected utility

$$U_i(\mu) = \sum_{a \in A} \mu(a) u_i(a).$$

Given a blocking plan  $\eta_S$  against the correlated strategy  $\mu$ , the induced distribution over actions is given by

$$\hat{\mu}(a) = \sum_{\alpha_S \in A_S} \mu(\alpha_S, a_{-S}) \eta_S(a_S | \alpha_S)$$

and players evaluate the blocking plan according to

$$U_i(\eta_S) = \sum_{a \in A} \hat{\mu}(a) u_i(a)$$

If coalitions form at the ex ante stage, players must decide on their blocking plans in a state of symmetric information. Hence, one need not worry about the sharing of information inside a coalition. This implies that the ex ante definitions of strong correlated equilibrium and coalition proof correlated equilibrium are considerably simpler than in our model.

**Definition 6** *A correlated strategy  $\mu$  is an ex ante strong correlated equilibrium (ESCE) if there exists no coalition  $S$  and blocking plan  $\eta_S$  such that  $U_i(\eta_S) > U_i(\mu)$  for all  $i \in S$ .*

In order to economise on notation, we define self-enforcing ex ante blocking plans informally. As before, they are defined recursively. Any blocking plan by a one-player coalition is self-enforcing. Given that self-enforcing blocking plans have been defined for all coalitions  $T$  with  $|T| < |S|$ , a blocking plan  $\eta_S$  generating a distribution  $\hat{\mu}$  is self-enforcing, if there exists no coalition  $T \subset S$ , and self-enforcing blocking plan  $\eta_T$  for  $T$  generating a distribution  $\hat{\mu}_T$  such that  $U_i(\hat{\mu}_T) > U_i(\hat{\mu})$  for all  $i$  in  $T$ .

---

<sup>3</sup>Ray (1996) also proposes a notion of coalition proof correlated equilibrium at the ex ante stage. Intuitively, his concept differs from Moreno and Wooders (1996)'s, Milgrom and Roberts (1996)'s and ours in that deviating coalitions cannot choose a new correlation device, but must abide by the fixed correlation device of the extended game.

**Definition 7** A correlated strategy  $\mu$  is a ex ante coalition proof correlated equilibrium (ECPCE) if there is no coalition  $S$  and self-enforcing blocking plan  $\eta_S$  such that  $U_i(\eta_S) > U_i(\mu)$  for all  $i \in S$ .

Coalitional incentives to block at the ex ante and interim stage cannot be compared. On the one hand, it may be easier for coalitions to block at the ex ante stage. Consider for example a correlated strategy in a two-player game putting equal weight on two outcomes with payoffs  $(0, 3)$  and  $(3, 0)$ . At the ex ante stage, this correlated strategy has expected value 1.5 for every player, and would be blocked by another outcome with payoffs  $(2, 2)$ . However, at the interim stage, neither of the two realizations can be blocked by both players. On the other hand, coalitions may find it easier to block at the interim stage, when a correlated strategy puts weight on an outcome with very low payoffs for the players. The following example illustrates this point in a two-player game where a correlated strategy puts positive weight on an outcome which is Pareto-dominated by another outcome. At the ex-ante stage, the correlated strategy is not dominated, but at the interim stage, for some realization, both players have an incentive to block. This example also highlights another difference between ICPCE and ECPCE - the former may fail to exist even in two-person games, whereas ECPCE always exist in such games.

**Example 1** Consider a two-player game where player 1 chooses the row and player 2 the column.

	$b_1$	$b_2$	$b_3$
$a_1$	4, 4	0, 0	0, 4.1
$a_2$	1, 1	1, 1	-1, 0
$a_3$	0, 0	0, -1	2, 2

This game possesses two pure strategy Nash equilibria  $(a_2, b_2)$  and  $(a_3, b_3)$ . Consider the correlated strategy  $\mu$  placing probability  $1/2$  on  $(a_1, b_1)$ , and  $1/4$  on  $(a_2, b_1)$  and  $(a_2, b_2)$ . It is clear that player 1 has no incentive to deviate from the recommendations of the correlated strategy. When the column player receives recommendation  $b_2$ , she has no incentive to deviate either. When she receives  $b_1$ , her expected payoff is  $4 * 2/3 + 1/3 = 3$ . By deviating to  $b_3$ , she would receive an expected payoff of  $4.1 * 2/3 = 8.2/3 < 3$ . Hence the correlated strategy  $\mu$  is a correlated equilibrium, which gives every player

an expected payoff of 3 and hence Pareto-dominates the two pure strategy Nash equilibria.

