Implementability of Correlated and Communication Equilibrium Outcomes in Incomplete Information Games

Igal Milchtaich*

May 2013

Abstract
In a correlated equilibrium, the players' choice of actions is directed by correlated random messages received from an outside source, or mechanism. These messages allow for more equilibrium outcomes than without any messages (pure-strategy equilibrium) or with statistically independent ones (mixed-strategy equilibrium). In an incomplete information game, the messages may also reflect the types of the players, either because they are affected by extraneous factors that also affect the types (correlated equilibrium) or because the players themselves report their types to the mechanism (communication equilibrium). Mechanisms may be further differentiated by the connections between the messages that the players receive and their own and the other players' types, by whether the messages are statistically dependent or independent, and by whether they are random or deterministic. Consequently, whereas for complete information games there are only three classes of equilibrium outcomes, with incomplete information the corresponding number is 14 or 15 for correlated equilibria and even larger – 15, 16 or 17 – for communication equilibria. For both solution concepts, the implication relations between the different kinds of equilibria form a two-dimensional lattice, which is considerably more intricate than the single-dimensional one of the complete information case.
JEL Classification: C72.

Keywords: Correlated equilibrium; Communication equilibrium; Incomplete information; Bayesian games; Mechanism; Implementation

1 Introduction

In complete information games, pure-strategy equilibria represent only a subset of the outcomes achievable by mixed-strategy equilibria, which in turn allow for fewer equilibrium outcomes than correlated equilibria. This is so whether 'outcome' is interpreted as referring to the distribution of the players' action profile or to their expected payoffs. Implementing the three solution concepts requires different, increasingly more general, mechanisms. Pure-strategy equilibria can do without any mechanism at all, but mixed-strategy equilibria require independent randomizations and correlated equilibria require centralized randomization. A similar connection between the allowed kinds of mechanisms and the implementable outcomes holds for incomplete information, or Bayesian, games, but it is considerably more complex than in the complete information case. This is because the set of equilibrium outcomes implementable by a mechanism may depend on the manner and the extent to which its output reflects the players' types. It may matter, for example, whether the messages that the players receive from the mechanism provide them with information about the other players' types, and whether they depend on the receiver's own type. The former may affect the mechanism's ability to implement type-dependent coordinated actions, and the latter reflects its ability to transmit information selectively, that is, to certain types of players only.

Dependence of the message received by a player on his own type does not logically require a type-aware mechanism. Type-dependent perceptual abilities may suffice. For example, the information about the value of an auctioned asset that a firm is able to extract from publicly presented data may depend on its level of experience or financial resources. By contrast, the

* Department of Economics, Bar-Ilan University, Ramat Gan 5290002, Israel
igal.milchtaich@biu.ac.il http://faculty.biu.ac.il/~milchti

received message may provide information about the *other* players' types only if the mechanism "knows" something about them. This may be so if the message is affected by factors that also affect the types. For example, firms entering an auction may factor in certain macroeconomic indicators when deciding on their bids, because, among other things, the state of the economy as a whole may bear on their competitors' situation. Similarly, the asking price of a farmer, hunter or trawler dealing with a wholesaler may be influenced by the recent weather or sea condition, because the latter is correlated with the market supply.

An alternative, straightforward way to construct a type-aware mechanism is through self-reporting of types. The reporting may be explicit or implicit. The former may hold, for example, for a mechanism that takes the form of an impartial mediator, and the latter holds for direct exchange of messages between players.

The two different kinds of type-aware mechanism correspond to two distinct solution concepts for incomplete information games: correlated equilibrium and communication equilibrium. The former employs one-way communication only – messages from the mechanism to the players – whereas the latter also involves (explicit or implicit) type reports from the players. In both cases, the messages that the mechanism sends to the players are not necessarily concrete action recommendations. The translation into actions is described by the players' strategies, which, for each player type, associate an action with each possible message. In a correlated equilibrium, taking the associated action is required to be incentive compatible. Communication equilibrium adds the requirement that truthful type reports are incentive compatible: players never have an incentive to misrepresent their types.

Correlated and communication equilibria, as defined here, share a single notion of correlated strategy, which is a pair consisting of a mechanism and a strategy profile. Where they differ is in the interpretation of the type profile on which the mechanism's messages are based: true vs. reported types.[1] Other solutions concepts, such as Bayesian equilibrium, may be viewed as special cases, obtained by limiting in one or more ways the mechanism's capabilities. Depending on the context, the potency of these limitations may stem from their influence on the mechanisms' ability to orchestrate certain joint actions, to make the actions incentive incompatible, or to elicit truthful type reports. The main objective of this paper is to present a taxonomy of correlated and communication equilibrium outcomes that is based on the properties of the mechanisms capable of implementing them. The questions this goal raises are quite different – both in substance and in the relevant techniques – from those arising in the context of complete information games. They have some formal similarity, which is reflected in the similar terminology, to the question of implementability of social choice functions studied in mechanism design theory. However, the subject matter here is not a special case of the latter (see also Kar et al., 2010).

The "plan of attack" of the present work is to separate the question of the implementability of correlated and communication equilibrium outcomes into three interrelated questions. The first question is the implementability of correlated strategy distributions, where only the joint distribution of the players' types and actions matters and payoffs are irrelevant. The second question, which does take payoffs and incentive compatibility into account, is the implementability of correlated and communication equilibrium distributions. The third, and arguably most important, question is the implementability of payoff vectors. The players' (expected) payoffs are uniquely determined by the joint distribution of types and actions but not conversely. The advantages of this three-part approach, when compared with directly addressing the third problem, are that it makes certain issues significantly more manageable and provides insights about the roots of non-implementability where the latter occurs.

The framework laid out in this paper accommodates the majority of the previously described varieties of correlated strategies, correlated equilibria and communication equilibria in

[1] The two solution concepts may be viewed as special cases of a third, more general, one, where both the true and the reported types may affect the messages that the players receive from the mechanism. See Section 8.3.

incomplete information games, and adds a good number of new ones. Different kinds of equilibria may be related in that one is a special case of another. As this paper shows, these relations form a rich and intricate structure, and they are not always obvious or perfectly intuitive. This is quite different from the complete information case, where the three kinds of equilibria are simply linearly ordered.

The next section illustrates the various issues involved in this work by means of examples. Section 3 describes the connections with the existing literature on incomplete information games. Section 4 lays out the formal framework. The three subsequent sections present results, primarily in the form of three Hasse diagrams that show the relations between the different kinds of correlated strategy outcomes (Section 5), correlated equilibrium outcomes (Section 6) and communication equilibrium outcomes (Section 7). Section 8 synthesizes and extends these results, and in particular uses them as the basis for the classification of the possible outcomes according to the properties of the implementing mechanisms. The paper's four appendices include the majority of the proofs and several additional examples.

2 Illustrative Examples

The following four examples illustrate some of the questions studied in this paper and the framework used for tackling them. In addition, each example points to a specific non-obvious way in which the properties of a mechanism affect the correlated strategies, correlated equilibria or communication equilibria implementable by it.

Example 1

Consider the following simple 2×2 Bayesian game (Milchtaich, 2004, Example 6). The two players have identical two-element action spaces, $A_1 = A_2 = \{L, R\}$, and identical two-element type spaces, $T_1 = T_2 = \{+1, -1\}$. The four type profiles are equally probable, so that types are independent. (Independence is not a crucial assumption. It would suffice to assume that all type profiles have positive probability.) A correlated strategy distribution is defined as follows: (i) If both players have type $+1$, the action profile is either (L, L) or (R, R), each with probability 0.5, and (ii) if the type profile is any of the other three, the action profile is either (L, R) or (R, L), each with probability 0.5.

Since in the above distribution the players' actions are correlated but their types are not, implementation requires an extraneous correlation mechanism. For example, the players may use as cues for their choice of actions certain signals, or messages, that they receive from the same outside source or different sources. The message that each player receives may conceivably depend on his type (e.g., because of type-specific perceptual abilities; see the Introduction), but it may be affected by the other player's type only if the sending source is somehow aware of the latter. It may seem doubtful that such effect is necessary. This is because the above distribution has the so-called conditional independence property (see Section 5.1): the action of a player of a given type is statistically independent of the other player's type. Indeed, the probability that the action is L or R is always 0.5, regardless of *both* players' types. It turns out, however, that even though each player's action does not reflect the other player's type, the latter *must* have an effect on the cue used for selecting the action. The distribution specified above cannot be implemented by any mechanism where a player's type has no effect whatsoever on the message received by the other player. (In other words, this is not a type correlated distribution; see Section 4.3.)

The last assertion can be proved as follows. Consider any mechanism that sends to the players messages that are affected only by their own types. These messages may be random, and so are described by four, possibly statistically dependent, random variables $\boldsymbol{m}_1^+$, $\boldsymbol{m}_1^-$, $\boldsymbol{m}_2^+$ and $\boldsymbol{m}_2^-$, where the subscript identifies the player and the superscript identifies his type. Each possible realization $(m_1^+, m_1^-, m_2^+, m_2^-)$ of these random variables determines a quartet of actions $(a_1^+, a_1^-, a_2^+, a_2^-)$ according to the players' strategies, which specify the way each player uses the message he receives as a cue for choosing his action. (For example, a_1^+ is the action that the message m_1^+ prompts player 1 to choose when his type is $+1$.) For the quartet to be consistent with the specified distributions, it must satisfy $a_1^+ = a_2^+$, $a_1^+ \neq a_2^-$, $a_1^- \neq a_2^+$ and $a_1^- \neq a_2^-$.

However, since there are only two possible actions, the equality and the three inequalities are contradictory. The contradiction proves the above assertion.

A simple mechanism capable of implementing the above distribution is one that, for each type profile, simply randomizes between the two possible action profiles and tells each player his action. The mechanism must be aware the players' types, since they determine the allowed action pairs. Nevertheless, as indicated, no information about the types is revealed by the action of any single player, even if his own type his known.

Example 2

The message that a mechanism sends to a player may reflect not only the other players' types but also the player's own type. The latter is an expression of the *selectivity* of the message, in particular, differences between player types in the information they get about the other players. As the following example demonstrates, selectivity may be necessary for making certain choices of actions incentive compatible.

