

CORE DISCUSSION PAPER

2006/41

CORRELATED EQUILIBRIUM IN GAMES WITH INCOMPLETE
INFORMATION REVISITED*

Françoise Forges[†]

April 2006

Abstract

A mistake in "Five legitimate definitions of correlated equilibrium in games with incomplete information" motivates a re-examination of some extensions of the solution concept that Aumann introduced.

Keywords: correlated equilibrium, Bayesian rationality, games with incomplete information

*"Five legitimate definitions of correlated equilibrium in games with incomplete information" first appeared as CORE Discussion Paper 9309. I prepared this note during a short visit at CORE in February 2006. I thank Qing Min Liu, Jean-François Mertens and Peter Vida for their comments.

[†]CEREMADE, Université Paris-Dauphine, Place du Maréchal de Lattre de Tassigny, 75775 Paris cedex 16.

1 Introduction

Aumann (1974) introduced the *correlated equilibrium* (C.E.) and showed that it could be interpreted as an expression of Bayesian rationality in games with complete information (see Aumann (1987)). In Forges (1993), henceforth [5def], I surveyed three standard extensions of Aumann’s solution concept in games with incomplete information (the *strategic form* C.E, the *agent normal form* C.E. and the *communication equilibrium*) and proposed a fourth extension, the *Bayesian solution*, in order to capture Bayesian rationality in a similar way as in Aumann (1987)¹. More precisely, I showed that a natural formulation of Bayesian rationality *à la* Aumann in games with incomplete information led to the *belief invariant* Bayesian solution, in which, at the time of making decisions, the posterior probability distribution of every player over the others’ types is the same as the prior². It follows from the definitions that every strategic form C.E. induces an agent normal form C.E., which in turn induces a belief invariant Bayesian solution. I claimed that the latter was in fact *equivalent* to the agent normal form C.E. As Lehrer et al. (2006) show on a counter-example, this claim is wrong. Hence, as argued in [5def], the agent normal form C.E. can be viewed as an expression of Bayesian rationality, but this property characterizes a larger set of solutions. The non-equivalence of the belief invariant Bayesian solution and the agent normal form C.E., as well as recent research on related topics, calls for a careful re-interpretation of both solution concepts. This is the purpose of this note. I first recall the definitions of C.E. in games with incomplete information and Lehrer et al. (2006)’s counter-example. Then I show that many differences between the extensions of the C.E. disappear in a standard cheap talk extension of the game, which does not affect the communication equilibrium outcomes nor the Bayesian solutions. Roughly, only the latter two solution concepts need to be considered if the number of players is at least three³. In particular, the equivalence claimed in [5def] can be recovered in that framework.

¹[5def] contains a fifth definition, in the framework of Mertens and Zamir (1985)’s universal beliefs space, but I will not refer to it here.

²In [5def], I used the less precise terminology “conditional independence property” instead of “belief invariance”. Nau (1992) considered general Bayesian solutions, which do not satisfy this property.

³In the more general cheap talk extension constructed by Vida (2006), the same holds with two players.

2 Model and solution concepts

The notations are as in [5def], but in order to fully clarify the differences between the various definitions, games involve an arbitrary number of players instead of two.

Model

Following Harsanyi (1967), we focus on a Bayesian game G , described by a finite set of players N , finite sets of types T_i ($i \in N$), a common probability distribution P over $T = \prod_{i \in N} T_i$, finite sets of actions A_i and payoff functions $v_i : T \times A \rightarrow \mathbb{R}$ ($i \in N$), where $A = \prod_{i \in N} A_i$. G is implicitly the reduced form of an interactive decision problem defined on a set of basic parameters K , so that the types of the players summarize their hierarchies of beliefs over K .

Without loss of generality, the definitions below are given in “canonical form” (see, e.g., Forges (1986b)).