However the correlated equilibrium  $\mu$  can be blocked by both players at the realization  $(a_2, b_2)$  where they both receive an expected payoff of 1, which is Pareto-dominated by the pure strategy Nash equilibria  $(a_3, b_3)$ . Hence, neither the two pure strategy Nash equilibria nor the correlated equilibrium  $\mu$  are CPCE of the game.

We still need to check that there is no other correlated equilibrium which would be immune to coalitional deviations. By the argument above, any correlated equilibrium putting weight on the cell  $(a_2, b_2)$  is dominated. Furthermore, if  $b_2$  is not played,  $a_2$  is strictly dominated by  $a_1$  and if  $a_2$  is not played,  $b_1$  and  $b_2$  are strictly dominated by  $b_3$ . Hence, if a correlated equilibrium does not put weight on the outcome  $(a_2, b_2)$ , it cannot put weight on the strategies  $a_2$ ,  $b_1$  and  $b_2$ . But this implies that the only other candidates for correlated equilibrium must put all the weight on the column player choosing  $b_3$  and the only possible outcome is  $(a_3, b_3)$ . Hence, any correlated equilibrium in the game must either put weight on  $(a_2, b_2)$  or concentrate all the weight on  $(a_3, b_3)$ , so that the game has no ICPCE.

Einy and Peleg (1995) define an interim notion of strong and coalition proof correlated equilibrium. Their concept differs from ours in two important respects. First, they assume that members of a blocking coalition freely share information about their recommendations.<sup>4</sup> Second, they assume that a coalition blocks if all its members are made better off *for any realization* of the initial correlated strategy. Formally, they define a blocking plan as a mapping from  $A_S$  (the set of recommended strategies in  $\mu$ ) to  $\Delta A_S$ . In their equilibrium concept, a coalition  $S$  blocks, if for all possible realizations  $a_S$ , the blocking plan is a strict improvement for all players in  $S$ .

There is no inclusion relation between the set of strong (and coalition proof) correlated equilibria defined by Einy and Peleg (1995) and the set of strong (and coalition-proof) correlated equilibria defined in this paper. On the one hand, the fact that members can freely share information about their recommendations makes deviation easier in Einy and Peleg (1995)'s sense. On the other hand, their – very strong – requirement that coalitional members

---

<sup>4</sup>In the context of exchange economies with private information, this is equivalent to the notion of "fine" core proposed by Wilson (1978). The problem of course is that players' announcements about the recommendation they received is not verifiable, and blocking plans may not be credible in our sense.

are better off for any realization of the correlated strategy makes deviations harder. Consider for instance the following example of a three-player game due to Einy and Peleg (1995).

**Example 2** (*Einy and Peleg (1995)*) Consider the following three-player game, where player I chooses rows  $(a_1, a_2)$ , player II chooses columns  $(b_1, b_2)$  and player III choose matrices  $(c_1, c_2)$ .

3,2,0	0,0,0
2,0,3	2,0,3

3,2,0	0,3,2
0,0,0	0,3,2

Einy and Peleg (1995) argue that the following is a strong correlated equilibrium.

1/3	0	0	1/3
0	1/3	0	0

To prove their claim, they note that for any two-player coalition, there exists one realization of the correlated strategies for which no strict improvement is possible. (For  $S = \{1, 2\}$ , the realization  $(a_1, b_1)$ , for  $S = \{2, 3\}$ , the realization  $(b_2, c_2)$  and for  $S = \{1, 3\}$ , the realization  $(a_2, c_1)$ .)

With our definition, we claim that this correlated strategy is *not* a ISCE. Consider the coalition  $S = \{1, 2\}$  and the realization  $(a_2, b_2)$ . Player 1 then knows that 3 has received the recommendation  $c_1$  and that 2 has received the recommendation  $b_2$ . Player 2 puts equal probability to  $(a_2, c_1)$  and  $(a_1, c_2)$ . Consider the admissible event  $E_1 = \{a_1, a_2\}$ ,  $E_2 = \{b_2\}$ . For this event, both players have a blocking plan  $(a_1, b_1)$  and  $E$  satisfies self-selection, as player 1 knows that player 2 has received recommendation  $b_2$ . Hence, coalition  $\{1, 2\}$  blocks the correlated strategy at the realization  $(a_2, b_2)$  and the correlated strategy is not a strong correlated equilibrium.<sup>5</sup>

---

<sup>5</sup>In fact, Moreno and Wooders (1996) use this example to show that there exist games with no ex ante coalition proof correlated equilibrium. The fact that a strong correlated equilibrium exists in Einy and Peleg (1995)'s sense shows that their definition makes blocking extremely difficult.