Consider a symmetric 2×2 Bayesian game with the same game structure and distribution of type profiles as in Example 1, and where the two players always get equal payoffs, which for a type profile (t_1, t_2) are given by the payoff matrix

$$\begin{matrix} & L & R \\ L & \multirow{2}{*}{} -t_1t_2 & 0 \\ R & 0 & 2t_1t_2 \end{matrix}$$

Thus, depending on the type profile, L or R is a dominant action for both players. A mechanism sends to each player a message, which is a type profile. For type $+1$ of player 1 and for both types of player 2, the type profile is the real one (t_1, t_2). For type -1 of player 1, the message is always $(-1, -1)$. This mechanism and the pair of strategies that simply instruct each player to choose the dominant action for the type profile that is specified by the message he receives together constitute a correlated equilibrium, which has the following distribution: if the type profile is $(+1, +1)$, $(+1, -1)$, $(-1, +1)$ or $(-1, -1)$ (each possibility has probability $1/4$), the action profile is (R, R), (L, L), (R, L) or (R, R), respectively. The expected payoffs of type $+1$ of player 1 and types $+1$ and -1 of player 2 are $3/2$, 1 and $3/2$, respectively, and these playoffs are clearly the highest these player types could achieve by choosing any actions. The same is true for type -1 of player 1, whose expected payoff would decrease from 1 to $1/2$ if he switched from his (constant) action R to L.

The mechanism described above sends information about player 2 only to type $+1$ of player 1. Such a differential treatment of 1's types is necessary. The above correlated equilibrium distribution is not implementable by any mechanism where for each type of player 2 the message that player 1 receives is statistically independent of his own type. To see this, suppose that such a mechanism exists, and let m_1^+ and m_1^- be some specific messages that player 1 receives with positive probability (which is the same for both types of that player) when player 2's type is $+1$ and -1, respectively. The two messages cannot be identical, for otherwise type $+1$ of player 1 would not always be able to tell 2's type, which is inconsistent with the fact that (according to the above distribution) with probability 1 he plays R if the type is $+1$ and L if it is -1. Therefore, every such m_2^+ and m_2^- must be distinct, which entails that player 1 can always tell by his message the type, and hence the action, of player 2. However, this is inconsistent with incentive compatibility, since it implies that, by always choosing the same action as player 2, type -1 of player 1 could increase his payoff from 1 to $3/2$. This contradiction proves that a mechanism that implements the above correlated equilibrium distribution cannot have the property that (if the type of player 2 is known) the message to player 1 never reveals any information about his own type.

Example 3

A non-obvious, yet important distinction exists between the requirement that the message to a particular player does not *reveal* any information about the type of that player (as in the case considered in the last paragraph) or about another player's type and the stronger requirement that the message is not *affected* by that type (which is the property considered in Example 1). The significance of this distinction is illustrated by the following example (see also Section 4.2).

			Player 2						
			Type $+1$				Type -1		
			L	R			L	R	
Player 1	Type $+1$	L	0.75	0	0.75	L	0.25	0.25	0.5
		R	0	0.25	0.25	R	0.5	0	0.5
			0.75	0.25			0.75	0.25	
			L	R			L	R	
	Type -1	L	0.25	0.5	0.75	L	0	0. 5	0.5
		R	0.25	0	0.25	R	0.5	0	0.5
			0.5	0.5			0.5	0.5	

Table 1. A correlated equilibrium distribution for Example 3. The four type profiles are equally probable. For each of them, the joint distribution of the players' actions (each of which can be L or R), as well as the marginal distributions, are shown.

The game is the same as in Example 2. A correlated equilibrium distribution in this game is given by Table 1, which specifies the (conditional) distribution of the players' action profile for each type profile. The marginal distributions, i.e., the probability that a player of a given type plays L or R, depend only on the opponent's type. Specifically, if the latter is $+1$ or -1, L has probability 0.75 or 0.5, respectively. Therefore, a mechanism that randomly chooses an action profile (a_1, a_2) according to the probabilities that the table specifies for the players' actual type profile, and reports a_1 to player 1 and a_2 to player 2, has the property that the message to a player reveals no information about his own type. This mechanism and the strategies of acting according to the messages it sends together constitute a correlated equilibrium, as it is not difficult to check that a player can never increase his expected payoff by deviating to the other action. It is also true, but less easy to check, that the correlated equilibrium distribution in Table 1 is not implementable by any mechanism with the stronger property that the message sent to each player is unaffected by the player's type. In fact, it takes a computer to check this. Although the problem is a standard linear programming one – it needs to be checked that a particular system of linear equalities and inequalities does not have a solution – the number of variables and equalities/inequalities involved (at least 256 and 20, respectively) is far too great for manual calculations.

Some intuition about why a mechanism as above cannot implement the correlated equilibrium distribution in Table 1 can be gained by considering two conceivable mechanisms where a player's type does not affect the message he receives. The first mechanism randomly selects four actions, a_1^+, a_1^-, a_2^+ and a_2^-, according to some suitable joint distribution, and then, depending on whether the type of player i $(= 1,2)$ is $+1$ or -1, instructs the *other* player j to take the action a_j^+ and a_j^-, respectively. However, such a mechanism cannot yield the distribution in Table 1. The reason is similar to that in Example 1, and in particular, it does not involve incentives (i.e., payoffs).

The second conceivable implementing mechanism for the distribution in Table 1 sends as a message to each player not a single action but a pure strategy, which is a specification of action for each of the player's types. It leaves it to the player to choose the action corresponding to his actual type. Specifically, the mechanism first chooses action profiles according to the probabilities specified by Table 1, independently for each of the four type profiles. Then, based on the players' actual type profile (t_1, t_2), it tells player 1 both his action for the type profile $(+1, t_2)$ and his action for $(-1, t_2)$, and similarly for player 2 (so that each player's own type does not affect the message he receives). If the players use the messages in the intended manner, that is, they take the first or second action if their type is $+1$ or -1, respectively, then the action distributions for the four type profiles are indeed as in Table 1. However, this is not a correlated equilibrium. The reason is that the double message conveys too much information about the other player's type. By assumption, the prior probability that player 2 has type $+1$ is 0.5. By Bayes' rule, the posterior probability that 2 has that type given that player 1 plays L is $(0.75/(0.75 + 0.5) =)$ 0.6. Thus, telling player 1 which action he should take gives him some information about 2's type, but not too much information, in the sense that taking the action is still optimal for him. A double message as above amounts to two independent draws from the same unknown distribution, which provide more information about the underlying distribution than a single draw. For example, telling player 1 that he should play L whether his type is $+1$ or

-1 increases the (posterior) probability that 2's type is $+1$ to almost 0.7. This probability is greater than $2/3$, which means that, regardless of the actions that the two types of play 2 take, L is *not* the optimal action for type $+1$ of player 1. Thus, a player may deduce from the additional information conveyed by the double message he receives that his expected payoff from taking the action he is supposed to take is actually less than for the alternative action.

Example 4

The following example, which is taken from Gerardi (2004, Example 2), illustrates the difference between implementability of correlated equilibrium distribution and of communication equilibrium distribution. By definition, in a communication equilibrium, the players report their types to the mechanism and truth telling is incentive compatible. The example shows that to satisfy this requirement, any implementing mechanism must use randomization. Thus, randomization has a different role than in Examples 1 and 3, where it is required simply because for some type profiles the player's actions are not unique. Here (as in Example 2), a single action profile is assigned to each type profile. However, without randomization, incentive compatibility cannot be upheld.

In a three-player Bayesian game, player 1 has two types, t_1' and t_1'', player 2 has two types, t_2' and t_2'', and player 3 has a single type. Types t_1'' and t_2'' cannot occur together, but the other three type profiles are all possible and equally probable. Players 1 and 2 have a single action, and player 3 has four actions, a_3^1, a_3^2, a_3^3 and a_3^4. The four payoff vectors corresponding to these actions, for each type profile, are given by the following table:

	t_2'	t_2''
t_1'	$(-1,-1,1), (1,1,0), (0,0,-1), (0,0,-1)$	$(1,0,0), (1,1,1), (-3,0,1), (1,0,-1)$
t_1''	$(0,1,0), (1,1,1), (0,1,-1), (0,-3,1)$	N/A

According to the table, for each of the three possible type profiles, player 3 receives his maximum payoff of 1 only by choosing one or two particular actions, and the former holds only for (t_1', t_2') (for which the unique optimal action is a_3^1). Thus, there are four different ways to choose optimal actions for each type profile. Each such choice of actions specifies a correlated strategy distribution, which is implementable by a mechanism that simply tells player 3 the types of the other players. However, it can be shown that only one of these four is also a communication equilibrium distribution, namely, the one in which player 3 chooses action a_3^2 if the other players' types are either (t_1', t_2'') or (t_1'', t_2'). In addition, in any communication equilibrium, player 3 randomizes exactly fifty-fifty between a_3^3 and a_3^4 if the impossible type combination (t_1'', t_2'') is reported, for otherwise truthful type reports would not be incentive compatible for player 1 or 2. It follows that any mechanism that implements the above communication equilibrium distribution necessarily involves randomization.

3 Related Literature

As Aumann (1987) showed, a correlated equilibrium may be viewed as an expression of Bayesian rationality. A rational player's choice of action is optimal given his knowledge of the state of the world. The latter includes a specification of the knowledge of the other players as well as their actions, which must also be optimal. Aumann's paper is concerned only with complete information games; types of players and type-dependent payoffs are not part of the setting. However, since the state-space formulation is a standard model for Bayesian games, the paper indicated the logical next step, which was to merge the two settings by allowing the states of the world to determine not only the players' types but also any additional information that they possess and may use for choosing their actions. Crucially, that additional information is not specified by the game – it is part of the solution concept. Herein lies the major difference between correlated equilibrium and the narrower concept of Bayesian equilibrium. In the latter, players have to do with the information that they receive within the game, and possibly private coin tosses.

