All-stage strong correlated equilibrium

Heller, Yuval

Tel-Aviv University, School of Mathematical Sciences

1 June 2008

All-Stage Strong Correlated Equilibrium

Yuval Heller

School of Mathematical Sciences, Tel-Aviv University, Tel-Aviv 69978, Israel.

Phone: 972-3-640-5386. Fax: 972-3-640-9357.

Email: helleryu@post.tau.ac.il

Abstract

A strong correlated equilibrium is a correlated strategy profile that is immune to joint deviations. Different notions of strong correlated equilibria have been defined in the literature. One major difference among those definitions is the stage in which coalitions can plan a joint deviation: before (*ex-ante*) or after (*ex-post*) the deviating players receive their part of the correlated profile. In this note we show that an *ex-ante* strong correlated equilibrium (Moreno D., Wooders J., 1996. Games Econ. Behav. 17, 80-113) is immune to deviations at all stages of any pre-play signalling process that implements it. Thus the set of *ex-ante* strong correlated equilibria is included in all other sets of strong correlated equilibria.

Key words: coalition-proofness, strong correlated equilibrium, common knowledge, incomplete information, non-cooperative games. JEL classification: C72, D82.

[1] This work is in partial fulfillment of the requirements for the Ph.D. in mathematics at Tel-Aviv University. I would like to thank Eilon Solan for his careful supervision, for the continuous help he offered, and for many insightful discussions. I would also like to express my deep gratitude to to Ehud Lehrer and Yaron Azrieli for many useful comments, discussions and ideas.

1 Introduction

The ability of players to communicate prior to playing a non-cooperative game, influences the set of self-enforcing outcomes of that game. Communication allows players to correlate their play, and to implement a correlated strategy profile as a non-binding agreement. For such an agreement to be self-enforcing, it has to be stable against coalitional deviations. Two notions in the literature describe such self-enforcing agreements: a *strong correlated equilibrium* is a correlated profile that is stable against *all* coalitional deviations, while a *coalition-proof correlated equilibrium* is stable against *self-enforcing* coalitional deviations (Bernheim et al., 1987). For a coalition of a single player, any deviation is self-enforcing. For a larger coalition, a deviation is self-enforcing if there is no further self-enforcing and improving deviation by one of its proper sub-coalitions. The main focus of this note is on the former notion.

A correlated strategy profile can be implemented by a mediator who privately recommends each player which action to play. It can also be implemented by a pre-play signaling process, *a revealing protocol*, that includes payoff-irrelevant private and public signals ("sunspots"). Each player deduces his recommended action from the signals he has received. In the existing literature (referred to below) it is assumed that all signals are simultaneously received by all players. However, a revealing protocol may be more complex. Few examples are: The recommendations may be revealed consecutively by private signals in a pre-specified order (see e.g., the polite cheap-talk protocol that implement strong correlated equilibria in Heller, 2008); Each private signal may include partial information about the player's recommended action; The order in which recommendations are revealed may depend on a private lottery.

So that a revealing protocol can implement a correlated equilbirum it should satisfy two properties. First, at the end of the protocol each player should know the action recommended to him. Second, no player should obtain any information about the actions recommended to the other players, except the conditional probability, given his own recommended action.

When all the players receive their recommended actions simultaneously, a coalition of players may communicate, share their information, and plan a joint deviation before, or after, the recommendations are revealed. In Milgrom and Roberts (1996), Moreno and Wooders (1996), and Ray (1996) it is assumed that players may only plan deviations at the *ex-ante* stage, before receiving the recommendations. In Einy and Peleg (1995), Ray (1998), and Bloch & Dutta (2008) it is assumed that players may only plan deviations at the *ex-post* stage, after receiving the recommendations. [2]

[2] This stage is called *interim* stage in some of referred papers.

When the players receive several signals, not necessarily simultaneously, they may communicate, share information, and plan coalitional deviations at different stages of the revealing protocol. By sharing information, a coalition of players may get information about the actions recommended to players outside the coalition, and may use this information to implement profitable deviations. Similar to the existing literature of simultaneous revealing protocols, we focus on protocols in which sharing information among deviating players does not allow them to obtain any information about the actions recommended to the other players, except the conditional probability, given their own recommended actions.