## 4 Strong Correlated Equilibria

Strong correlated equilibrium is a very demanding concept as any joint correlated deviation can be used to upset the initial correlated strategy. In fact, even games possessing pure strategy strong Nash equilibria may not have any ISCE as we show in the following examples. We then go on to show that there is in fact no relationship at all between the existence of ISCE and strong Nash equilibrium by constructing a game which has an ISCE but no strong Nash equilibrium. Finally, we show the existence of a class of games where the set of ISCE is non-empty.

**Example 3** *Let  $n = 2$ , and each player can take action  $a$  or  $b$ .*

	a	b
a	$(1 + \alpha, \alpha)$	$(1, 1)$
b	$(0, 0)$	$(\alpha, 1 + \alpha)$

**Example 4**

	a	b
a	$(1 - \alpha, 1 - \alpha)$	$(1, 0)$
b	$(0, 1)$	$(-\alpha, -\alpha)$

Example 3 is a game where both players have a favorite action (action  $a$  for player 1 and action  $b$  for player 2), and enjoy a positive externality of  $\alpha < 1$  if the other player chooses the same action. Example 4 is a game where both players prefer action  $a$  and suffer a negative externality of  $\alpha$  if the other player chooses the same action. Both games admit a unique strong Nash equilibrium,  $(a, b)$  in Example 3 and  $(a, a)$  in Example 4. However, if  $\alpha > 1/2$ ,  $(a, b)$  is not a strong correlated equilibrium in Example 3 because both players have an incentive to deviate to the correlated strategy putting equal weight on  $(a, a)$  and  $(b, b)$ . Similarly, if  $\alpha < 1/2$ ,  $(a, a)$  is not a strong correlated equilibrium in Example 4 because it is blocked by a correlated strategy putting equal weight on  $(a, b)$  and  $(b, a)$ . Furthermore, because  $(a, b)$  in Example 3 and  $(a, a)$  in Example 4 are dominant strategy equilibria, there is no other candidate correlated equilibrium, and hence the games do not admit strong correlated equilibria for some ranges of the parameter  $\alpha$ .



Examples 3 and 4 illustrate a very simple fact. When agents can choose a correlated deviation, they can block more easily than when they can only deviate by choosing mixed strategies. This explains why strong Nash equilibria may fail to be strong correlated equilibria.

On the other hand, the initial correlated strategy may give agents a higher payoff than a pair of mixed strategies. Hence, as the following example shows, there exist games which do not admit a strong Nash equilibrium, but for which a strong correlated equilibrium exists.

**Example 5** Consider a three-player game where player 1 chooses the row, player 2 the column and player 3 the matrix, with payoffs:

	$b_1$	$b_2$
$a_1$	(1.5, 1.5, 0.5)	(0, 0, 0)
$a_2$	(0, 2, 0)	(1, 1, 1)
	$b_1$	$b_2$
$a_1$	(0.5, 1.5, 1)	(0, 0, 0)
$a_2$	(1, 0, 0)	(2, 1, 0.5)

This game does not admit a strong Nash equilibrium. To check this, note that there is no pure strategy Nash equilibrium. There is a unique mixed strategy equilibrium in which 1 plays  $a_1$  with probability 0.136, 2 plays  $b_1$  with probability 0.864 and 3 plays  $c_1$  with probability 0.378.<sup>6</sup> The expected payoff vector corresponding to these probabilities is not efficient since the strategy vector  $(a_2, b_2, c_1)$  yields a payoff of 1 to each of the players.

Consider the correlated strategy  $\mu$  which places equal probability on the strategy vectors  $\{(a_1, b_1, c_1), (a_2, b_2, c_1), (a_1, b_1, c_2), (a_2, b_2, c_2)\}$ . This yields the expected payoff vector (1.25, 1.25, 0.75).

Note first that  $\mu$  is a correlated equilibrium, and hence is immune to deviations by singletons. Consider next deviations by pairs of players. The realized recommendations for 1 and 2 are either  $(a_1, b_1)$  or  $(a_2, b_2)$ . So, 1 and 2 can both infer the recommendation received by the other player from their own recommendation. If the joint recommendations are  $(a_1, b_1)$ , then the expected payoffs are 1 and 1.5 to 1 and 2 respectively since 3 receives the recommendations  $c_1$  and  $c_2$  with equal probability. No correlated deviation

---

<sup>6</sup>These numbers are correct to three decimal places.

by 1 and 2 can bring both of them higher payoffs. A similar analysis holds for  $(a_2, b_2)$ .