Two models of the kind outlined above were proposed by Cotter (1991, 1994). They differ from one another in the restrictions they put on the players' additional information. In a strategy correlated equilibrium (Cotter, 1991), the additional information takes the form of random messages that the players receive from an outside correlation mechanism, which is ignorant of their types. A correlated strategy with such a mechanism is a prescription of a (pure or randomized) action for each type of each player, which depends on the message he receives. The equilibrium condition is that taking the prescribed action is incentive compatible in that no player can increase his expected payoff by taking a different action. A type correlated equilibrium (Cotter, 1994; see also Samuelson and Zhang, 1989) can be described as a strategy correlated equilibrium in a version of the game in which each type of each player is an independent agent, so that the message a player receives may depend on his type. However, as in a strategy correlated equilibrium, the message is unaffected by the other players' types. Consequently, the player's action is necessarily conditionally independent of the other players' types, given the player's own type. Cotter asserted that this conditional independence property is characteristic of type correlated equilibria, in that any correlated equilibrium distribution that has it can be implemented by a mechanism as above. However, Example 1 shows that this assertion is incorrect (see also Section 5.1).

A different extension of correlated equilibrium to games with incomplete information is communication (or mediated) equilibrium (Myerson, 1994). This solution concept is characterized by bidirectional communication. The players first send private messages to, and then receive such messages from, a particular mechanism, which thus serves as a mediator as well as a correlation device. According to the revelation principle (see Myerson, 1994), without loss of generality the messages sent by the players to the mediator may be assumed type reports. The message that each player gets from the mediator indicates a particular action for that player. The mechanism is required to be incentive compatible in that it is in each player's best interest to report his type honestly and take the indicated action if all the others do the same.

A comprehensive account of correlated and communication equilibria in games with incomplete information is Forges' (1993) paper, which compares strategy correlated equilibrium, type (or agent normal form) correlated equilibrium, communication equilibrium, and 'Bayesian solution'. (A fifth solution concept considered in the paper concerns hierarchies of beliefs.) Bayesian solution is a very general solution concept, which fully realizes the extension of Aumann's notion of Bayesian rationality to incomplete information games. It includes strategy and type correlated equilibria as special cases. A Bayesian solution extends a given incomplete information game by introducing a state space where several states may correspond to a single type profile. This allows players to have partial or complete information about the other players' types as well as about outside events. As in Aumann's (1987) model, the information structure is complemented by a mapping from states to action profiles that is required to satisfy the obvious incentive compatibility condition. A Bayesian solution may be implemented by an omniscient mediator, who knows the players' types. Thus, unlike in a communication equilibrium, only one-way communication is required.[2]

As indicated, the messages that the players receive from the mediator are part of the solution concept and are thus distinct from any signals they receive as part of the game itself, which are completely specified by their types. However, the potential impact of the former (the mediator's messages) may depend on the degree of dependence among the latter (the players' types). For example, with perfectly correlated types, the players effectively already know each other's type when they receive the mediator's messages, which can therefore only help them to coordinate their actions. Conversely, if the types are independent, the mediator's messages may also inform players about the other players' types. However, this may be so only if the solution concept allows the messages to depend on the other players' types. Therefore, depending on

[2] As noted by Bergemann and Morris (2011), Forges' formal definition of Bayesian solution is in fact not as general as it could be, in that the mediator knows only the players' types, and not any other payoff-relevant information that is not extractable from the type profile. They call the generalization that allows the mediator to possess and use such information *Bayes correlated equilibrium*.

the solution concept, garbling, or randomly perturbing, in a particular way the signals that the players receive as part of the game may or may not change the set of equilibrium outcomes.[3] Lehrer et al. (2006) identified the kinds of garbling that do not affect the equilibrium outcomes for three kinds of correlated equilibrium in two-player Bayesian games: mixed-strategy equilibrium, type correlated equilibrium, and a special kind of Bayesian solution (called belief invariant Bayesian solution by Forges, 2006), which satisfies a condition similar to the conditional independence property. They showed, for example, that garbing has no effect on mixed-strategy equilibria (regardless of the payoff functions) if and only if it is performed independently for each player, that is, without taking into account the other player's signal.

Identification of information types (Milchtaich, 2004) is a special kind of garbling. It removes distinctions between player types that are interchangeable in terms of their effect on the player's own payoff and on those of the others and differ, say, only in what the player knows about the other players' types. Identification of information types may transform one kind of equilibrium into another. For example, a pure-strategy equilibrium (with different actions for different information types) may become a mixed-strategy equilibrium (with several possible actions for the single type that results from the identification). Therefore, the collection of pure-strategy equilibria is not closed under identification of information types. The same is true for certain more general solution concepts. For example, this is so for type correlated equilibrium, since when information types are identified, the conditional independence property may cease to hold (Milchtaich, 2004, Examples 7). In fact (see Milchtaich, 2004, Propositions 4 and 5), the narrowest extension of pure-strategy equilibrium that is closed under identification of information types is the notion of correlated equilibrium used in the present work, which is similar to Forges' (1993, 2006) Bayesian solution, or global equilibrium in the terminology of Lehrer et al. (2006). Thus, in this respect at least, this solution concept is not excessively broad.

4 Preliminaries

4.1 Bayesian games

An n-player (finite) *pre-Bayesian game* is a function $u = (u_1, u_2, \dots, u_n): T \times A \to \mathbb{R}^n$, where $T = T_1 \times T_2 \times \cdots \times T_n$ and $A = A_1 \times A_2 \times \cdots \times A_n$ are each the Cartesian product of n finite sets and $\mathbb{R}$ is the real line. $N = \{1,2, \dots, n\}$ is the set of *players*. For each player i, the sets T_i and A_i and the function u_i are respectively the *type space*, *action space* and *payoff function* of player i. For each *type profile* $t = (t_1, t_2, \dots, t_n) \in T$ and *action profile* $a = (a_1, a_2, \dots, a_n) \in A$, $u(t, a)$ is the resulting n-tuple of *payoffs*.

One interpretation of player types is that they represent signals that players receive about the "state of nature" (Osborne and Rubinstein, 1994). The latter may conceivably have a direct effect on the players' payoffs, which is apparently not the case in the above "reduced" formulation, where the payoffs are completely determined by the type and action profiles. However, the general case can be formulated simply by adding nature as a (dummy) player, whose type space consists of all states of nature. A state may represent complete resolution of all uncertainty in the game, and in particular determine the (real) players' types. Alternatively, a state may represent only certain "hidden variables" that affect the players' payoffs but cannot be extracted from the type profile. Nature does not take action, and so the corresponding action space is necessarily a singleton and the payoff function is irrelevant.

A *pure strategy* for a player i is a function from the player's type space to his action space, i.e., an element of $A_i^{T_i}$. Using some fixed indexing of the (finite) type space, $T_i = \{t_i^1, t_i^2, \dots\}$, any pure-strategy can be written as $(a_i^1, a_i^2, \dots)$, where, for each j, a_i^j is the action of the jth type of player i. A *pure-strategy profile* is an assignment of a pure strategy to each player, i.e., an element of $A_1^{T_1} \times A_2^{T_2} \times \cdots \times A_n^{T_n}$. It can be written as $(a_1^1, a_1^2, \dots; a_2^1, a_2^2, \dots; \dots; a_n^1, a_n^2, \dots)$.

[3] This assumes that the players' types only represent information and do not have a *direct* effect on payoffs.

A pre-Bayesian game becomes a (finite) *Bayesian game* when it is coupled with a specified probability measure on T, the players' *common prior* η_T. The collection of all type profiles to which the common prior assigns positive probability, i.e., its support $\mathrm{supp}(\eta_T)$, may be a proper subset of T. However, it is assumed (essentially without loss of generality) that every type t_i of every player i is supported, in the sense that $(t_i, t_{-i}) \in \mathrm{supp}(\eta_T)$ for some *partial type profile* $t_{-i} = (t_1, \dots, t_{i-1}, t_{i+1}, \dots, t_n)$. It is useful to view the common prior as the distribution of some random variable with values in T,[4] the *random type profile* $\boldsymbol{t} = (\boldsymbol{t}_1, \boldsymbol{t}_2, \dots, \boldsymbol{t}_n)$.[5] Thus,

$$\eta_T(\{t\}) = \Pr(\boldsymbol{t} = t), \qquad t \in T.$$

For each player i, the conditional distribution of $\boldsymbol{t}$, given $\boldsymbol{t}_i$, may be interpreted as the player's posterior beliefs about the type profile after he learns his own type.

4.2 Mechanisms

A mechanism for an n-player Bayesian game is an extraneous source of messages[6], which the players receive before they choose their actions. The message m_i that each player i receives is an element of some finite set M_i, the player's (received) *message space*. It may reflect to a lesser or greater extent the players' type profile. As indicated, one of these players may be nature, and in this case, the messages may provide the real players with payoff-relevant information that would not otherwise be available to them even through information pooling. Formally, a mechanism is a random variable $\boldsymbol{m}$ that is (statistically) independent of the random type profile $\boldsymbol{t}$, with values that are functions from T to the product set $M = M_1 \times M_2 \times \cdots \times M_n$. Evaluation at (in other words, projection on) any type profile $t \in T$ gives the corresponding *random message profile* $\boldsymbol{m}(t) = (\boldsymbol{m}_1(t), \boldsymbol{m}_2(t), \dots, \boldsymbol{m}_n(t))$, which describes the messages that the players receive when their types are t. Since the type profile is itself random, the actual messages are given by the random variable $\boldsymbol{m}(\boldsymbol{t})$, whose value is the message profile obtained by evaluating the function returned by $\boldsymbol{m}$ at the type profile returned by $\boldsymbol{t}$. Note that the assumption that $\boldsymbol{m}$ and $\boldsymbol{t}$ are (statistically) independent, which expresses the extraneous nature of the messages (see Section 4.2.1 for further discussion of this point), does not contradict the assertion that the latter may reflect the type profile. To take an extreme example, if the mechanism is an outside observer who is capable of finding out the players' types, his message to each player i may fully convey that information:

$$\boldsymbol{m}_i(t) = t, \qquad t \in T. \tag{1}$$

In this example, independence holds simply because the mechanism does not perform any randomization, and thus serves purely as a source of information about the other players' types. Other mechanisms may partially or exclusively serve as randomization devices, and convey little or no information about the types. A finer, more exact classification of mechanisms is facilitated by the following list of possible properties, which are referred to in this paper as the *fundamental properties* of mechanisms. Each property is expressed by a condition that the message that each player i receives from the mechanism is required to satisfy for every pair of type profiles $t = (t_i, t_{-i})$ and $t' = (t'_i, t'_{-i})$.