Strategic form correlated equilibrium

G can be viewed as a game in strategic form with sets of pure strategies $\Sigma_i = A_i^{T_i}$ ($i \in N$) and payoffs computed as expectations w.r.t. P . A C.E. of this game, in the sense of Aumann (1974), defines a *strategic form* C.E. of G . A strategic form C.E. is thus implemented by means of

(i) a correlation device, namely a probability distribution Q over $\prod_{i \in N} \Sigma_i$, which selects an N -tuple of pure strategies $(\sigma_i)_{i \in N}$

(ii) a mediator, who privately recommends σ_i to player i ($i \in N$) such that the players cannot gain in deviating unilaterally from the recommendations.

Let $C(G)$ be the set of strategic form C.E. (*interim* expected) payoffs of G .

The prior probability distribution P and a strategic form C.E. distribution Q induce the following probability distribution $\Pi_{P,Q}$ over $T \times A$

$$\Pi_{P,Q}(t, a) = P(t)Q(\sigma_i(t_i) = a_i, i \in N)$$

where $t = (t_i)_{i \in N}$ and $a = (a_i)_{i \in N}$. $\Pi_{P,Q}$ satisfies a conditional independence property (C.I.P.): *Player i 's action (a_i) is conditionally independent of the other players' types ($t_{-i} = (t_j)_{j \neq i}$), given his type (t_i).* In other words, a strategic form C.E. is *belief invariant*: the players' posterior beliefs $\Pi_{P,Q}(t_{-i}|t_i, a_i)$ coincide with their priors $P(t_{-i}|t_i)$.

Agent normal form correlated equilibrium

In the agent normal form of G , every type t_i of player i is viewed as a different player ($t_i \in T_i$, $i \in N$). An agent normal form C.E. is implemented by means of

(i) a correlation device, namely a probability distribution Q over $\prod_{i \in N} \Sigma_i$, as above. Q selects $(\sigma_i)_{i \in N} = ([\sigma_i(t_i)]_{t_i \in T_i})_{i \in N}$

(ii) a mediator, who privately recommends $\sigma_i(t_i)$ to agent (i, t_i) of player i ($t_i \in T_i$, $i \in N$)

such that the (agents of the) players cannot gain in deviating unilaterally from the recommendations.

Let $C_a(G)$ be the set of agent normal form C.E. (*interim* expected) payoffs of G .

An agent normal form C.E. distribution Q is the same object as a strategic form C.E. distribution, but satisfies weaker non-deviation conditions. In particular, the probability distribution it induces, together with the prior P , over $T \times A$, satisfies the C.I.P., so that a normal form C.E. is belief invariant.

Communication equilibrium

A communication equilibrium is implemented by means of

(i) a communication device, namely a system q of probability distributions $q(\cdot|t)$ over A , $t \in T$

(ii) a mediator, who invites every player i to report his type t_i , selects an N -tuple of actions a according to $q(\cdot|t)$ and privately recommends a_i to player i ($i \in N$)

such that the players cannot gain in unilaterally lying on their type nor deviating from the recommended action.

The probability distribution $\Pi_{P,q}$ induced by the prior P and q over $T \times A$, namely,

$$\Pi_{P,q}(t, a) = P(t)q(a|t) \tag{1}$$

does not necessarily satisfy the C.I.P. Let $M(G)$ be the set of (*interim* expected) communication equilibrium payoffs of G .

Bayesian solution

As in Aumann (1987), let Y be the (finite) set of all states of the world, \mathcal{S}_i be player i 's information partition of Y and Π be the common prior probability of the players over Y . Y contains in particular the players' types and actions in G . Player i 's type can thus be viewed as a random variable $\tau_i : Y \rightarrow T_i$; similarly, player i 's action is a random variable $\alpha_i : Y \rightarrow A_i$.