Also, 1 and 3 cannot construct any blocking plan. For suppose 1 has received the recommendation  $a_1$ . If 3 has received  $c_1$ , then this gives 1 the highest possible payoff given that 2 has received recommendation  $b_1$ . Hence, 1 does not want to deviate. If 3 has received  $c_2$ , then 3 has received the highest possible payoff. This is known to both and so there is no credible blocking plan if 1 receives recommendation  $a_1$ . A similar analysis holds if 1 has received  $a_2$ .

Consider the coalition  $\{2, 3\}$ . If 2 receives the recommendation  $b_1$ , then 2 receives 1.5 irrespective of what action is chosen by 3 given that 1 will choose  $a_1$ . Similarly, 2 will receive 1 irrespective of the action chosen by 3 since 1 chooses action  $a_2$  in this case. Hence, 2 and 3 have no blocking plan.

Finally consider the coalition  $\{1, 2, 3\}$ . If 1 receives the recommendation  $a_1$ , he knows that 2 receives  $b_1$ . Furthermore, by self-selection, the only recommendation which leads player 3 to deviate is  $c_1$ . At this event, the players receive payoffs  $(1.5, 1.5, 0.5)$ . We show that there is no correlated strategy which gives all players a higher payoff. The only pure strategy profiles we need to consider are profiles which give one of the players a higher payoff than  $(a_1, b_1, c_1)$ . Let then  $p_1 = \eta(a_2, b_1, c_1)$ ,  $p_2 = \eta(a_2, b_2, c_1)$ ,  $p_3 = \eta(a_1, b_1, c_2)$  and  $p_4 = \eta(a_2, b_2, c_2)$ . For a correlated strategy to dominate  $(a_1, b_1, c_1)$  for all the players we then need:

$$\begin{aligned} p_2 + 0.5p_3 + 2p_4 &\geq 1.5 \\ 2p_1 + p_2 + 1.5p_3 + p_4 &\geq 1.5 \\ p_2 + p_3 + 0.5p_4 &\geq 0.5 \\ p_1 + p_2 + p_3 + p_4 &= 1. \end{aligned}$$

The third inequality implies that

$$p_1 \leq p_2 + p_3$$

so  $p_1 \leq 1/2$ . The first inequality amounts to

$$2p_1 + 0.5p_3 \geq 1.5$$

which cannot be satisfied if  $p_1 \leq 1/2$ . If 1 receives the recommendation  $a_2$ , the only event which satisfies self-selection is  $(a_2, b_2, c_2)$  at which 1 receives his optimal payoff and hence has no incentive to deviate.

Although the previous examples show that there is no logical relationship between the existence of strong Nash equilibria and ISCE, we now show that there is an interesting class of games for which strong Nash equilibria are also ISCE. These are the class of games with positive externalities for which strong Nash equilibria are known to exist. While example 3 show that in general ISCE do not exist in such games, we provide sufficient conditions for the existence of ISCE.

Following Konishi et al. (1997a) and (1997b), we consider games where all agents have the same action set,  $A_i = A$  for all  $i$  in  $N$ , and agents' utilities *only depend on their own action and the number of players who have chosen the same action*,  $u_i(a) = V_i(a_i, n(a_i))$  where  $n(a_i)$  denotes the number of agents who have chosen action  $a_i$ . If  $V_i$  is increasing in  $a_i$ , the game is a game with positive externalities. If  $V_i$  is decreasing in  $a_i$ , it is a game with negative externalities.

A simple illustration of games with positive externalities are the choices of standards (by firms) or products (by consumers) when consumers derive a positive utility from the number of consumers choosing the same product (Farrell and Saloner (1985) and Katz and Shapiro (1985)). Positive externalities also arise in local public good economies, when agents choose a jurisdiction, and share the fixed cost of the public project with other members of the jurisdiction (Alesina and Spolaore (2003), Haimanko, Le Breton and Weber (2004)). When  $|A| = 2$ , Konishi et al. (1997a) show that every game admits a pure strategy strong Nash equilibrium (Proposition 2.2 p. 168).

In the next section, we will show that ICPCCE exist in *all* games with positive externalities provided each  $A_i$  consists of the same two actions. In the next proposition, we further restrict the class of games with positive externalities. In particular, we assume that the positive externality accruing to individual  $i$  is *separable* in the number of individuals taking the same action as individual  $i$ , so that

$$V_i(a_i, n(a_i)) = v_i(a_i) + f(n(a_i)) \text{ where } f' > 0$$

Now, let  $N_1 = \{i \in N | v_i(a_1) > v_i(a_2)\}$ , while  $N_2 = \{i \in N | v_i(a_2) > v_i(a_1)\}$ .