(S) The message is unaffected by the receiving player's type:

$$\boldsymbol{m}_i(t) = \boldsymbol{m}_i(t'_i, t_{-i}).$$

---

[4] A *random variable*, in this work, is any function from a finite probability space where each point has positive probability to a finite set. Enumerating the range would in principle enable viewing the random variable as real- or integer-valued, but doing so has little practical value. Random variables are denoted by boldface letters and, following common practice, their arguments are always suppressed.

[5] One way to obtain a random type profile is by restricting the common prior to its support and defining $\boldsymbol{t}$ as the identity map on $\mathrm{supp}(\eta_T)$. However, since only the distribution of the random type profile is specified, there is actually more than one version of it. Any reference to the random type profile is taken to mean that differences between versions are irrelevant, which entails that the reference is ultimately to the common prior itself.

[6] The term 'messages', rather than 'signals', is used here to emphasize the assumption that the sending mechanism is part of a solution concept rather than the game. As indicated, 'signals' in an incomplete information game are often synonymous with the players' types, which *are* part of the game.

($\tilde{S}$) The message provides no information about the receiving player's type:

$$\boldsymbol{m}_i(t) \stackrel{d}{=} \boldsymbol{m}_i(t_i', t_{-i}).$$[7]

(O) The message is unaffected by the other players' types:

$$\boldsymbol{m}_i(t) = \boldsymbol{m}_i(t_i, t_{-i}').$$

($\tilde{O}$) The message provides no information about the other players' types:

$$\boldsymbol{m}_i(t) \stackrel{d}{=} \boldsymbol{m}_i(t_i, t_{-i}'). \tag{2}$$

(D) The message is non-random:

$\boldsymbol{m}_i(t)$ has a degenerate distribution.[8]

(I) The message is (statistically) independent of the messages that the other players receive:

$$\boldsymbol{m}_1(t), \boldsymbol{m}_2(t), \ldots, \boldsymbol{m}_n(t) \text{ are independent.} \tag{3}$$

The equality between random variables required by properties S and O means equality with probability 1 (equivalently, pointwise equality). Conditions $\tilde{S}$ and $\tilde{O}$, by contrast, only require equality between the distributions of two random variables, for specified type profiles. Since the players' types are actually random, the latter translates into equality between *conditional* distributions. Specifically, $\tilde{S}$ entails that the message that player i receives is conditionally independent of his type $\boldsymbol{t}_i$, given the other players' types $\boldsymbol{t}_{-i}$. In property $\tilde{O}$, $\boldsymbol{t}_i$ and $\boldsymbol{t}_{-i}$ are interchanged.

As this work shows, the distinction between equality with probability 1 and equality in distribution has significant implications for correlated strategies and equilibria.[9] Nevertheless, it does not seem to have received sufficient attention in the existing literature on games with incomplete information. Indeed, properties S and O cannot even be stated using the formalism employed, e.g., by Myerson (1994). That formalism specifies the distribution of the random message profile separately for each type profile, and thus provides no language for describing connections between the messages corresponding to different type profiles.

The list of properties of mechanisms is potentially extendable beyond the fundamental properties. Two conceivable additions are the "duals" of O and $\tilde{O}$, which assert that the type of a player does not affect the other players' messages or (at least) is not reflected in their joint distribution. These additions may be particularly meaningful if the player in question is nature (see Section 4.1). They mean, in this case, that even if the real players pooled their information, their messages would not tell them anything about the state of nature beyond what they can learn already from their types.

4.2.1 Independence lemma and the default mechanism

The definition of mechanism in effect posits that the randomness or uncertainty regarding the messages that the players receive has two independent potential sources: the inherent randomness of the type profile, which the messages may reflect, and residual randomness, which persists also with a specified type profile (and reflects, for example, the use of noisy communication channels). The following lemma shows that the assumption that the two sources are independent involves no loss of generality. Any joint distribution of types and

[7] The symbol $\stackrel{d}{=}$ denotes equality in distribution.

[8] A distribution is degenerate if it assigns probability 1 to some value. For example, this is the case in (1).

[9] As an example of this distinction, if a mechanism satisfies $\tilde{O}$, and the types of all but two players change, the distribution of the message that each of these two receives does not change. However, the *joint* distribution of the two messages may change, for example, uncorrelated messages may become correlated. By contrast, O would imply that the joint distribution also does not change when the other players' types change. A subtler, yet highly consequential (See Section 5.1), difference between O and $\tilde{O}$ applies also to two-player games.

messages can be produced by "mixing" the random type profile $\boldsymbol{t}$ with a suitable mechanism $\boldsymbol{m}$, which by definition is independent of $\boldsymbol{t}$.

Lemma 1. Let $M = M_1 \times M_2 \times \cdots \times M_n$ be any finite product set and η any probability measure on $T \times M$ whose marginal on T is equal to the common prior. There exists a random variable $\boldsymbol{m}$ that is independent of the random type profile $\boldsymbol{t}$, with values that are functions from T to M, such that the joint distribution of $\boldsymbol{t}$ and $\boldsymbol{m}(\boldsymbol{t})$ is equal to η and, in addition:

(i) The random profiles $\{\boldsymbol{m}(t)\}_{t\in T}$ are independent. (4)

(ii) For every $t \in T$ and $i \in N$ there is some $t' \in T$ such that $(t_i, t'_{-i}) \in \operatorname{supp}(\eta_T)$ and (2) holds. (5)

(iii) For every $t \notin \operatorname{supp}(\eta_T)$, (3) holds. (6)

Proof. Given a probability measure η as above, construct a family $\{\boldsymbol{m}(t)\}_{t\in T}$ of independent random variables with values in M, indexed by the type profiles, that are collectively independent of $\boldsymbol{t}$ and are distributed as follows. For $t \in \operatorname{supp}(\eta_T)$, the distribution of $\boldsymbol{m}(t)$ is the probability measure on M that assigns to each element m the (conditional) probability

$$\frac{\eta(\{(t,m)\})}{\eta_T(\{t\})}.$$

Note that this already implies that the joint distribution of $\boldsymbol{t}$ and $\boldsymbol{m}(\boldsymbol{t})$ is equal to η:

$$\Pr(\boldsymbol{t} = t, \boldsymbol{m}(\boldsymbol{t}) = m) = \eta_T(\{t\}) \Pr(\boldsymbol{m}(t) = m) = \eta(\{(t,m)\}), \qquad t \in T, m \in M.$$

For $t \notin \operatorname{supp}(\eta_T)$, choose for each $i \in N$ some type profile t' such that $(t_i, t'_{-i}) \in \operatorname{supp}(\eta_T)$. The distribution of $\boldsymbol{m}(t) = (\boldsymbol{m}_1(t), \boldsymbol{m}_2(t), \dots, \boldsymbol{m}_n(t))$ is then completely specified the requirement that (2) and (3) hold. ■

A mechanism $\boldsymbol{m}$ as in Lemma 1 is referred to in this paper as the *default mechanism* of the measure η.[10] The three special properties that the default mechanism is required to possess may seem purely technical. Property (4) concerns relations between different type profiles, which by definition cannot coexist (since only one type profile is realized), while (5) (which is not trivial only for $t \notin \operatorname{supp}(\eta_T)$) and (6) concern type profiles t that cannot occur ($\Pr(\boldsymbol{t} = t) = 0$). However, property (4) is not innocuous, as it essentially prevents the mechanism from satisfying S or O unless it also satisfies D. This is because two random variables that are equal *and* independent necessarily have a degenerate distribution. Mainly because of this limitation, it is not possible to restrict attention to default mechanisms. They are, however, useful technical constructs.

4.3 Correlated strategies

A *correlated strategy* in an n-player Bayesian game is a pair $(\boldsymbol{m}, \sigma)$ consisting of a mechanism $\boldsymbol{m}$ and a profile of strategies $\sigma = (\sigma_1, \sigma_2, \dots, \sigma_n)$. The strategy of each player i is a function $\sigma_i: T_i \times M_i \to A_i$ that specifies how the player's type t_i and the message he receives m_i together determine the player's action a_i.[11] As indicated, the messages are part of the correlated strategy's specification rather than the game. Their (potential) randomness and that of the types means that the actions are also random. The *random action profile* corresponding to a correlated strategy $(\boldsymbol{m}, \sigma)$ is the random variable $\boldsymbol{a} = (\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_n)$ given by

[10] Since some aspects of the default mechanism are only partially specified, the definition actually allows for more than one version of it. Any reference to the default mechanism is taken to mean that differences between versions are irrelevant.