The use of a joint deviation requires the unanimous agreement of all members of the deviating coalition. A player agrees to be part of a joint deviation if, given his own information the deviation is profitable. His agreement to participate in the joint deviation is a public signal to all the deviators about that fact. Thus, if a joint deviation is implemented, then it is common knowledge among the deviators that each of them believes that the deviation is profitable (see the example in Sect. 3 for more details). We model the information structure of the deviating players by an incomplete information model (with the common prior assumption) *à* la Aumann (1987).

In the spirit of the concept of strong correlated equilibrium, we assume that deviations are binding: A deviation is implemented with the assistance of a new mediator. The deviators truthfully report their information to the new mediator, and they are bound to follow his recommendations, even if new information at a later stage makes it unprofitable. If the deviators are not bound to follow the new mediator's recommendations, the solution concept is close in spirit to the coalition-proof notion.

A correlated strategy profile is an *all-stage strong correlated equilibrium* if, for every revealing protocol that implements it, and for each stage of the protocol, there is no coalition with a profitable deviation. A correlated strategy profile is an *ex-ante strong correlated equilibrium* (Moreno and Wooders, 1996) if no coalition has a profitable deviation at the *ex-ante* stage. Our result shows that the two notions coincide: an *ex-ante* strong correlated equilibrium is resistant to deviations at all stages of any revealing protocol that implements it. This implies an inclusion relation among the different notions of strong correlated equilibria, and a "robustness" of the *ex-ante* notion (as discussed in Sect. 6).

One could hope that similar results may be obtained for the coalition-proof notions. However, in Sect. 5 we demonstrate that the *ex-ante* coalition-proof notion is not appropriate when deviations can be planned at all stages.

A related work is the seminal paper of Holmstrom and Myerson (1983) that developed a few concepts of efficiency according to how much information is revealed and shared among the players. The result presented in Sect. 4-5 of their paper, when being adapted to our framework, states that if a correlated profile is resistant to deviations of the grand coalition at the *ex-ante* stage, then it is also resistant to such deviations at the *ex-post* stage. The contribution of this note is twofold. First, the modeling of the different kinds of strong correlated equilibria by an incomplete information model *à* la Aumann (1987). Second, our result extends Holmstrom and Myerson's result in two ways: we prove the resistance at all stages (not only at the ex-post stage), and simultaniously against deviations of all coalitions.

The note is organized as follows. Section 2 presents the model and the result. The result is demonstrated with an example in Section 3, and proven in Section 4. We deal with the coalition-proof notion in Section 5, and discuss the implications of the result in Section 6.

2 Model and Definitions

A game in strategic form G is defined as: $G = \left(N, (A^i)_{i\in N}, (u^i)_{i\in N}\right)$, where N is the finite and non-empty set of players. For each $i \in N$, A^i is player i's finite and non-empty set of actions, and u^i is player i's utility (payoff) function, a real-valued function on $A = \prod_{i\in N} A^i$. The multi-linear extension of u^i to $\Delta(A)$ is still denoted by u^i. A member of $\Delta(A)$ is called a (correlated) strategy profile. A coalition S is a non-empty member of 2^N. Given a coalition S, let $A^S = \prod_{i\in S} A^i$, and let $-S = \{i \in N \mid i \notin S\}$ denote the complementary coalition. A member of $\Delta(A^S)$ is called a (correlated) S-*strategy profile*. Given $q \in \Delta(A)$ and $a^S \in A^S$, we define $q_{|S} \subseteq \Delta(A^S)$ to be $q_{|S}(a^S) = \sum_{a^{-S}\in A^{-S}} q(a^S, a^{-S})$, and for simplicity we omit the subscript: $q(a^S) = q_{|S}(a^S)$. Given a^S s.t. $q(a^S) > 0$, we define $q(a^{-S}|a^S) = {q(a^S,a^{-S})}/{q(a^S)}$.

A *state space* is a probability space, $(\Omega, \mathcal{B}, \mu)$ that describes all parameters that may be the object of uncertainty on the part of the players. We interpret Ω as the space of all possible states of the world, $\mathcal{B}$ as the σ-algebra of all measurable events, and μ as the common prior. Given a non-null event $E \in \mathcal{B}$ and a random variable $\mathbf{x} : \Omega \to X$ (where X is a finite set), let $\mathbf{x}(E) \in \Delta(X)$ denote the posterior distribution of $\mathbf{x}$ conditioned on the event E. The implementation of an *agreement* (a correlated strategy profile) by a mediator or by a signaling process is modeled by a random variable $\mathbf{a} : \Omega \to A$, which satisfies that the prior distribution $\mathbf{a}(\Omega)$ is equal to the agreement distribution.