Notice that if  $a$  is to be a strong Nash equilibrium, then individuals in  $N_1$  and  $N_2$  must choose actions  $a_1$  and  $a_2$  respectively when  $f(k) = 0$  for all  $k$ . Of course, this must also be an ISCE. Now, if the effect of th externality is small, then this will continue to be an ISCE. This is basically the content of the next proposition.

**Proposition 1** *Let  $A_i = \{a_1, a_2\}$  for each  $i \in N$ . Then, there exists  $\bar{f}$  with  $\bar{f}' > 0$  such that for all  $f \leq \bar{f}$  with  $f' > 0$ , an ISCE exists.*

**Proof.** Since each  $A_i$  has only 2 elements, we can normalise utility functions by setting  $u_i(a_1) = 0$  for all  $i$ . Moreover, without loss of generality, assume that  $u_i(a_2) \leq u_{i+1}(a_2)$  for all  $i = 1, \dots, n - 1$ .

Let  $a^*$  be a strong Nash equilibrium. Then, either all individuals take the same action or there is  $k^*$  such that all  $i \leq k^*$  choose  $a_1$ , while the others choose  $a_2$ . If all individuals choose the same action, then that must also be an ISCE.

Otherwise, let  $(B_1, B_2)$  be the non-empty sets of individuals choosing actions  $a_1, a_2$  respectively. Then,  $u_{k^*}(a_2) < u_{k^*+1}(a_2)$ , and either  $u_{k^*}(a_2) < 0$  or  $u_{k^*+1}(a_2) > 0$ . Without loss of generality, assume  $u_{k^*+1}(a_2) = d > 0$ .

Now, if a coalition  $T'$  has a profitable blocking plan, then  $T' \cap B_1 \neq \emptyset$  and  $T' \cap B_2 \neq \emptyset$ . Moreover, if any coalition  $T'$  has a profitable blocking plan, then there is also a *connected*<sup>7</sup> coalition  $T$  including  $k^*$  and  $k^* + 1$  which has a profitable blocking plan  $\eta_T$ .

Let  $p > 0$  be the total probability with which  $k^* + 1$  chooses action  $a_1$ . Then,  $pd$  is the loss in utility suffered by  $k^* + 1$  from the switch in choice of action. Let  $\alpha_1, \dots, \alpha_J$  be the probabilities with which  $k^* + 1$  chooses the same action as sets of individuals of size  $s_1, \dots, s_J$ . Then, in order for  $\eta_T$  to be improving for  $k^* + 1$ , we need

$$\sum_{j=1}^J \alpha_j f(s_j) > pd + f(|B_1|) \quad (2)$$

Consider *all* coalitions  $\bar{T}$  including  $k^* + 1$  which have improving blocking plans. For each such coalition, and each improving blocking plan, there will be an inequality of the form 2. Clearly, one can choose an “externality” function which violates all these inequalities. This function fulfills the role of  $\bar{f}$  in the proposition. ■

## 5 Coalition Proof Correlated Equilibrium

We now turn our attention to the less demanding concept of coalition proof correlated equilibrium. Our first result parallels the main existence result of

---

<sup>7</sup>That is, if  $T$  includes  $i$  and  $i + 2$ , then  $T$  also includes  $i + 1$  for all  $i$ .

Moreno and Wooders (1996) and establishes a connection between interim coalition proof correlated equilibrium, and the elimination of strictly dominated strategies.

**Definition 8** Let  $B = \prod_{i \in N} B_i \subset A$ . An action  $a'_i \in B_i$  is strictly dominated in  $B$  if there exists  $\sigma_i \in \Delta B_i$  such that for each  $a_{-i} \in B_{-i}$ ,

$$\sum_{a_i \in B_i} \sigma_i(a_i) u_i(a_i, a_{-i}) > u_i(a'_i, a_{-i}).$$

**Definition 9** The set  $A^\infty$  of action profiles surviving iterated elimination of dominated strategies is defined by  $A^\infty = \prod_{i \in N} A_i^\infty$  where  $A_i^\infty = \bigcap_{t=0}^\infty A_i^t$  and  $A_i^t$  is the set of actions that are not strictly dominated in  $A^{t-1} = \prod_{i \in N} A_i^{t-1}$  with  $A_i^0 = A_i$ .

The following Proposition is analogous to the main existence result of Moreno and Wooders (1996) (Corollary page 92).

**Proposition 2** Let  $a^*$  be a pure strategy action profile that Pareto-dominates any other pure strategy action profile in  $A^\infty$ . Then  $a^*$  is an interim coalition proof correlated equilibrium.