A first natural requirement is that these random variables be \mathcal{S}_i -measurable, namely that every player knows his type and his action. Consistency conditions express that the beliefs of the players in the enlarged model are the same as in G :

$$\begin{aligned} \Pi(\tau = t) &= P(t) \\ \Pi(\tau_{-i} = t_{-i} | \mathcal{S}_i) &= \Pi(\tau_{-i} = t_{-i} | \tau_i) \quad i \in N \end{aligned} \tag{2}$$

Finally, every player must be Bayesian rational, i.e., player i 's action maximizes his expected payoff given his information:

$$E_{\Pi}(v_i(\tau, \alpha) | \mathcal{S}_i) = \max_{a_i \in A_i} E_{\Pi}(v_i(\tau, a_i, \alpha_{-i}) | \mathcal{S}_i) \quad i \in N$$

Let $B_I(G)$ (resp., $B(G)$) be the set of all (*interim* expected) payoffs that can be achieved in a world satisfying all the previous conditions (resp., all conditions but (2)). The payoffs in $B(G)$ are canonically achieved by means of

(i) a system q of probability distributions $q(\cdot | t)$ over A , $t \in T$ (formally, as a communication device)

(ii) an *omniscient mediator*, who knows the N -tuple of realized types \hat{t} , selects an N -tuple of actions a according to $q(\cdot | \hat{t})$ and privately recommends a_i to player i ($i \in N$)

such that the players cannot gain in unilaterally deviating from the recommended action.

For $B_I(G)$, one further requires belief invariance, namely that the probability distribution induced by P and q over $T \times A$, as in (1), satisfies the C.I.P. In other words, the omniscient mediator can use his knowledge of the types to make his recommendations but the players should not be able to infer anything on the others' types from these recommendations. Following Lehrer et al. (2006)'s terminology, the "garbling" q , which transforms the initial information structure, in which every player i is informed of t_i , into the information structure in which every player learns his action a_i , is "non-communicating".

$B(G)$ (resp., $B_I(G)$) will be referred to as the set of Bayesian solutions (resp., belief invariant Bayesian solutions). Nau (1992) justifies $B(G)$ in terms of no arbitrage conditions. According to a completely different approach, $B(G)$ is also the set of all "certification equilibrium" payoffs (see Forges and Koessler (2005)). Several recent papers (e.g., Dekel et al. (2005),

Ely and Peski (2006) and Liu (2005)) emphasize the importance of solutions that preserve the players' hierarchies of beliefs, which in our reduced framework, amounts to belief invariance⁴.

3 Comparison of the solution concepts

3.1 Comparison in the original game

In a strategic form C.E., the lottery Q can be performed before the move of nature P and the correlation device is autonomous, in the sense that it does not require any input from the players. The mediator just needs to be able to identify every player i , $i \in N$.

In an agent normal form C.E., the lottery Q is the same as in a strategic form C.E. and the correlation device is again autonomous. However the mediator needs to be able to identify the agents of every player i . In practice, this means that the players' types must be verifiable by the mediator. However, an agent normal form C.E. can be implemented by $|N|$ *different mediators, one for every player*, who must only be able to verify the type of the player they are taking care off. Player i 's mediator observes $\sigma_i = [\sigma_i(t_i)]_{t_i \in T_i}$, verifies player i 's realized type \hat{t}_i and recommends him $\sigma_i(\hat{t}_i)$. The implementation of an agent normal form C.E. does not rely on an omniscient mediator who would know the N -tuple of realized types.

The communication equilibrium deeply differs from the previous two solution concepts in that it allows for information transmission from the players to the mediator. In particular, it is not necessarily belief invariant.

In a Bayesian solution, the probability distribution over actions given types is the same as if there were an omniscient mediator making recommendations to the players. If the solution is belief invariant, these recommendations do not modify the players beliefs over types.

There is no inclusion relationship between $C_a(G)$ and $M(G)$ (see [5def] for examples). But, as a consequence of the definitions,

$$\begin{aligned} C(G) &\subseteq C_a(G) \subseteq B_I(G) \subseteq B(G) \\ C(G) &\subseteq M(G) \subseteq B(G) \end{aligned}$$

⁴Dekel et al. (2005), Ely and Peski (2006) and Liu (2005) start with a basic parameter space and consider the full belief hierarchies of the players without imposing a common prior.

All these inclusions may be strict. In particular, proposition 3 of [5def], according to which $C_a(G) = B_I(G)$ in every two-person game G , is wrong as soon as both players have private information. This is illustrated on the following counter-example from Lehrer et al. (2006).