Definition 1 Let G be a game, $q \in \Delta(A)$ an agreement, and $(\Omega, \mathcal{B}, \mu)$ a state space. A *recommendation profile that implements* q is a random variable $\mathbf{a} = (\mathbf{a}^i)_{i\in N} : \Omega \to A$ that satisfies: $\mathbf{a}(\Omega) = q$.

A (joint) deviation of a coalition S is a random variable (in Ω) that is conditionally independent of $\mathbf{a}^{-S}$ given $\mathbf{a}^S$.

Definition 2 Let G be a game, $q \in \Delta(A)$ an agreement, $S \subseteq N$ a coalition, $(\Omega, \mathcal{B}, \mu)$ a state space, and $\mathbf{a} : \Omega \to A$ a recommendation profile that implements q. A *deviation* (of S from $\mathbf{a}$) is a random variable $\mathbf{d}^S = (\mathbf{d}^i)_{i \in S} : \Omega \to A^S$ that is conditionally independent of $\mathbf{a}^{-S}$ given $\mathbf{a}^S$.

The interpretation is as follows: If the players in S agree to use deviation $\mathbf{d}^S$, they implement it with the assistance of a new mediator. The new mediator receives the S-part of the recommendation profile, but he does not receive any information about the actions recommended to the other players. Thus, $\mathbf{d}^S$ may depend only on $\mathbf{a}^S$, but not on $\mathbf{a}^{-S}$.

When the members of a coalition S consider the implementation of a joint deviation, they are in a situation of incomplete information: each player may know his recommended action, and may have additional private information acquired when communicating with the other deviating players. We assume that the deviating players have no information about the actions recommended to the non-deviating players, except the conditional probability given the information they have about their recommended actions. We model this by the following definition of a consistent information structure.

Definition 3 Let G be a game, $q \in \Delta(A)$ an agreement, $S \subseteq N$ a coalition, $(\Omega, \mathcal{B}, \mu)$ a state space, and $\mathbf{a} : \Omega \to A$ a recommendation profile that implements q. *An information structure* (of S) is a $|S|$-tuple of partitions of Ω $(\mathcal{F}^i)_{i \in S}$, whose join ($\bigwedge_{i \in S} \mathcal{F}^i$, the coarsest common refinement of $(\mathcal{F}^i)_{i \in S}$) consists of non-null events. We say that $(\mathcal{F}^i)_{i \in S}$ is a *consistent information structure,* if $\forall \omega \in \Omega$, $\forall i \in S$, $\forall a \in A$, $\mathbf{a}\left(F^i(\omega)\right)(a) = \mathbf{a}^S\left(F^i(\omega)\right)\left(a^S\right) \cdot q(a^{-S} \mid a^S)$.

We interpret $\mathcal{F}^i$ as the information partition of player i; that is, if the true state of the world is $\omega \in \Omega$ then player i is informed of that element $F^i(\omega)$ of $\mathcal{F}^i$ that contains ω.

When each player considers whether the implementation of a deviation is profitable, he compares his conditional expected payoff when playing the original agreement and when implementing the deviation. A player agrees to deviate, only if the latter conditional expectation is larger. Formally, let G be a game, $q \in \Delta(A)$ an agreement, $S \subseteq N$ a coalition, $i \in S$ a player, $(\Omega, \mathcal{B}, \mu)$ a state space, $\mathbf{a} : \Omega \to A$ a recommendation profile, $\mathbf{d}^S : \Omega \to A^S$ a deviation, and $(\mathcal{F}^i)_{i \in S}$ a consistent information structure. The conditional expected payoff when all the players follow the agreement is:

$$u_f^i(\omega) = \int_{F^i(\omega)} u^i\left(\mathbf{a}(\omega)\right) d\mu$$

The conditional expected payoff when the members of S deviate, by implementing $\mathbf{d}^S$, and the players in $-S$ follow the agreement:

$$u_d^i(\omega) = \int_{F^i(\omega)} u^i\left(\left(\mathbf{d}^S, \mathbf{a}^{-S}\right)(\omega)\right) d\mu$$