**Proof.** We first show that if there exists a self-enforcing blocking plan  $\eta_S$  against  $a^*$  which is preferred by the deviating coalition, then this blocking plan must have a support contained in  $A_S^\infty$ . Suppose by contradiction that  $\text{supp}(\eta_S)$  is not contained in  $A_S^\infty$  and let  $t^*$  be the largest value for which  $\text{supp}(\eta_S) \subset A_S^{t^*}$ . Then there exists an agent  $i$  in  $S$ , and an action  $a'_i$  that agent  $i$  uses in the blocking plan  $\eta_S$  such that  $a'_i$  is strictly dominated by  $\sigma_i$  in  $A^{t^*}$ . Let  $\mu_i$  be player  $i$ 's prior beliefs about the recommendation of other players, after  $a$  and  $\eta_S$  have been chosen. Suppose that, after she receives recommendation  $a'_i$ , player  $i$  deviates to playing  $\sigma_i$ . This plan is self-enforcing (because it only involves one player), and the event  $E = \{a'_i\} \times A_{-i}$  trivially satisfies self-selection. Hence, to show that that player  $i$  has a self-enforcing blocking plan, we only need to check that her utility strictly increases after the deviation. First notice that, because  $a^* \in A^\infty$ ,  $a_{-S}^* \in A_{-S}^{t^*}$ , and, by

construction, for any  $a_{S \setminus i}$  in the support of the blocking plan  $\eta_S$ ,  $a_{S \setminus i} \in A_{S \setminus i}^{t^*}$ . Hence,

$$\begin{aligned} U_i(\mu_i|a'_i, A_{-i}) &= \sum_{a_{-i} \in A_{-i}} \tilde{\mu}(a_{-i}|a'_i, A_{-i}) u_i(a'_i, a_{-i}) \\ &= \sum_{a_{-i} \in A_{-i}^{t^*}} \tilde{\mu}(a_{-i}|a'_i, A_{-i}) u_i(a'_i, a_{-i}). \end{aligned}$$

Similarly, as  $\sigma_i \in \Delta A_i^{t^*}$

$$\begin{aligned} U_i(\sigma_i|a'_i, A_{-i}) &= \sum_{a_{-i} \in A_{-i}} \sum_{a_i \in A_i} \tilde{\mu}(a_{-i}|a'_i, A_{-i}) \sigma_i(a_i) u_i(a_i, a_{-i}) \\ &= \sum_{a_{-i} \in A_{-i}^{t^*}} \sum_{a_i \in A_i^{t^*}} \tilde{\mu}(a_{-i}|a'_i, A_{-i}) \sigma_i(a_i) u_i(a_i, a_{-i}). \end{aligned}$$

Because  $\sigma_i$  strictly dominates  $a'_i$  in  $A^{t^*}$ ,  $u_i(a'_i, a_{-i}) < \sum_{a_i \in A_i^{t^*}} \sigma_i(a_i) u_i(a_i, a_{-i})$  for all  $a_{-i}$  in  $A_{-i}^{t^*}$ . Hence,  $U_i(\sigma_i|a'_i, A_{-i}) > U_i(\mu_i|a'_i, A_{-i})$ , establishing the result.

Next, it is clear that if a pure action profile  $a^*$  Pareto-dominates any other pure action profile in  $A^\infty$ , it also Pareto-dominates any correlated strategy with support contained in  $A^\infty$ . This suffices to show that there does not exist any coalition  $S$  and blocking plan  $\eta_S$  with support in  $A^\infty$  for which

$$u_i(a^*) < \sum_{a_S \in A_S} \eta_S(a_S) u_i(a_{-S}^*, a_S) \text{ for all } i \text{ in } S.$$

so that  $a^*$  is a coalition proof correlated equilibrium. ■

Some remarks are in order. First, our sufficient condition is stronger than that of Moreno and Wooders (1996) who do not require the existence of a *pure action profile* which Pareto-dominates all other pure action profiles in  $A^\infty$ , but only the existence of a correlated strategy which Pareto-dominates all other action profiles in  $A^\infty$ . Second, our result shows that if a game is dominance-solvable (by iterative elimination of *strictly* dominated strategies), then the unique outcome surviving the elimination of dominated strategies is a coalition-proof correlated equilibrium. Finally, the existence results of Milgrom and Roberts (1996), show existence of a pure strategy ex ante coalition proof correlated equilibrium in games with strategic complementarities admitting a unique Nash equilibrium, or for which utilities are monotonic

in the actions of the other players (Theorem 2, P. 124). These results rely on the same argument as the one given here – the existence of a pure action profile which Pareto-dominates any other action profile, and hence also apply to our setting where coalitions deviate at the interim stage.

The next proposition shows that games with positive externalities and two actions always admit a coalition proof correlated equilibrium.