[11] By assumption, the specification of the actions is deterministic; randomized actions are not allowed. This assumption involves no loss of generality and, in particular, does not preclude mixed strategies. It only means that even private randomization is viewed as part of a single large mechanism. If there is already a mechanism from which the players receive messages, randomization may be relegated to it. This may be done by appending a random number to the massage that each player receives, such that the n random numbers are independent. A mechanism modified in this way does not satisfy D, but the modification has no effect on properties $S, \tilde{S}, O, \tilde{O}$ or I.

$$\boldsymbol{a}_i = \sigma_i(\boldsymbol{t}_i, \boldsymbol{m}_i(\boldsymbol{t})), \qquad i \in N. \tag{7}$$

Pure- and mixed-strategy profiles are essentially special cases of correlated strategy. If the message that each player receives is unaffected by his or the others' types and is also non-random, that is, if the mechanism satisfies S, O and D (which effectively means that there are no messages at all), then, for some fixed $m_1, m_2, \dots, m_n$,

$$\boldsymbol{a}_i = \sigma_i(\boldsymbol{t}_i, m_i), \qquad i \in N.$$

In this case, for each player i, σ_i associates a (deterministic) action a_i with each type t_i, which means that the correlated strategy is effectively a pure-strategy profile. A mixed-strategy profile is described by a correlated strategy with a mechanism that satisfies S, O and I. These properties of the mechanism mean that the messages are independent and equal to $\boldsymbol{m}_1(t'), \boldsymbol{m}_2(t'), \dots, \boldsymbol{m}_n(t')$, where t' is any fixed type profile. The corresponding random action profile is

$$(\sigma_1(\boldsymbol{t}_1, \boldsymbol{m}_1(t')), \sigma_2(\boldsymbol{t}_2, \boldsymbol{m}_2(t')), \dots, \sigma_n(\boldsymbol{t}_n, \boldsymbol{m}_n(t'))).$$

This is effectively the same as, and it can be implemented by, private randomization over pure strategies independently for each player.[12]

The *correlated strategy distribution* (CSD) of a correlated strategy $(\boldsymbol{m}, \sigma)$ is the joint distribution of the random type profile $\boldsymbol{t}$ and the random action profile $\boldsymbol{a}$ specified by (7). Many different correlated strategies, employing different mechanisms, may have the same CSD. A CSD is *implementable* by a mechanism $\boldsymbol{m}$ (which *implements* the distribution) if it is the distribution of some correlated strategy of the form $(\boldsymbol{m}, \sigma)$, that is, if suitable strategies $\sigma_1, \sigma_2, \dots, \sigma_n$ exist. A CSD is a *pure-strategy distribution* or a *mixed-strategy distribution* if it is implementable by some mechanism $\boldsymbol{m}$ with properties S, O and D or properties S, O and I, respectively. The former is obviously a special case of the latter, which in turn is a special case of *type correlated distribution*, which is defined as a CSD that is implementable by a mechanism with property O (that is, one whose message to each player may reflect the player's type but is not affected by the types of the other players).

Every CSD η is a probability measure on $T \times A$ whose marginal on T coincides with the common prior η_T.[13] Hence, it has a default mechanism $\boldsymbol{m}$ (see Section 4.2.1), in which the message space M_i of each player i is his action space A_i. Thus, the default mechanism simply tells each player what action to take. The *default correlated strategy* of η is defined as $(\boldsymbol{m}, \sigma)$, where

$$\sigma_i(t_i, m_i) = m_i, \qquad i \in N. \tag{8}$$

Thus, each player's strategy simply instructs him to obey the default mechanism. The resulting random action profile $\boldsymbol{a}$, called the *default random action profile* of η, satisfies

$$\boldsymbol{a} = (\sigma_1(\boldsymbol{t}_1, \boldsymbol{m}_1(\boldsymbol{t})), \sigma_2(\boldsymbol{t}_2, \boldsymbol{m}_2(\boldsymbol{t})), \dots, \sigma_n(\boldsymbol{t}_n, \boldsymbol{m}_n(\boldsymbol{t}))) = \boldsymbol{m}(\boldsymbol{t}). \tag{9}$$

The second equality shows that the default mechanism $\boldsymbol{m}$ of a CSD η implements it. Obviously, a similar equality holds for the default mechanism and strategy of any probability measure on $T \times A$ with the marginal η_T. Hence, every such measure is a CSD. This conclusion establishes the following simple result.

Lemma 2. In a Bayesian game, a probability measure on $T \times A$ is a correlated strategy distribution if and only if its marginal on T is equal to the common prior.

[12] A correlated strategy with a mechanism as above may alternatively be viewed as a *behavior strategy* for each player, that is, a randomized action for each of the player's types.

[13] Since the common prior is given as part of the specification of the game, a CSD may also be viewed as an assignment of a probability measure on A to every type profile t, namely, the distribution of the players' action when their types are given by t. If t lies outside the support of η_T, the probability measure may be chosen arbitrarily.

4.4 Correlated equilibria

The players' incentives in a Bayesian game are embodied by their payoff functions, $u_1, u_2, \dots, u_n$. For a correlated strategy $(\boldsymbol{m}, \sigma)$, the expected payoff (hereafter referred to usually simply as the payoff) of each player i is the expectation of the random variable $u_i(\boldsymbol{t}, \boldsymbol{a})$, where $\boldsymbol{t}$ is the random type profile and $\boldsymbol{a}$ is the random action profile given by (7). The correlated strategy is incentive compatible if none of the players i can increase his payoff by unilateral deviation, that is, by switching from σ_i to some other strategy $\sigma_i'\colon T_i \times M_i \to A_i$. Equivalently, incentive compatibility means that the action that the correlated strategy specifies for each player i always maximizes the conditional expectation[14] of $u_i(\boldsymbol{t}, \boldsymbol{a})$, given the player's type and the message he receives. The latter condition is used in the following definition.

Definition 1. In a Bayesian game, a correlated strategy $(\boldsymbol{m}, \sigma)$ is a *correlated equilibrium* if the corresponding random action profile $\boldsymbol{a}$ is such that, for every player i and action a_i' for that player,

$$E(\, u_i(\boldsymbol{t}, \boldsymbol{a}) - u_i(\boldsymbol{t}, (a_i', \boldsymbol{a}_{-i})) \mid \boldsymbol{t}_i, \boldsymbol{m}_i(\boldsymbol{t})\,) \geq 0. \tag{10}$$

The correlated strategy distribution of a correlated equilibrium is referred to as a *correlated equilibrium distribution* (CED). A CED is *implementable* by a mechanism $\boldsymbol{m}$ if it is the CSD of some correlated equilibrium with that mechanism. A CED is a *pure-equilibrium distribution* or a *mixed-equilibrium distribution* if it is implementable by some mechanism $\boldsymbol{m}$ with properties S, O and D or properties S, O and I, respectively, in other words, if it is the CSD of some correlated equilibrium employing a mechanism with the first or second set of properties (a correlated equilibrium of either kind is called a *Bayesian equilibrium*). Two additional kinds of CEDs are named in Section 6.3.

As the following lemma shows, to check whether a given correlated strategy distribution is a correlated equilibrium distribution it suffices to consider its default correlated strategy. The lemma also presents an equivalent characterization of CED, which does not explicitly refer to correlated equilibrium. Like Definition 1, it is formulated in terms of random variables (whose joint distribution is the CED). However, the characterization could also be expressed in purely measure theoretic terms.

Lemma 3. For a correlated strategy distribution η in a Bayesian game, the following conditions are equivalent:

(i) η is a correlated equilibrium distribution.

(ii) Its default correlated strategy is a correlated equilibrium.

(iii) For some (equivalently, every[15]) random variable $\boldsymbol{a} = (\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_n)$ such that the joint distribution of $\boldsymbol{t}$ and $\boldsymbol{a}$ is equal to η,

$$E(\, u_i(\boldsymbol{t}, \boldsymbol{a}) - u_i(\boldsymbol{t}, (a_i', \boldsymbol{a}_{-i})) \mid \boldsymbol{t}_i, \boldsymbol{a}_i\,) \geq 0, \qquad i \in N, a_i' \in A_i. \tag{11}$$

Proof. Condition (i) means that η is the CSD of some correlated strategy $(\boldsymbol{m}, \sigma)$ that satisfies the condition in Definition 1. For every player i, taking the condition expectation of both sides of (10), given $\boldsymbol{t}_i$ and $\boldsymbol{a}_i$, yields

$$E\big(\, E\big(\, u_i(\boldsymbol{t}, \boldsymbol{a}) - u_i(\boldsymbol{t}, (a_i', \boldsymbol{a}_{-i})) \,\big|\, \boldsymbol{t}_i, \boldsymbol{m}_i(\boldsymbol{t}) \,\big) \mid \boldsymbol{t}_i, \boldsymbol{a}_i\,\big) \geq 0. \tag{12}$$

[14] For a random variable $\boldsymbol{x}$ and a real-valued random variable $\boldsymbol{y}$ that are defined on the same probability space, the conditional expectation $E(\boldsymbol{y} \mid \boldsymbol{x})$ is also a random variable on that space. It is constant on every event of the form $[\boldsymbol{x} = x]$ (where $\boldsymbol{x}$ takes a particular value x), and its value there is $E(\boldsymbol{y} \mid \boldsymbol{x} = x)$, i.e., the conditional expectation of $\boldsymbol{y}$, given that $\boldsymbol{x} = x$. The meaning of equalities and inequalities involving conditional expectations is that they hold with probability 1 (equivalently, hold pointwise).

[15] The equivalence holds since whether or not the inequality in (11) holds only depends on the joint distribution of $\boldsymbol{t}$ and $\boldsymbol{a}$.

Since, by (7), $\boldsymbol{a}_i$ can be expressed as a function of $\boldsymbol{t}_i$ and $\boldsymbol{m}_i(\boldsymbol{t})$, the iterated conditional expectation in (12) is equal to the single one in (11) (see Shiryaev, 1996, Chapter I, §8). Hence, (iii) holds.

Conversely, if η satisfies (iii), then in particular (11) holds when $\boldsymbol{a}$ is the default random action profile. By (9), the inequality in this case coincides with (10), which shows that the default correlated strategy $(\boldsymbol{m}, \sigma)$ is a correlated equilibrium. This proves that (iii) implies (ii), which in turn implies (i). ∎

A probability measure η on $T \times A$ that satisfies the equivalent conditions in Lemma 3 is sometimes referred to itself as a correlated equilibrium (Bergemann and Morris, 2007). Another reasonable alternative definition of this concept would be a mechanism that recommends an action to each player, which together with the strategies of following the recommendations constitutes a correlated equilibrium in the sense of Definition 1. Lemma 3 shows that these two alternative definitions of correlated equilibrium are not fundamentally different from Definition 1. However, this paper emphatically distinguishes between correlated equilibrium and correlated equilibrium distribution, and between correlated equilibrium and the mechanism it employs. These distinctions are instrumental for the paper's primary objective of studying the implementability relation between a correlated equilibrium distribution and a mechanism, which is the existence of *some* correlated equilibrium with that mechanism that has that distribution.