If the players in S unanimously decide to implement a deviation in some state $\omega \in \Omega$, then it is common knowledge (in ω) that each player believes to earn more if the deviation is implemented. In that case we say that the joint deviation is profitable. Formally:

Definition 4 (Aumann 1976) Let G be a game, $S \subseteq N$ a coalition, $(\Omega, \mathcal{B}, \mu)$ a state space, $(\mathcal{F}^i)_{i \in S}$ an information structure, and $\omega \in \Omega$ a state. An event $E \in \mathcal{B}$ is *common knowledge* at ω if E includes that member of the meet $\mathcal{F}^{meet} = \bigwedge_{i \in S} \mathcal{F}^i$ that contains ω.

Definition 5 Let G be a game. $q \in \Delta(A)$ an agreement, $S \subseteq N$ a coalition, $(\Omega, \mathcal{B}, \mu)$ a state space, and $\mathbf{a} : \Omega \to A$ a recommendation profile that implements q. A deviation (of S) $\mathbf{d}^S$ is *profitable*, if there exists a consistent information structure $(\mathcal{F}^i)_{i \in S}$ and a state $\omega_0 \in \Omega$ such that it is common knowledge in ω_0 that $\forall i \in S$, $u_d^i(\omega) > u_f^i(\omega)$. In that case, we say that $\mathbf{d}^S$ *is* a *profitable deviation* (from the reccomendation profile $\mathbf{a}$) with respect to the information structure $(\mathcal{F}^i)_{i \in S}$.

We now define an all-stage strong correlated equilibrium as a strategy profile, from which no coalition has a profitable deviation.

Definition 6 Let G be a game. A strategy profile $q \in \Delta(A)$ is an *all-stage strong correlated equilibrium* if for every reccomendation profile $\mathbf{a} : \Omega \to A$ that implements q, no coalition $S \subseteq N$ has a profitable deviation.

A profile is an *ex-ante* strong correlated equilibrium, if no coalition has a profitable deviation at the *ex-ante* stage, when the players have no information about the recommendations.

Definition 7 Let G be a game and $(\Omega, \mathcal{B}, \mu)$ a state space. A profile $q \in \Delta(A)$ is an *ex-ante strong correlated equilibrium* if for every reccomendation profile $\mathbf{a} : \Omega \to A$ that implements q, no coalition S has a profitable deviation with respect to the *ex-ante* information structure $(\mathcal{F}^i)_{i \in S}$ that satisfies $\forall i, \mathcal{F}^i = \Omega$.

One can verify that Def. 7 is equivalent to the definition of Moreno and Wooders (1996). The definition immediately implies that an all-stage strong correlated equilibrium is also an *ex-ante* strong correlated equilibrium. The main result shows that the converse is also true, and thus the two notions coincide.

Theorem 8 *A correlated strategy profile is an ex-ante strong correlated equi-*

librium if and only if it is an all-stage strong correlated equilibrium.

3 An Example of the Main Result

In the following example we present an *ex-ante* strong correlated equilibrium in a 3-player game, and a specific deviation that is considered by the grand coalition during a revealing protocol. At first glance, one may think that this deviation is profitable to all the players conditioned on their posterior information at that stage, but a more thorough analysis reveals that this is not the case. The analysis in this example provides the intuition for the use of a model of incomplete information *à* la Aumann (1987), for the common knowledge requirement in Def. 5 of a profitable deviation, and for the main result.

Table 1 presents the matrix representation of a 3-player game, where player 1 chooses the row, player 2 chooses the column, and player 3 chooses the matrix.