Example 1

$N = \{1, 2\}$, $T_i = \{t_i, t'_i\}$, types are independent and equiprobable, $A_i = \{a_i, a'_i\}$, $i = 1, 2$. If the pair of types is (t_1, t_2) , the payoffs are

$$\begin{array}{cc} & \begin{array}{c} a_2 \\ a'_2 \end{array} \\ \begin{array}{c} a_1 \\ a'_1 \end{array} & \begin{pmatrix} (1, 1) & (-1, -1) \\ (-1, -1) & (1, 1) \end{pmatrix} \end{array}$$

For the other three pairs of types, the payoffs are

$$\begin{array}{cc} & \begin{array}{c} a_2 \\ a'_2 \end{array} \\ \begin{array}{c} a_1 \\ a'_1 \end{array} & \begin{pmatrix} (-1, -1) & (1, 1) \\ (1, 1) & (-1, -1) \end{pmatrix} \end{array}$$

The best *ex ante* expected payoff that every player can achieve with a pair of pure strategies (in $\Sigma_1 \times \Sigma_2$), and thus at an agent normal form C.E., is $\frac{1}{2}$. However, if an omniscient mediator makes recommendations according to

$$\begin{array}{cc} & \begin{array}{c} a_2 \\ a'_2 \end{array} \\ \begin{array}{c} a_1 \\ a'_1 \end{array} & \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix} \end{array}$$

if the pair of types is (t_1, t_2) and

$$\begin{array}{cc} & \begin{array}{c} a_2 \\ a'_2 \end{array} \\ \begin{array}{c} a_1 \\ a'_1 \end{array} & \begin{pmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{pmatrix} \end{array}$$

otherwise, the *ex ante* expected payoff of both players is 1. None of the players can gain in deviating unilaterally from the recommendation. Furthermore, the corresponding Bayesian solution is belief invariant.

Proposition 3 in [5def], which is stated for two-person games, holds if there is a single informed player. Even with this assumption, it does not extend if $|N| > 2$ (see example 3 in the appendix).

Proposition 1 *If $|N| = 2$ and only one player has private information, then $C_a(G) = B_I(G)$.*

Proof: Let player 1 be informed (i.e., T_2 is a singleton). Consider a solution in $B_I(G)$, and let $q(\cdot|t_1)$, $t_1 \in T_1$, be the corresponding system of probability distributions over A . For every $t_1 \in T_1$, $a_1 \in A_1$, $a_2 \in A_2$,

$$q(a_1, a_2|t_1) = q(a_1|t_1, a_2)q(a_2|t_1)$$

By belief invariance,

$$q(a_1, a_2|t_1) = q(a_1|t_1, a_2)q(a_2)$$

Hence one can define Q over $\Sigma_1 \times A_2$ by

$$Q([\sigma_1(t_1)]_{t_1 \in T_1}, a_2) = q(a_2) \prod_{t_1 \in T_1} q(\sigma_1(t_1)|t_1, a_2)$$

If Q is used to recommend a_2 to player 2 and $\sigma_1(t_1)$ to player 1 of type t_1 , both players have the same information as in the initial Bayesian solution; hence, Q defines an agent form correlated equilibrium.

3.2 Comparison in a cheap talk extension of the game

Let us pursue our comparison of the different definitions of C.E. by considering a *cheap talk extension* of G , denoted as $extG$. In $extG$, the players are informed of their types as in G ; there are two stages of cheap talk: at every stage $t = 1, 2$, the players simultaneously send a message to each other. To keep the approach as simple as possible, we assume that there is a continuum of messages. By the arguments that are used to derive canonical representations, one can show that cheap talk has no effect on communication equilibria nor on Bayesian solutions: $M(extG) = M(G)$ (see, e.g., Forges (1990)) and $B(extG) = B(G)$ (see the appendix). Intuitively, in both solution concepts, the mediator is powerful enough to mimic a plain conversation between the players. Forges (1990) proves that, if $|N| \geq 3$, $C(extG) = M(G)$.⁵ In other words, as soon as cheap talk is allowed, the strategic form C.E. and the communication equilibrium are equivalent. We will establish a similar property for the agent normal form C.E. and the Bayesian solution.