**Proposition 3** *Suppose  $A_i = \{a_1, a_2\}$  for all  $i \in N$ . Then, any game with positive externality admits an interim coalition proof correlated equilibrium.*

**Proof.** From Konishi et al. (1997a), we know that the game admits pure strategy strong Nash equilibria. Pick one of these strong Nash equilibria, (characterized by a partitioning of the agents,  $\{B_1, B_2\}$ ) with the property that  $\max\{b_1, b_2\} \geq \max\{c_1, c_2\}$  for all other strong Nash equilibria  $\{C_1, C_2\}$ .<sup>8</sup> In words, among all strong Nash equilibria, we choose one with the largest number of players choosing the same action. Let  $T$  be a coalition which has a profitable blocking plan  $\eta_T$  against the pure strategy recommendation which results in the partition  $\{B_1, B_2\}$ . We first claim that  $T \cap B_i \neq \emptyset$  for  $i = 1, 2$  – the deviating coalition must involve players from both sides moving. Suppose by contradiction that  $T \subset B_1$ . (A similar argument would hold if  $T \subset B_2$ ). Because  $\{B_1, B_2\}$  is a strong Nash equilibrium, there must exist an agent  $i \in T$  for whom  $u_i(a_1, b_1) \geq u_i(a_2, b_2 + t)$ . But if  $T \subset B_1$ , then for any outcome  $\{C_1, C_2\}$  in the support of the blocking plan,  $c_1 \leq b_1$ . Hence, for all outcomes in the support of the blocking plan, agent  $i$  either chooses action  $a_1$  in a group containing  $c_1 \leq b_1$  agents, or chooses action  $a_2$  in a group containing  $c_2 \leq b_2 + t$  agents. In either case, his utility is less than or equal to  $u_i(a_1, b_1)$  and he cannot participate in the blocking plan.

Let  $T_1 = T \cap B_1$  and  $T_2 = T \cap B_2$ . Without loss of generality, suppose that  $b_1 \geq b_2$ . Consider the partition  $\{B_1 \cup T_2, B_2 \setminus T_2\}$ . By assumption, this partition is not a strong Nash equilibrium. We will show that there exists a deviating coalition  $S \subset T_2$ . First notice that  $B_1 \cap S = \emptyset$ . If members of  $B_1$  had an incentive to deviate collectively in the partition  $\{B_1 \cup T_2, B_2 \setminus T_2\}$ , they would also have an incentive to deviate in the partition  $\{B_1, B_2\}$ , contradicting the fact that  $\{B_1, B_2\}$  is a strong Nash equilibrium. Notice furthermore that if there exists a deviating coalition  $S$  containing members of  $T_2$  and

---

<sup>8</sup>Here,  $B_i, C_i$  denote the set of agents choosing action  $a_i$  for  $i = 1, 2$ . Also, we use  $b, c$  for the cardinality of the sets  $B, C$ .

$B_2 \setminus T_2$ , there also exists another deviating coalition  $S'$  only containing members of  $T_2$ . Hence, if there is no deviating coalition  $S$  satisfying  $S \subset T_2$ , it must be that all deviating coalitions are included in  $B_2 \setminus T_2$ . Consider then the largest deviating coalition,  $S$ , for which  $u_i(a_1, b_1 + t_2 + s) > u_i(a_2, b_2 - t_2)$  for all  $i \in S$ , and the resulting partition  $\{B_1 \cup T_2 \cup S, B_2 \setminus T_2 \setminus S\}$ . Again, this partition is not a strong Nash equilibrium, and there must exist a deviating coalition  $U$ . Now, as no subset of players of  $B_1 \cup T_2$  wanted to deviate from the partition  $\{B_1 \cup T_2, B_2 \setminus T_2\}$ , there is no collective deviation including members of  $B_1 \cup T_2$ . Furthermore, as  $u_i(a_1, b_1 + t_2 + s) > u_i(a_2, b_2 - t_2)$ , there is no collective deviation from members of  $S$  either. Hence, we must have  $U \subset B_2 \setminus T_2 \setminus S$ . The process can be repeated until the formation of the partition  $\{N, \emptyset\}$ , at which point we reach a contradiction, because this partition is not a strong Nash equilibrium, and it is impossible to construct a deviating coalition. Hence, there must exist a deviating coalition  $S$  from  $\{B_1 \cup T_2, B_2 \setminus T_2\}$  such that  $S \subset T_2$ .