Table 1
A 3-Player Game With An Ex-Ante Strong Correlated Equilibrium

	c_1			c_2			c_3		
	b_1	b_2	b_3	b_1	b_2	b_3	b_1	b_2	b_3
a_1	10,10,10	5, 20,5	0,0,0	5,5,20	0,0,0	0,0,0	0,0,0	0,0,0	0,0,0
a_2	20,5,5	0,0,0	0,0,0	0,0,0	0,0,0	0,0,0	0,0,0	0,0,0	0,0,0
a_3	0,0,0	0,0,0	0,0,0	0,0,0	0,0,0	0,0,0	0,0,0	0,0,0	7,11,12

Let q be the profile: $\left(\frac{1}{4}(a_1, b_1, c_1), \frac{1}{4}(a_2, b_1, c_1), \frac{1}{4}(a_1, b_2, c_1), \frac{1}{4}(a_1, b_1, c_2)\right)$, with an expected payoff of 10 to each player. Observe that q is an *ex-ante* strong correlated equilibrium:

- The profile q is a correlated equilibrium.
- No coalition of two players has a profitable deviation, because their uncertainty about the action recommended to the third player prevents them from earning together more than 20 by a joint deviation.
- The grand coalition cannot earn more than a total payoff of 30.

Now, consider a stage of a revealing protocol in which player 1 has received a recommendation to play a_1, player 2 has received a recommendation to play a_2, and player 3 has not received a recommendation yet. No player knows whether the other players have received their recommended actions.[3] At first glance, the implementation of the deviation $\mathbf{d}(\cdot) = (a_3, b_3, c_3)$, which gives a payoff of $(7, 11, 12)$, may look profitable to all the players: Conditioned on his

[3] It is common knowledge that each player has either received his recommended action or has not recieved any information about it.

recommended action (a_1), player 1 has an expected payoff of $6\frac{2}{3}$, and thus $\mathbf{d}$ is profitable to him. The same is true for player 2; Player 3 does not know his recommended action. His *ex-ante* expected payoff is 10, and he would earn a payoff of 12 by implementing $\mathbf{d}$.

However, a more thorough analysis reveals that $\mathbf{d}$ is unprofitable for player 3. Player 1 can only earn from $\mathbf{d}$ if he has received a recommendation to play a_1. Thus, if player 1 agrees to implement $\mathbf{d}$, then it is common knowledge that he has received a_1. The expected payoff of players 2 and 3, conditioned on that player 1 has received a_1, is $11\frac{2}{3}$. Thus, if player 2 agrees to implement $\mathbf{d}$ (with a payoff of 11) it is common knowledge that he has more information: his recommended action is a_2. Therefore player 3 knows that if the others agree to implement $\mathbf{d}$, then their recommended actions are (a_1, a_2). Conditioned on that, his expected payoff is 15, and thus $\mathbf{d}$ is unprofitable for himself.

4 Proof of the Main Result

We now prove the main result. As discussed earlier, one direction immediately follows from the definitions, and we only have to prove the other direction:

Theorem 9 *Every ex-ante strong correlated equilibrium is an all-stage strong correlated equilibrium.*

In other words: If a profitable deviation from an agreement $q \in \triangle(A)$ exists, then there also exists a profitable *ex-ante* deviation from q.

PROOF. Let $q \in \Delta(A)$ be a correlated profile that is not an all-stage strong correlated equilibrium in a game G, $(\Omega, \mathcal{B}, \mu)$ the state space, and $\mathbf{a} : \Omega \to A$ a recommendation profile that implements q. There exists a coalition $S \subseteq N$ with a profitable deviation $\mathbf{d}^S : \Omega \to A^S$ with respect to a consistent information structure $(\mathcal{F}^i)_{i \in S}$. This implies that there is a state $\omega_0 \in \Omega$, such that it is common knowledge in ω_0 that $\forall i$, $u^i_d(\omega) > u^i_f(\omega)$, i.e., $F^{meet}(\omega_0) \subseteq \left\{\omega \mid u^i_d(\omega) > u^i_f(\omega)\right\}$. For each deviating player $i \in S$, write $F^{meet} = F^{meet}(\omega_0) = \dot{\bigcup_j} F^i_j$ where the F^i_j are disjoint members of $\mathcal{F}^i$, and let $\omega^i_j \in F^i_j$ be a state in F^i_j. We now construct an *ex-ante* profitable deviation $\mathbf{d}^S_e$ with respect to the *ex-ante* information structure $(\mathcal{F}^i_e)_{i \in S}$, which satisfies $\forall i$, $\mathcal{F}^i_e = \Omega$: $\mathbf{d}^S_e(\omega) = \begin{cases} \mathbf{d}^S(\omega) \ \omega \in F^{meet}, \\ \mathbf{a}^S(\omega) \ \omega \notin F^{meet}. \end{cases}$