4.5 Communication equilibria

Communication equilibrium differs from correlated equilibrium in that the players self-report their types to the mechanism. Correspondingly, the incentive-compatibility condition of correlated equilibrium is augmented by the requirement that a player cannot gain from being the only one to misreport his type (and possibly also deviate from the correlated strategy). The reliance on the players' reports turns the mechanism from a primary source of information about the (other) players' types to a secondary source – a mediator. The mediator may be a physical entity, such as a disinterested third party or a machine, or it may be a communication protocol, such as one-shot direct exchange of messages between players.[16] Note that in the last example, the assumption that the mechanism receives type reports as input does not exclude the exchange of more complicated messages. This is because each player could in principle use a gadget that takes type as input and outputs any required message. The players' individual gadgets could then be viewed collectively as a single mechanism, in the sense of the definition in Section 4.2. This mechanism is still rather special in that it has property S: the massage that a player receives is unaffected by his own type. In the context of communication equilibrium, this and the other fundamental properties of mechanisms described in Section 4.2 may be interpreted as specific limitations on the allowed kinds of communication protocols.

With a correlated strategy $(\boldsymbol{m}, \sigma)$, a player i who misreports his type as t_i' changes the (random) messages that the mechanism sends to the players from $\boldsymbol{m}(\boldsymbol{t})$, where $\boldsymbol{t} = (\boldsymbol{t}_i, \boldsymbol{t}_{-i})$ is the (true) random type profile, to $\boldsymbol{m}(t_i', \boldsymbol{t}_{-i})$. The player may take advantage of the resulting change of the other players' actions by appropriately changing the strategy he himself uses for determining the response to the mechanism's message, from σ_i to some $\sigma_i' : T_i \times M_i \to A_i$. The resulting random action profile $\boldsymbol{a}' = (\boldsymbol{a}_1', \boldsymbol{a}_2', \dots, \boldsymbol{a}_n')$ is given by

$$\begin{aligned} \boldsymbol{a}_i' &= \sigma_i'(\boldsymbol{t}_i, \boldsymbol{m}_i(t_i', \boldsymbol{t}_{-i})), \\ \boldsymbol{a}_j' &= \sigma_j(\boldsymbol{t}_j, \boldsymbol{m}_j(t_i', \boldsymbol{t}_{-i})), \qquad j \in N \setminus \{i\}. \end{aligned} \tag{13}$$

Communication equilibrium is defined by the requirement that, regardless of player i's true type, misreporting it in the above manner and changing the strategy cannot increase i's expected payoff.

[16] The question of how, and to what extent, can unmediated communication between players replace a mediator or a correlation device lies outside to scope of this paper. This question has been extensively studied in the contexts of both complete and incomplete information games. See, for example, Forges (1990), Ben-Porath (2003), Gerardi (2004) and references therein.

Definition 2. In a Bayesian game, a correlated strategy $(\boldsymbol{m}, \sigma)$ is a *communication equilibrium* if, for every player i, type t_i' for that player and function $\sigma_i': T_i \times M_i \to A_i$,

$$E(\, u_i(\boldsymbol{t}, \boldsymbol{a}) - u_i(\boldsymbol{t}, \boldsymbol{a}') \mid \boldsymbol{t}_i \,) \geq 0,$$

where $\boldsymbol{a}$ and $\boldsymbol{a}'$ are given by (7) and (13), respectively.

A generalization of the revelation-principle argument invoked in the first paragraph of this subsection shows that the set of possible communication equilibrium outcomes would not change if players were allowed to send to the mechanism arbitrary (rather than only type) reports, possibly affected by private randomization. For a player i of type t_i, such a general report might be described as a random variable $\boldsymbol{r}_i(t_i)$, with values in some finite (say) set R_i. A corresponding generalized mechanism would be a random variable $\widehat{\boldsymbol{m}}$, with values that are functions from $R_1 \times R_2 \times \cdots \times R_n$ to M. Thus, if the players send to the mechanism the reports $r = (r_1, r_2, \dots, r_n)$, its response is the random profile of messages $\widehat{\boldsymbol{m}}(r)$. The action a_i of each player i would then be determined as some function $\hat{\sigma}_i$ by the player's type t_i, the report r_i he sent to the mechanism and the message m_i he got in response:

$$a_i = \hat{\sigma}_i(t_i, r_i, m_i).$$

The reason this setup is in fact no more general than the one described above is that there exists a "normal" correlated strategy that induces identical actions. The mechanism $\boldsymbol{m}$ it employs simply internalizes the players' reporting process. That is, for each type profile t, the random message it sends to each player i is the pair

$$(\boldsymbol{r}_i(t_i), \widehat{\boldsymbol{m}}_i(\boldsymbol{r}_1(t_1), \boldsymbol{r}_2(t_2), \dots, \boldsymbol{r}_n(t_n))).$$

A player i of type t_i who receives a message (r_i, m_i) then determine his action according to the strategy σ_i defined by

$$\sigma_i(t_i, (r_i, m_i)) = \hat{\sigma}_i(t_i, r_i, m_i).$$

It is a simple and standard exercise to show that if with the generalized mechanism none of the players i could increase his expected payoff by changing $(\boldsymbol{r}_i(t_i))_{t_i \in T_i}$ (which specifies the report he sends) and/or $\hat{\sigma}_i$ (which specifies his response to the mechanism's messages), then the correlated strategy $(\boldsymbol{m}, \sigma)$ is a communication equilibrium in the sense of Definition 2.

A similar argument, which is also part of the revelation principle, shows that in every Bayesian game the set of communication equilibrium outcomes would not change also if the messages that the mechanism sends to the players were required to be concrete action recommendations rather than arbitrary objects. (Such a mechanism is also called a *mediation plan*. See Myerson, 1994.) More specifically, if a correlated strategy distribution is a *communication equilibrium distribution* (MED), that is, if it is the CSD of some communication equilibrium, then it is also the CSD of a communication equilibrium where (i) the message spaces coincide with the respective players' actions spaces and (ii) each player's equilibrium strategy is to act according to the message he receives. This observation shows that communication equilibrium as defined in this paper is not a fundamentally different concept than, e.g., in Myerson (1994), where (i) and (ii) are part of the definition. However, as for correlated equilibria, this does not mean that attention can be restricted to communication equilibria of this special kind. Doing so, and thus effectively dispensing with the distinction between mechanism (which is the mediator or communication protocol used) and correlated strategy (which also describes how the players react to the messages they receive), might affect in unwarranted ways the properties of the implementing mechanisms. This is demonstrated by the simple example of two players who base their choice of actions on their own type and on a type report that they receive from the other player. A mechanism that simply transmits the reports has property S, but a mechanism that is required to explicitly indicate each player's action cannot have that property.

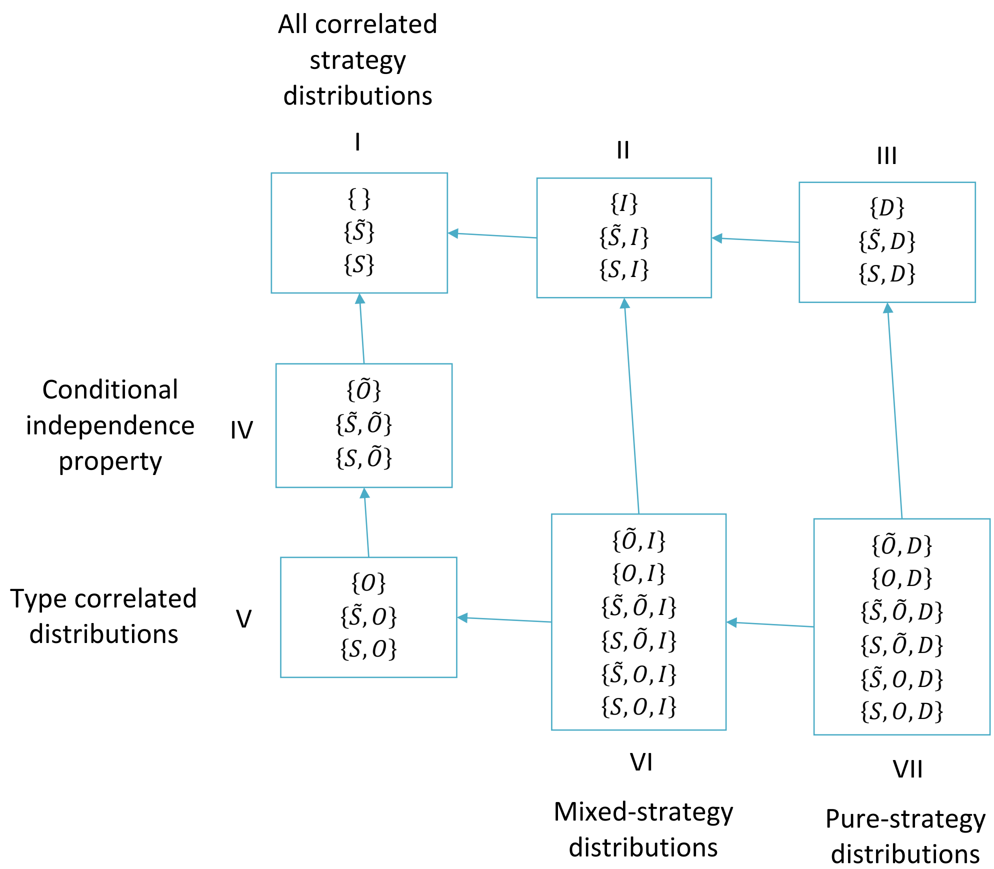

Figure 1. Hasse diagram of the different attributes of correlated strategy distributions (CSDs) in Bayesian games, ordered by implication. An attribute is represented by a box that contains its equivalent definitions, each of which is a set of fundamental properties possessed by some mechanism that implements the CSD. Two sets in the same box identify kinds of mechanisms that implement exactly the same CSDs. For those in different boxes, the implementable CSDs are different. An arrow from one box to another indicates that the first attribute implies the second one but the reverse implication does not hold.