Finally, we show that this implies that there exists a self-enforcing blocking plan,  $\eta_S$  against the original deviation  $\eta_T$ . Consider the plan where members of  $S$  always choose action  $a_2$ . Every member  $i$  of  $S$  will then receive at least  $u_i(a_2, b_2 - t_2 + s)$  after deviating. By sticking to the recommendation  $a_1$ , he would receive at most  $u_i(a_1, b_1 + t_2 - (s - 1)) \leq u_i(a_1, b_1 + t_2)$ . Because  $S$  is a deviating coalition from the partition  $\{B_1 \cup T_2, B_2 \setminus T_2\}$ ,  $u_i(a_2, b_2 - t_2 + s) > u_i(a_1, b_1 + t_2)$  for all  $i \in S$ , and hence the blocking plan  $\eta_S$  is profitable. Finally, the blocking plan is self-enforcing, because no subcoalition of  $S$  can guarantee a higher payoff to all its members, as this would involve some players moving back to action  $a_1$ . ■

## 6 Conclusion

This paper proposes new concepts of strong and coalition-proof correlated equilibria where agents form coalitions at the interim stage and share information about their recommendations in a credible way. Our analysis highlights the difference between the coalitional deviations at the ex ante stage studied by Moreno and Wooders (1996) and Milgrom and Roberts (1996), and the coalitional deviations at the interim stage. Whereas ex ante coalition-proof correlated equilibria always exist in two-player games, we provide an example of a two-player game which does not admit any interim coalition-proof correlated equilibrium. Following the same line of argument



as Moreno and Wooders (1996), we provide a sufficient condition for existence based on the existence of a Pareto-dominant strategy in the set of strategies surviving iterative elimination of dominated strategies. However, the sufficient condition for existence of an interim coalition-proof correlated equilibrium is strictly stronger than the sufficient condition uncovered by Moreno and Wooders (1996) for the existence of an ex ante coalition-proof correlated equilibrium. Finally, we identify a class of games with positive externalities, already studied by Konishi et al. (1997), which always admit interim coalition-proof correlated equilibria.

In our view, the study of coalitional deviations in games with communication is a first step towards the study of coalitional deviations in general Bayesian games. Our definition of credible information sharing could easily be adapted to a setting where agents have different (privately known) types, and our equilibrium concepts could easily be applied to general games with incomplete information. We plan to pursue this agenda in future research, thereby making progress on the study of cooperation and coalition formation among agents with incomplete information.

## 7 References

- Alesina, A. and E. Spolaore (2003) *The Size of Nations*, Cambridge and London: MIT Press.
- Bernheim, D. , B. Peleg and M. Whinston (1987) "Coalition-proof Nash Equilibria: Concepts", *Journal of Economic Theory* **42**, 1-12.
- Chwe, M. (1994) "Farsighted Coalitional Stability", *Journal of Economic Theory* **64**, 299-325.
- Dutta, B. and R. Vohra (2005) "Incomplete Information, Credibility and the Core", *Mathematical Social Sciences* **50**, 148-165.
- Einy, E. and B. Peleg (1995) "Coalition Proof Communication Equilibria", in *Social Choice, Welfare and Ethics*, (W. Barnett, H. Moulin, M. Salles and N. Schofield, eds.), Cambridge, New York and Melbourne: Cambridge University Press.
- Farrell, J. and G. Saloner (1985) "Standardization, Compatibility and Innovation", *Rand Journal of Economics* **16**, 70-83.
- Haimanko, O. , M. Le Breton and S. Weber (2004) "Voluntary Formation of Communities for the Provision of Public Projects", *Journal of Economic Theory* **115**, 1-34.

- Katz, M. and C. Shapiro (1986) "Technology Adoption in the Presence of Network Externalities", *Journal of Political Economy* **94**, 822-841.
- Konishi, H. , M. Le Breton and S. Weber (1997a) "Pure Strategy Nash Equilibria in Group Formation Games with Positive Externalities", *Games and Economic Behavior* **21**, 161-182.
- Konishi, H., Le Breton, M. and S. Weber (1997b) "Equilibria in a Model with Partial Rivalry", *Journal of Economic Theory* **72**, 225-237.
- Milgrom, P. and J. Roberts (1996) "Coalition Proofness and Correlation with Arbitrary Communication Possibilities", *Games and Economic Behavior* **17**, 113-128.
- Moreno, D. and J. Wooders (1996) "Coalition Proof Equilibrium", *Games and Economic Behavior* **17**, 80-112.
- Ray, I. (1996) "Coalition Proof Correlated Equilibrium: A Definition", *Games and Economic Behavior* **17**, 56-79.
- Ray, D. and R. Vohra (1997) "Equilibrium Binding Agreements", *Journal of Economic Theory* **73**, 30-78.
- Wilson, R. (1978) "Information, Efficiency and the Core of an Economy", *Econometrica* **46**, 807-816.