Observe that $\mathbf{d}^S_e$ and $\mathbf{a}^{-S}$ are conditionally independent given $\mathbf{a}^S$, thus $\mathbf{d}^S_e$ is well-defined. Let $u^i_{d_e}(\omega)$, $u^i_{f_e}(\omega)$ be the conditional utilities of the players with respect to $(\mathcal{F}^i_e)_{i \in S}$. We finish the proof by showing that $\mathbf{d}^S_e$ is profitable:

$$
\begin{aligned}
u^i_{d_e}(\omega) - u^i_{f_e}(\omega) &= \int_{F^i_e(\omega)} \left(u^i\left(\left(\mathbf{d}^S_e, \mathbf{a}^{-S}\right)(\omega)\right) - u^i\left(\mathbf{a}(\omega)\right)\right) d\mu && (1)\\
&= \int_{\Omega} \left(u^i\left(\left(\mathbf{d}^S_e, \mathbf{a}^{-S}\right)(\omega)\right) - u^i\left(\mathbf{a}(\omega)\right)\right) d\mu && (2)\\
&= \int_{F^{meet}} \left(u^i\left(\left(\mathbf{d}^S_e, \mathbf{a}^{-S}\right)(\omega)\right) - u^i\left(\mathbf{a}(\omega)\right)\right) d\mu && (3)\\
&= \int_{F^{meet}} \left(u^i\left(\left(\mathbf{d}^S, \mathbf{a}^{-S}\right)(\omega)\right) - u^i\left(\mathbf{a}(\omega)\right)\right) d\mu && (4)\\
&= \sum_j \int_{F^i_j} \left(u^i\left(\left(\mathbf{d}^S, \mathbf{a}^{-S}\right)(\omega)\right) - u^i\left(\mathbf{a}(\omega)\right)\right) d\mu && (5)\\
&= \sum_j u^i_d(\omega^i_j) - u^i_f(\omega^i_j) > 0 && (6)
\end{aligned}
$$

Equation (2) is due to the equality $F^i_e(\omega) = \Omega$, (3) holds since $\mathbf{d}^S_e = \mathbf{a}^{-S}$ outside F^{meet}, (4) holds since $\mathbf{d}^S_e = \mathbf{d}^S$ in F^{meet}, (5) follows from $F^{meet} = \dot{\bigcup}_j F^i_j$, and the last inequality is implied by $F^{meet} \subseteq \left\{\omega \mid u^i_d(\omega) > u^i_f(\omega)\right\}$. **QED**

5 Coalition-Proof Correlated Equilibria

In Sect. 4 we have shown that an *ex-ante* strong correlated equilibrium is also appropriate to frameworks in which players can plan deviations at all stages. A natural question is whether a similar result holds for the notion of coalition-proof correlated equilibrium. [4] We show that the answer is negative, by presenting an example, adapted from Bloch and Dutta (2008), in which there is an *ex-ante* coalition-proof correlated equilibrium that is not a self-enforcing agreement in a framework in which communication is possible at all stages. Table 2 presents a two-player game and an *ex-ante* coalition-proof correlated equilibrium.

Table 2
A Two-Player Game and an *Ex-ante* Coalition-Proof Correlated Equilibrium

	b_1	b_2	b_3
a_1	6,6	-2,0	0,7
a_2	2,2	2,2	0,0
a_3	0,0	0,0	3,3

	b_1	b_2	b_3
a_1	1/2	0	0
a_2	1/4	1/4	0
a_3	0	0	0

[4] An *ex-ante* coalition-proof correlated equilibrium is a correlated strategy profile from which no coalition has a self-enforcing_and improving *ex-ante* deviation (Moreno and Wooders, 1996). For a coalition of a single player any deviation is self-enforcing. For a larger coalition, a deviation is self-enforcing if there is no further self-enforcing and improving *ex-ante* deviation by one of its proper sub-coalitions.