The question of which fundamental properties a mechanism that implements a given communication equilibrium distribution can possibly have is central for this work. To put it somewhat simplistically, these properties describe how simple the implementing mechanisms can be. The classification of MEDs presented in Section 7 below may correspondingly be viewed as reflecting the extent to which the workload of the mediator or communication protocol can be reduced by off-loading some of it to the individual players – in the sense of employing a simpler mechanism but possibly more complicated strategies – without affecting the outcome.[17] This classification reflects and is partially based on similar classifications of correlated strategies and correlated equilibria, which are presented in the next two sections.

5 Attributes of Correlated Strategy Distributions

The six fundamental properties of mechanisms described in Section 4.2 are not independent. Property S implies $\tilde{S}$, property O implies $\tilde{O}$, and D implies I. Therefore, for each of the three pairs, a mechanism may satisfy both properties, only the second one, or none of them. Altogether, there are (3^3 =) 27 possibilities. This classification of mechanisms induces a classification of correlated strategy distributions. For a set of properties $\mathcal{P} \subseteq \{S, \tilde{S}, O, \tilde{O}, D, I\}$, a CSD is $\mathcal{P}$-implementable if it is implementable by some mechanism that has all the properties in $\mathcal{P}$. For example, a CSD is *S-implementable* if it is implementable by some mechanism with property S, and it is *S, O-implementable* if it is implementable by some mechanism that has both property S and property O.[18]

The various attributes of CSDs are not independent. For example, S, O-implementability implies $\tilde{S}, O$-implementability, since S is a more stringent requirement than $\tilde{S}$, and it is implied

[17] Note that this issue is quite different from that mentioned in footnote 16. In particular, the latter concerns changes of mechanism that typically make it more rather than less complex.

[18] Since implementability concerns *sets* of properties, a more accurate, though unwieldy, notation would be $\{S\}$- and $\{S, O\}$-implementability.

by S, O, I-implementability, which involves the additional requirement that the implementing mechanism also satisfies I. A natural question for each of these implications is whether the reverse implication also holds, so that the two attributes are actually equivalent. As the Hasse diagram in Figure 1 shows, the answer is affirmative for S, O- and $\tilde{S}, O$-implementability (which are equivalent) but negative for S, O- and S, O, I-implementability (which are not equivalent). Thus, every CSD that is implementable by some mechanism that satisfies $\tilde{S}$ and O is also implementable by a mechanism that satisfies S and O (and, obviously, vice versa), and there is some such CSD in some Bayesian game that is not implementable by any mechanism that also satisfies I.

As Figure 1 shows, there are not 27 but only 7 distinct (i.e., nonequivalent) attributes of CSDs that can be defined in terms of fundamental properties of the implementing mechanisms. The attributes are not all comparable. That is, some of them neither imply nor are implied by certain other attributes. Each attribute can be described in several equivalent ways by using different combinations of properties. For example, S-implementability and $\tilde{S}$-implementability are both equivalent to the attribute of simply being a CSD, which is denoted in Figure 1 by the empty set (of fundamental properties of implementing mechanisms) $\{\,\}$. Thus, the limitations that these two properties put on the implementing mechanisms are inconsequential. This result and all the others that are presented by Figure 1 are proved in Appendix A.

Additional attributes of correlated strategy distributions in Bayesian games may conceivably be defined by conjunction. For example, a CSD may be both $\tilde{O}$-implementable *and* D-implementable. A natural question is whether this attribute is equivalent to $\tilde{O}, D$-implementability. More generally, if there is some implementing mechanism with a particular set of properties and another mechanism with some other properties, does it follow that the CSD is implementable by a single mechanism that has all the properties of the other two? Lemma 5 in Appendix A answers this question in the affirmative. Hence, conjunctions do not in fact define new attributes of correlated strategy distributions. In this, the latter differ from correlated equilibrium distributions, for which conjunctions do give rise to new attributes (see Section 6).

5.1 Intrinsic characterizations

Each of the seven attributes of CSDs in Figure 1 can also be characterized *intrinsically*, that is, without explicitly referring to implementing mechanisms.[19] The significance of intrinsic characterization is that it gives an additional, concrete meaning to the attribute, and may help to check whether particular distributions have it. Lemma 2 may be viewed as an intrinsic characterization of the weakest attribute (I in Figure 1), which is simply being a CSD. The following proposition characterizes the strongest attribute, which is being a pure-strategy distribution (attribute VII), as well as the attribute of being a mixed-strategy distribution (VI). The proofs of this and the next two propositions are given in Appendix A.

Proposition 1. A correlated strategy distribution η is a mixed-strategy distribution if and only if the following holds for some (equivalently, every) random variable $\boldsymbol{a} = (\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_n)$ such that the joint distribution of $\boldsymbol{t}$ and $\boldsymbol{a}$ is equal to η:

(i) For each player i, $\boldsymbol{a}_i$ and $(\boldsymbol{t}_{-i}, \boldsymbol{a}_{-i})$ are conditionally independent, given $\boldsymbol{t}_i$.

A correlated strategy distribution is a pure-strategy distribution if and only if it satisfies the stronger condition in which (i) is replaced by:

(ii) For each player i, the conditional distribution $\boldsymbol{a}_i$, given $\boldsymbol{t}_i$, is degenerate.

An intrinsic characterization of type correlated distributions is given by the next proposition. The characterizing property in this case is the existence of a probability measure μ on $A_1^{T_1} \times A_2^{T_2} \times \cdots \times A_n^{T_n}$ that satisfies a certain condition. By definition, μ assigns a probability to each pure-

[19] Note that 'intrinsic' is not a strict, formal notion. However, the discussion below and in Appendix A should make its meaning clear.

strategy profile $(a_1^1, a_1^2, \ldots; a_2^1, a_2^2, \ldots; \ldots; a_n^1, a_n^2, \ \ldots)$ (see Section 4.1). For each type profile $t = (t_1^{j_1}, t_2^{j_2}, \ldots, t_n^{j_n})$, there is a corresponding marginal measure μ^t on $A_1 \times A_2 \times \cdots \times A_n$, which assigns to each action profile $a = (a_1, a_2, \ldots, a_n)$ the probability that the actions associated with the types specified by t are those specified by a. Formally,

$$\mu^t(\{a\}) = \mu\left(\left\{(a_1^1, a_1^2, \ldots; a_2^1, a_2^2, \ldots; \ldots; a_n^1, a_n^2, \ldots) \in A_1^{T_1} \times A_2^{T_2} \times \cdots \times A_n^{T_n} \mid (a_1^{j_1}, a_2^{j_2}, \ldots, a_n^{j_n}) = a\right\}\right).$$

Proposition 2. A correlated strategy distribution η is a type correlated distribution if and only if there is a probability measure μ on $A_1^{T_1} \times A_2^{T_2} \times \cdots \times A_n^{T_n}$ such that the marginal measure defined above coincides with the conditional probability of action profiles, that is,

$$\mu^t(\{a\}) = \frac{\eta(\{(t, a)\})}{\eta_T(\{t\})}, \qquad t \in \operatorname{supp}(\eta_T), a \in A. \tag{14}$$

The intrinsic characterization in Proposition 2 is not as straightforward as those in Proposition 1. According to the latter, pure-strategy distributions are characterized by the property that knowing a player's type entails knowing his action, and mixed-strategy distributions are characterized by the weaker property that if a player's type is known, his action does not add any information about the other players' types or actions.[20] It would seem that type correlated distributions ought to be characterized by the even weaker property that, if a player's type is known, his action does not add any information about the other players' types (but may do so for the actions). This property can be formally expressed as follows (Forges, 1993).

Definition 3. A correlated strategy distribution η has the *conditional independence property* if for some (equivalently, every) random variable $\boldsymbol{a}$ such that the joint distribution of $\boldsymbol{t}$ and $\boldsymbol{a}$ is equal to η, the action $\boldsymbol{a}_i$ of each player i and the types $\boldsymbol{t}_{-i}$ of the other players are conditionally independent, given i's own type $\boldsymbol{t}_i$.

Every type correlated distribution has the conditional independence property. This is because, if a correlated strategy distribution is implementable by a mechanism whose message to each player is unaffected by the other players' types, it is impossible to learn from the player's action anything about the others' types. However, somewhat surprisingly, the converse it not true. As Example 1 demonstrates[21] (see also Lehrer at al., 2010, and Forges, 2006), even in a two-player game a correlated strategy distribution may have the conditional independence property without being a type correlated distribution.[22] As the next proposition shows, the non-equivalence of these two attributes of CEDs reflects a difference between properties O and $\tilde{O}$ of mechanisms. Whereas being a type correlated distribution by definition means O-implementability (attribute V in Figure 1), the conditional independence property is equivalent to the weaker requirement of $\tilde{O}$-implementability (attribute IV).

Proposition 3. A correlated strategy distribution is $\tilde{O}$-implementable if and only if it has the conditional independence property.

Intrinsic characterizations for the remaining two attributes of CSDs (II and III in Figure 1) are given by Proposition 4 in Appendix A.

6 Attributes of Correlated Equilibrium Distributions

A correlated equilibrium distribution in a Bayesian game is also a correlated strategy distribution but the opposite is not always true. However, since every correlated strategy distribution can be made a correlated equilibrium distribution simply by replacing the payoff functions with constant ones, the number of distinct (i.e., nonequivalent) attributes of CEDs that can be

[20] For an extension of this result to games with a random number of players, see Milchtaich (2004, Theorem 2).

[21] Note that the demonstration can be read as a proof that a probability measure μ as in Proposition 2 does not exist for Example 1, because no strategy profile can belong to its support.

[22] Forges (1993) and Cotter's (1994) suggestion that the conditional independence property characterizes type correlated distributions is mistaken. The mistake was corrected in Forges (2006).