⁵The result also holds for $|N| = 2$ if further assumptions are imposed on G or if more general cheap talk extensions are allowed (see Vida (2006)). Versions are also available for Nash, or sequential equilibria, of $extG$ (see, e.g., Ben Porath (2003), Gerardi (2004), Forges (1990)).

Let us say that $x_i = [x_i(t_i)]_{t_i \in T_i}$ is *interim individually rational* for player i if there exists a system of probability distributions $\mu_{-i}(\cdot|t_{-i})$ over A_{-i} such that

$$x_i(t_i) \geq \sum_{t_{-i}} P(t_{-i}|t_i) \sum_{a_{-i}} \mu_{-i}(a_{-i}|t_{-i}) v_i(t_i, t_{-i}, a_i, a_{-i}) \quad \forall t_i, a_i \quad (3)$$

This definition makes sense in extensions G' of the game G such that, if player i unilaterally decides not to participate in G' , the other players can punish player i , whatever his type t_i , without knowing t_i .

Let $INTIR(G)$ be the set of payoffs, in $\prod_{i \in N} R^{|T_i|}$, that are *interim individually rational* for every player. It is not difficult to show that $C_a(G) \cup M(G) \subseteq INTIR(G)$ and that $B(G)$ is not necessarily included in $INTIR(G)$ (see the appendix).

Proposition 2 *If $|N| \geq 3$, $C_a(extG) = B(G) \cap INTIR(G) = B(extG) \cap INTIR(G)$*

The result can be proved following the lines of Forges (1990) (see the appendix). It also holds for $|N| = 2$ under appropriate assumptions (see Forges (1986a)) or by relying on more complex cheap talk protocols, as in Vida (2006). An immediate consequence of the previous proposition is that *agent normal form correlated equilibria and belief invariant Bayesian solutions are essentially equivalent if $|N| \geq 3$ and cheap talk is allowed: $C_a(extG) = B_I(extG) \cap INTIR(G)$* . However, the scope of this corollary is limited by the fact that, as stated in proposition 2, cheap talk destroys belief invariance.

The lesson from the exercise is rather that, if players are at least three and can talk to each other before making their decisions, only two extensions of Aumann's solution concept survive: the communication equilibrium and the Bayesian solution. Nau (1992) ends up with exactly the same solution concepts by a completely different approach. As suggested above, solutions relying on an *omniscient mediator* do not make much sense in reduced games as G , described by fixed sets of types, but can only result from the identification of an appropriate enlarged game. This in turn requires the explicit description of the underlying parameter space and the players' hierarchies of beliefs.

4 Appendix

Proof of $B(\text{ext}G) \subseteq B(G)$

In $\text{ext}G$, a pure strategy of player i consists of mappings $(\sigma_i^1, \sigma_i^2, \tau_i)$ where $\sigma_i^1 : T_i \rightarrow [0, 1]$ (resp., $\sigma_i^2 : T_i \times [0, 1]^N \rightarrow [0, 1]$) chooses player i 's message at the first (resp., second) stage of cheap talk and $\tau_i : T_i \times [0, 1]^N \times [0, 1]^N \rightarrow A_i$ chooses player i 's action. Consider an omniscient mediator, who knows the N -tuple of realized types \hat{t} , selects strategies $(\sigma_i^1, \sigma_i^2, \tau_i)_{i \in N}$ as a function of \hat{t} , and recommends $(\sigma_i^1, \sigma_i^2, \tau_i)$ to player i , $i \in N$, so that none of the players can gain in deviating unilaterally from the recommendation. Given \hat{t} and $(\sigma_i^1, \sigma_i^2, \tau_i)_{i \in N}$, this mediator can evaluate the actions that the players would choose at the last stage of $\text{ext}G$. If he recommends directly these actions in G , none of the players can gain in deviating unilaterally from the recommendation. This induces a solution in $B(G)$.