We first show that the profile presented in Table 2 is an *ex-ante* coalition-proof equilibrium. Observe that it is a correlated equilibrium. Moreno and Wooders (1996) show that in a two-player game, every correlated equilibrium that is not Pareto-dominated by another correlated equilibrium is a coalition-proof correlated equilibrium. The profile gives each player a payoff of 4. Thus we prove that it is an *ex-ante* coalition-proof correlated equilibrium by showing that any correlated equilibrium q gives player 1 a payoff of at most 4. Let $x = q(a_1, b_1)$. Observe that $q(a_2, b_1) \geq x/2$ because otherwise player 2 would have a profitable deviation: playing b_3 when recommended b_1. This implies $q(a_2, b_2) \geq x/2$, because otherwise player 1 would have a profitable deviation: playing a_1 when recommended a_2. Thus q's payoff conditioned on that the recommendation profile is in $A = ((a_1, b_1), (a_2, b_1), (a_2, b_2))$ is at most 4, and the fact that the payoff of player 1 outside A is at most 3 completes the proof.

We now explain why this profile is not a self-enforcing agreement in a framework in which the players can also plan deviations at the *ex-post* stage. Assume that the players have agreed to play the profile, and player 1 has received a recommendation to play a_2. In that case, he can communicate with player 2 at the *ex-post* stage, tell him that he has received a_2 (and thus if the players follow the recommendation profile they would get a payoff of 2), and suggest a joint deviation: playing (a_3, b_3). As player 1 has no incentive to lie, player 2 would believe him, and they would both play (a_3, b_3). This *ex-post* deviation is self-enforcing because (a_3, b_3) is a Nash equilibrium.

Observe that the same deviation is not self-enforcing at the *ex-ante* stage. If the players agree at the *ex-ante* stage to implement a deviation that changes (a_2, b_1) into (a_3, b_3), then player 2 would have a profitable sub-deviation: playing b_3 when recommended b_1. Similarly, if they agree to implement a deviation that changes (a_2, b_2) into (a_3, b_3), then player 1 would have a profitable sub-deviation: playing a_1 when recommended a_2.

6 Discussion

Notions of *ex-ante* strong correlated equilibria have been presented in Moreno and Wooders (1996), Ray (1996), and Milgrom and Roberts (1996). Our *ex-ante* definition is equivalent to the definition of Moreno and Wooders. In Ray (1996) deviating coalitions are not allowed to construct new correlation devices, and are limited to use only uncorrelated deviations. In Milgrom and Roberts (1996) only some of the coalitions can coordinate deviations. In both cases the sets of feasible deviations are included in our set of deviations, and thus our set of *ex-ante* strong correlated equilibria is included in the other sets of equilibria.

An *ex-post* strong correlated equilibrium can be defined in our framework, as a profile that is resistant to deviations at the *ex-post* stage when each player knows his recommendation (i.e., no coalition $S \subseteq N$ has a profitable deviation with respect to an *ex-post* information structure $(\mathcal{F}^i)_{i \in S}$, in which: $\forall \omega \in \Omega$, $\forall i \in S$, $\exists a^i \in A^i$ s.t. $\mathbf{a}^i\left(F^i(\omega)\right)(a^i) = 1$).

Notions of *ex-post* strong correlated equilibria have been presented in Einy and Peleg (1995), Ray (1998), and Bloch and Dutta (2008). In Einy and Peleg (1995) a deviating coalition can only use deviations that improve the conditional utilities of all deviating players for *all possible* recommendation profiles.[5] In Ray (1998) a coalition S can only use *pure* deviations (functions $d^S : A^S \to A^S$). In Bloch and Dutta (2008), a coalition S can only use deviations that are implemented if the S-part of the recommendation profile a^S is included in some set $E^S \subseteq A^S$, which satisfies: (1) If $a^S \in E^S$, each player earns from implementing the deviation; (2) If $a^S \notin E^S$, at least one player looses from implementing the deviation. It can be shown that those conditions imply the existence of a profitable deviation with respect to an *ex-post* information structure.[6] Thus our set of *ex-post* strong correlated equilibria is included in the other sets of equilibria.

The main result reveals inclusion relations among the different notions of strong correlated equilibria, which described in Fig. 1.[7] Thus, the *ex-ante* notion of Moreno and Wooders is much more robust than originally presented: It is an appropriate notion not only for frameworks where players can only communicate before receiving the agreement's recommendations, but for any pre-play signaling process that is used to implement the agreement, and for any communication possibilities among the players.

Three possible extensions of the Main Result are:

(1) *Bayesian games*: Moreno and Wooders (1996) present a notion of *ex-ante strong communication equilibrium* in Bayesian games. The main result can be extended to this setup as well, to show that an *ex-ante* strong communication equilibrium is resistant to deviations at all stages.