Relationships between correlated equilibrium payoffs and *interim* individually rational ones

$C_a(G) \cup M(G) \subseteq \text{INTIR}(G)$. If $|N| = 2$, $B(G)$ is not necessarily included in $\text{INTIR}(G)$ (see example 2) but $B_I(G) \subseteq \text{INTIR}(G)$. However, example 3 will illustrate that $B_I(G)$ may not be included in $\text{INTIR}(G)$ if $|N| > 2$, even if only one player has private information. Hence, $B_I(G)$ may not be included in $C_a(G)$ in this case.

To show that $C_a(G) \subseteq \text{INTIR}(G)$ for every $|N|$ and that $B_I(G) \subseteq \text{INTIR}(G)$ if $|N| = 2$, we establish more generally that an appropriate subset of $B(G)$ is included in $\text{INTIR}(G)$. Consider the following property (C.I.P.): player i 's type t_i is conditionally independent of the other players' actions a_{-i} given the other players' types t_{-i} . If $|N| = 2$, this is just the C.I.P. above; but if $|N| > 2$, it is a quite different property. The probability distributions over $T \times A$ associated with $C_a(G)$ satisfy both. Fix a solution in $B(G)$ such that (C.I.P.) holds. Let $i \in N$; assume that player i , of type t_i , takes no account of the recommendation and plays some given action $a_i^* \in A_i$, while

the other players follow the recommendation. Player i 's expected payoff is

$$\begin{aligned}
& \sum_{t_{-i}} P(t_{-i}|t_i) \sum_{a_i, a_{-i}} q(a_i, a_{-i}|t_i, t_{-i}) v_i(t_i, t_{-i}, a_i^*, a_{-i}) \\
&= \sum_{t_{-i}} P(t_{-i}|t_i) \sum_{a_{-i}} q(a_{-i}|t_i, t_{-i}) v_i(t_i, t_{-i}, a_i^*, a_{-i}) \\
&= \sum_{t_{-i}} P(t_{-i}|t_i) \sum_{a_{-i}} q(a_{-i}|t_{-i}) v_i(t_i, t_{-i}, a_i^*, a_{-i})
\end{aligned}$$

where the last equality follows from C.I.P.'

$M(G) \subseteq INTIR(G)$: Let $i \in N$; assume that player i , of type t_i , reports that his type is t_i^* but takes no account of the recommendation and plays some given action $a_i^* \in A_i$. Player i 's expected payoff is

$$\begin{aligned}
& \sum_{t_{-i}} P(t_{-i}|t_i) \sum_{a_i, a_{-i}} q(a_i, a_{-i}|t_i^*, t_{-i}) v_i(t_i, t_{-i}, a_i^*, a_{-i}) \\
&= \sum_{t_{-i}} P(t_{-i}|t_i) \sum_{a_{-i}} q(a_{-i}|t_i^*, t_{-i}) v_i(t_i, t_{-i}, a_i^*, a_{-i})
\end{aligned}$$

Since the previous equality holds for every t_i and t_i^* , players $j \neq i$ can punish all types t_i of player i simultaneously by using $q(a_{-i}|t_i^*, t_{-i})$ for some given t_i^* .

Example 2

$N = \{1, 2\}$, $T_1 = \{t_1, t'_1\}$, types are equiprobable, $A_2 = \{a_2, a'_2\}$; all other sets are singletons. If player 1's type is t_1 , the payoffs are

$$\begin{array}{cc}
a_2 & a'_2 \\
(1, 0) & (0, 1)
\end{array}$$

while for t'_1 , they are

$$\begin{array}{cc}
a_2 & a'_2 \\
(0, 1) & (1, 0)
\end{array}$$

$((0, 0), 1) \in B(G)$, but $\notin INTIR(G)$ since *interim* individually rational payoffs for player 1 have the form $(\lambda, 1 - \lambda)$ for $0 \leq \lambda \leq 1$.