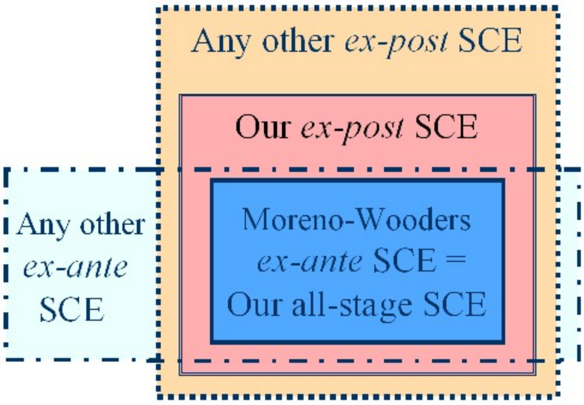
(2) *all-stage Coalition-proofness*: By using an appropriate notion of consistent refinements of information structures one can extend our model, and define a notion of *all-stage coalition-proof correlated equilibrium,* . However, the example in Sect. 5 shows that this notion does not coincide with the *ex-ante* coalition-proof notion, nor that there is any inclusion

[5] It is equivalent to requiring that $\forall i \in S, \omega \in \Omega$ $u^i_d(\omega) > u^i_f(\omega)$.

[6] The information structure is such that each deviator would know his recommendation and whether $\mathbf{a}^S(\omega) \in E^S$.

[7] See Moreno and Wooders (1996, Sect. 4) for an example of an *ex-post* strong correlated equilibrium that is not an *ex-ante* equilibrium.

Figure 1. Relations among Different Notions of Strong Correlated Equilibria (SCE)



relations among the different coalition-proof notions.[8]

(3) *k-strong equilibria*: In Heller (2008) an *ex-ante* notion of *k-strong correlated equilibrium* is defined as a strategy profile that is resistant to all coalitional deviations of up to k players. The main result can be directly extended to this notion: An *ex-ante* k-strong correlated equilibrium is resistant to deviations of up to k players at all stages.

References

[1] Aumann R., 1959. Acceptable Points in General cooperative n-person Games. In: Kuhn HW, Luce RD (eds) Contributions to the Theory of Games IV. Princeton University Press, NJ, 287-324.

[2] Aumann R., 1976. Agreeing to disagree. Ann. Statist. 4 (6), 1236-1239.

[3] Aumann R., 1987. Correlated equilibrium as an expression of Bayesian rationality. Econometrica 55, 1-18.

[4] Bernheim B.D., Peleg B., Whinston M., 1987. Coalition-proof Nash equilibria - I. concepts. J. Econ. Theory 42, 1-12.

[5] Bloch F., Dutta B., 2008. Correlated equilibria, incomplete information and coalitional deviations. Games Econ. Behav. (forthcoming).

[6] Einy E., Peleg B., 1995. Coalition-proof communication equilibria. In: Barnett W, Moulin H, Salles M, Schofield N (eds), Social Choice, Welfare & Ethics, Cambridge, New-York and Melbourne: Cambridge University Press.

[8] The example presents an *ex-ante* equilibrium that is not an all-stage equilibrium. Based on this, it is possible to construct a 3-player game with an all-stage equilibrium that is not an *ex-ante* equilibrium, in which a deviation of the coaltion $\{1, 2\}$ induces a situation similar to table 2.

[7] Heller Y., 2008. Minority-proof cheap-talk protocol (extended version), mimeo, http://www.tau.ac.il/~helleryu/minority-extended.pdf

[8] Holmstrom B., Myerson RB, 1983. Efficient and Durable Decision Rules with Incomplete Information, Econometrica 51, 1799-1819.

[9] Milgrom P., Roberts J., 1996. Coalition-proofness and correlation with arbitrary communication possibilities. Games Econ. Behav. 17, 113-128.

[10] Moreno D., Wooders J., 1996. Coalition-proof equilibrium. Games Econ. Behav. 17, 80-113.

[11] Ray I., 1996. Coalition-proof correlated equilibrium: a definition. Games Econ. Behav. 17, 56-79.

[12] Ray I., 1998. Correlated equilibrium as a stable standard of behavior. Rev. Econ. Design **3**, 257-269