Example 3

$N = \{1, 2, 3\}$, $T_1 = \{t_1, t'_1\}$, types are equiprobable, $A_i = \{a_i, a'_i\}$, $i = 2, 3$; all other sets are singletons. If player 1's types is t_1 , the payoffs are

$$\begin{array}{cc} & a_3 & a'_3 \\ a_2 & (0, 1, 1) & (1, 0, 0) \\ a'_2 & (1, 0, 0) & (0, 1, 1) \end{array}$$

Otherwise, they are

$$\begin{array}{cc} & a_3 & a'_3 \\ a_2 & (1, 0, 0) & (0, 1, 1) \\ a'_2 & (0, 1, 1) & (1, 0, 0) \end{array}$$

Assume that an omniscient mediator makes recommendations according to

$$\begin{array}{cc} & a_3 & a'_3 \\ a_2 & \frac{1}{2} & 0 \\ a'_2 & 0 & \frac{1}{2} \end{array}$$

if player 1's type is t_1 and

$$\begin{array}{cc} & a_3 & a'_3 \\ a_2 & 0 & \frac{1}{2} \\ a'_2 & \frac{1}{2} & 0 \end{array}$$

otherwise. This defines a solution in $B_I(G)$, with payoff $(0, 0)$ to player 1, while as in the previous example, *interim* individually rational payoffs for this player have the form $(\lambda, 1 - \lambda)$ for $0 \leq \lambda \leq 1$. In particular, the solution is not in $C_a(G)$.

Proof of proposition 2

By the definition of the solution concepts, $C_a(extG) \subseteq B(extG)$. We have shown above that $B(extG) \subseteq B(G)$, so that $C_a(extG) \subseteq B(G)$.

To see that $C_a(extG) \subseteq INTIR(G)$, fix a payoff in $C_a(extG)$. It is achieved through a probability distribution over the pure strategies in $extG$, which are described as above. Consider player i of type t_i and let $(\sigma_i^1(t_i), \sigma_i^2(t_i, \cdot), \tau_i(t_i, \cdot, \cdot))$ be his recommendation. Assume that player i ignores his type, this recommendation as well as the others' messages and chooses $(m_i^1, m_i^2, a_i) \in [0, 1] \times [0, 1] \times A_i$, while the other players follow their recommendations. The interim expected payoff of player i corresponding to this deviation does not depend on t_i (nor on his recommendation) and yields thus a level of the form in (3).

The proof that $B(G) \cap INTIR(G) \subseteq C_a(extG)$ is just a modification of the proof that $M(G) \subseteq C(extG)$ in Forges (1990) (sections 4.1 and 4.2)⁶. Let us fix a payoff in $B(G) \cap INTIR(G)$; let $q = q(\cdot|t)$, $t \in T$, be the corresponding system of conditional probability distributions over A . q has the same form as a communication device, but satisfies weaker non-deviation conditions than a communication equilibrium. Since the payoff is in $INTIR(G)$, there are also probability distributions $\mu_{-i}(\cdot|t_{-i})$ over A_{-i} for every i , t_{-i} such that the inequalities (3) hold. The proof of Forges (1990) constructs a correlation device (henceforth, c.d.) for $extG$ with q as system of conditionals. If q were a communication equilibrium, this c.d. would define a strategic form correlated equilibrium of $extG$. We will show that if q corresponds to a Bayesian solution, a slight modification of this c.d. induces an agent normal form of $extG$.

As in Forges (1990), let the c.d. select uniformly, independently of each other, $|N|$ random bijections $\gamma_i : T_i \rightarrow L_i$, where L_i is a copy of T_i , $i \in N$. In addition, let the c.d. choose code functions (as in Forges (1990), section 4.2) $k_i : L_i \rightarrow [0, 1]$ uniformly and independently of each other. Before the beginning of $extG$, the c.d. transmits to player i of type \hat{t}_i : $\gamma_i(\hat{t}_i)$, $k_i(\gamma_i(\hat{t}_i))$ and k_j , namely $k_j(l_j)$, $l_j \in L_j$, for every $j \neq i$. The c.d. also selects $a_{-j}(t_{-j})$ for every t_{-j} according to $\mu_{-j}(\cdot|t_{-j})$; all these choices are made independently of each other and transmitted to all players $i \neq j$. Finally, the c.d. picks mappings σ_i and τ_i , $i \in N$, that will be useful at the second and final stage of $extG$. This part is exactly as in Forges (1990) and will not be detailed here.

The strategy of player i of type \hat{t}_i prescribes to send $\gamma_i(\hat{t}_i)$ and $k_i(\gamma_i(\hat{t}_i))$ to the other players at the first stage of $extG$. If player i attempts to send $l_i \neq \gamma_i(\hat{t}_i)$, he must send the code $k_i(l_i)$ which is expected by the other players. But he has a zero probability to guess it correctly. If player i sends a wrong code, the other players punish him by revealing t_{-i} to each other at the second stage of $extG$ and by choosing $a_{-i}(t_{-i})$ at the last stage of $extG$ (as if they chose their actions in A_{-i} according to $\mu_{-i}(\cdot|t_{-i})$). The rest of the proof is exactly the same as in Forges (1990).

⁶Two key aspects of the reasoning already appear in [5def]: the first basic idea is to use a product distribution to induce the same conditionals as with a mediator (as in Forges (1990)); another useful feature is that in an agent normal form C.E., there are no incentive conditions w.r.t. the revelation of types. However, these ideas do not lead anywhere unless applied to $ext(G)$. The problem is that, in $extG$, cheap talk kills the C.I.P. of agent normal form correlated equilibria: the reasoning shows that $B_I(extG) \subseteq C_a(extG)$, but $B_I(extG) = B(extG) = B(G)$.

References

- [1] Aumann, R., 1974, "Subjectivity and correlation in randomized strategies", *Journal of Mathematical Economics* 1, 67-96.
- [2] Aumann, R., 1987, "Correlated equilibrium as an expression of Bayesian rationality", *Econometrica* 55, 1-18.
- [3] Ben Porath, E., 2003, "Cheap talk in games with incomplete information", *Journal of Economic Theory* 108, 45-71.
- [4] Dekel, E., D. Fudenberg and S. Morris, 2005, "Interim rationalizability", mimeo, Harvard University.
- [5] Ely and Peski, 2006, "Hierarchies of beliefs and interim rationalizability", *Theoretical Economics* 1, 19-65.
- [6] Lehrer, E., D. Rosenberg and E. Shmaya, 2006, "Signaling and mediation in games with common interests", mimeo, Tel Aviv University
- [7] Forges, F., 1986a, "Correlated equilibria in games with incomplete information: a model with verifiable types", *International Journal of Game Theory* 15, 65-82.
- [8] Forges, F., 1986b, "An approach to communication equilibria", *Econometrica* 54, 1375-1385.
- [9] Forges, F., 1990, "Universal mechanisms", *Econometrica* 58, 1341-1364.
- [5def] Forges, F., 1993, "Five legitimate definitions of correlated equilibrium in games with incomplete information", *Theory and Decision* 35, 277-310.
- [11] Forges, F. and F. Koessler, 2005, "Communication equilibria with partially verifiable types", *Journal of Mathematical Economics* 41, 793-811.
- [12] Gerardi, D., 2004, Unmediated communication in games with complete and incomplete information, *Journal of Economic Theory* 114, 104-131.

- [13] Harsanyi, J., 1967, "Games with incomplete information played by Bayesian players", *Management Science* 14, 159-182.
- [14] Liu, Qing Min, 2005, "Representation of beliefs hierarchies in games with incomplete information", mimeo, Graduate School of Business, Stanford University.
- [15] Mertens, J.-F. and S. Zamir, 1985, "Formulation of Bayesian analysis for games with incomplete information", *International Journal of Game Theory* 14, 1-29.
- [16] Nau, R., 1992, "Joint coherence in games with incomplete information", *Management Science* 38, 374-387.
- [17] Vida, P., 2006, "A detail-free mediator and the three player case", mimeo, PSE (Paris-Jourdan Sciences Economiques).