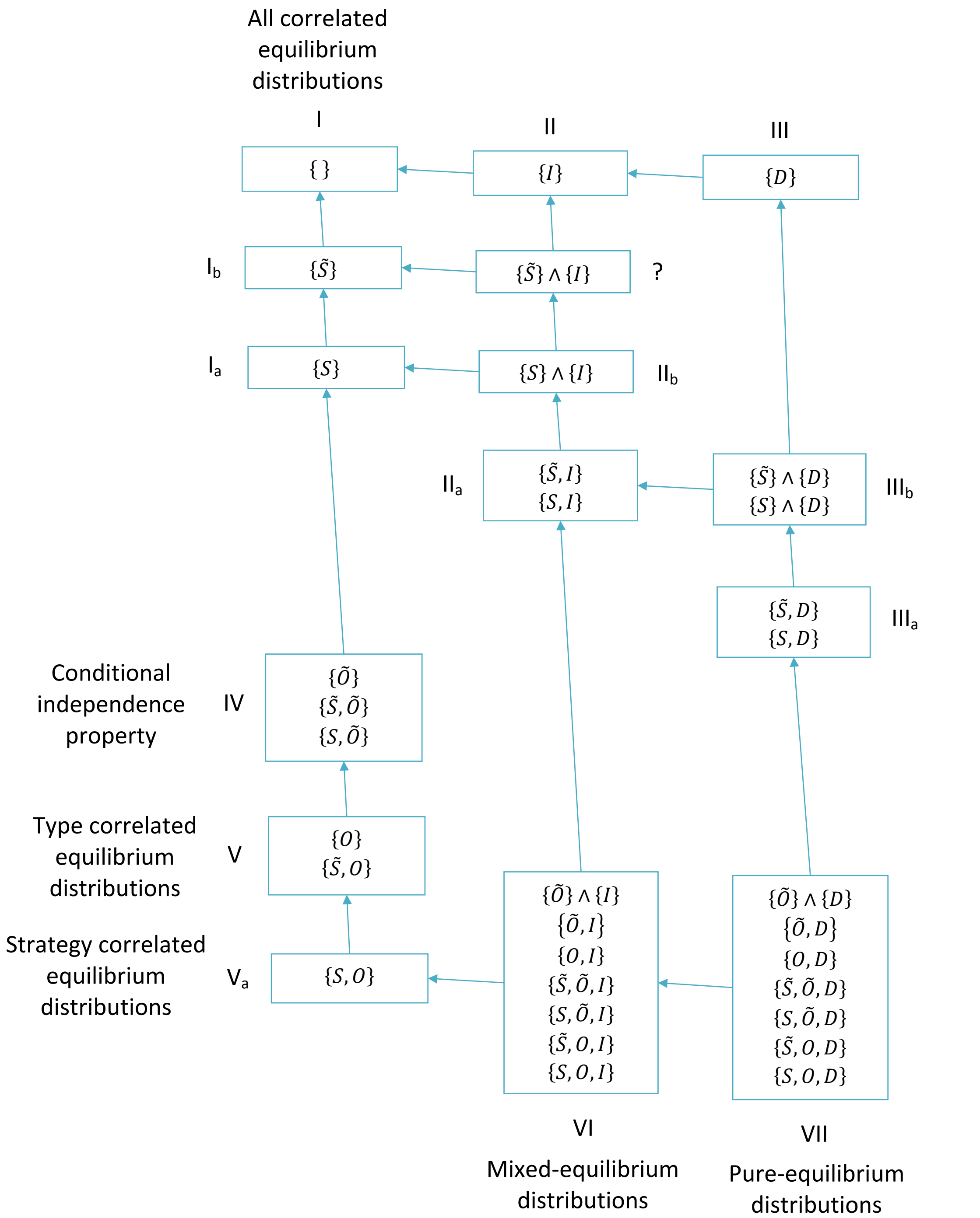

Figure 2. Hasse diagram of the different attributes of correlated equilibrium distributions (CEDs) in Bayesian games, ordered by implication. As in Figure 1, each attribute is represented by a box, and an implication relation is represented by an arrow. A conjunction symbol ∧ means that the CED is implementable both by a mechanism satisfying one property and by a mechanism satisfying the other property. The box marked with a question mark represents an attribute that may or may not be equivalent to the one below it. If the former holds, the two boxes need to be merged.

defined in terms of fundamental properties of the implementing mechanisms is not smaller than for CSDs. In fact, as Figure 2 shows, the number is significantly larger: 14 or 15 instead of 7. This reflects the fact that the classifications of CEDs can be viewed as consisting of two layers: (i) the classification induced by the different attributes of CSDs, and (ii) the refinement that results from also taking into account the incentive compatibility requirement ((iii) in Lemma 3). Thus, two CEDs in a Bayesian game may differ (i) in that even as CSDs they require different kinds of implementing mechanisms, or (ii) only in that different kinds of implementing mechanisms are compatible with the equilibrium condition. This is a useful distinction, which seems to be lacking in the existing literature on games with incomplete information.

Where the equilibrium condition is effective is the connection between a player's type and the messages he receives, i.e., properties S and $\tilde{S}$ of the mechanism. For CSDs these properties do not make any difference, as can be seen in Figure 1, but this is not so for CEDs. In particular, the three attributes of simply being a CED, S-implementability and $\tilde{S}$-implementability (I, I_a and I_b in Figure 2) are not equivalent. As Examples 2 and 3 show, there are CEDs in some Bayesian games that cannot be implemented by any mechanism satisfying $\tilde{S}$, and there are CEDs that are implementable by such a mechanism but cannot be implemented by any mechanism with the stronger property S.

The reason S and $\tilde{S}$ affect implementability of CEDs is that these properties may entail inability to restrict messages to certain types of players only. This is not a problem for correlated strategies, where information cannot do any harm, but it may be a problem for correlated equilibria, where incentive compatibility comes into play. However, as can be seen in Figure 2, whether this is actually so depends on the other properties of the mechanism. For example, for $\tilde{O}$-implementable CEDs (IV in Figure 2), requiring that the implementing mechanism also satisfies S or $\tilde{S}$ does not make any difference.

The various relations between attributes of correlated equilibrium distributions shown in Figure 2 are established in Appendix B. The next subsection takes a closer look at the connection between these attributes and those of correlated strategy distributions. The subsequent subsection presents another source of difference between the two sets of attributes. The last subsection examines in detail a couple of particular attributes.

6.1 Attributes inherited from correlated strategy distributions

As a correlated strategy distribution, a correlated equilibrium distribution has one or more of the attributes in Figure 1, which are based on the fundamental properties of the implementing mechanisms. However, a CED that, as a CSD, is implementable by a mechanism with particular properties is not necessarily implementable by such a mechanism also as a CED. That is, it may be impossible to find a correlated strategy with that kind of mechanism that is also a correlated equilibrium. For example, as indicated, the CED in Example 2 is not $\tilde{S}$-implementable, but it is $\tilde{S}$-implementable as a CSD (indeed, this is so for every CSD; see Figure 1). However, as the following theorem shows, this kind of discrepancy between the two notions of implementability may arise only when properties S or $\tilde{S}$ of mechanisms are involved. The proof of the theorem is given in Appendix B.

Theorem 1. For $\mathcal{P} \subseteq \{O, \tilde{O}, D, I\}$, a correlated equilibrium distribution is implementable by a mechanism with all the properties in $\mathcal{P}$ if and only if it satisfies a similar condition as a correlated strategy distribution.

Since the four fundamental properties considered in Theorem 1 are sufficient to characterize all the attributes of CSDs in Figure 1, it follows from the theorem that a CED has attribute I, II, III, IV, V, VI or VII in Figure 2 if and only if, as a CSD, it has the similarly numbered attribute in Figure 1. For example, a CED is a pure- or mixed-*equilibrium* distribution if and only if it is a pure- or mixed-*strategy* distribution, respectively. The intrinsic characterizations of the various attributes of CSDs given by Propositions 1 through 4 therefore apply also to the similarly numbered attributes of CEDs. For example, a CED is $\tilde{O}$-implementable (attribute IV in Figure 2) if and only if it has the conditional independence property. An intrinsic characterization for the very attribute of being a CED (I in Figure 2) is given by condition (iii) in Lemma 3, which says that a CSD is a CED if and only if it satisfies a certain incentive-compatibility condition (for distributions). It follows that, for example, an $\tilde{O}$-implementable CED can be (intrinsically) characterized as a CSD that satisfies the incentive-compatibility condition and has the conditional independence property.

6.2 Attributes defined by conjunction

Attributes of CSDs and those of CEDs show not only similarities but also differences. One such difference concerns the effect of conjunctions. As indicated in Section 5, if a CSD that is implementable by a mechanism with a particular fundamental property is also implementable by a mechanism with another property, then it is implementable by a single mechanism that has both properties. It follows from Theorem 1 that this remains true for CEDs as long as the fundamental properties concerned do not include S or $\tilde{S}$. However, Examples 6 and 7 in Appendix B show that, without the qualification, the result would not hold. Specifically, a CED that is implementable by a mechanism with property S as well as by a mechanism with property D may not be implementable by any mechanism with both properties, and the same is true with D replaced by I.

The conjunction of S- and D-implementability (as in Example 6) and the conjunction of S- and I-implementability (Example 7) are two attributes of CEDs that have no parallels among the attributes of CSDs. A third attribute that is defined in a similar manner may exist, namely, the conjunction of $\tilde{S}$- and I-implementability. However, its existence is still an open question: it is not known whether this third attribute is indeed distinct from the second one. This uncertainty is represented in Figure 2 by a question mark. In any case, by Theorem 2 below, these two or three attributes of CEDs are the only ones that can be defined (only) by conjunctions; additional such attributes do not exist.

6.3 Type correlated and strategy correlated equilibria

As an illustration of the discussion in the previous two subsections, this one describes in detail two of the attributes of correlated equilibrium distributions in Figure 2: O-implementability (attribute V), which is inherited from correlated strategy distributions, and the "spin-off" attribute S, O-implementability (V_a). Both attributes have been previously described in the literature, under various names. The O-implementable CEDs may be referred to as type correlated equilibrium distributions, as they are the distributions of type correlated equilibria (Cotter, 1994), also known as agent normal form correlated equilibria (Forges, 1993, 2006; Lehrer et al., 2006, 2010). The S, O-implementable CEDs may be referred to as strategy correlated equilibrium distributions, as they are the distributions of strategy correlated equilibria (Cotter, 1991), also known as strategic (normal) form correlated equilibria (Forges, 1993, 2006; Lehrer et al., 2010).

In a *strategy correlated equilibrium*, a referee who does not know the players' types confidentially recommends a strategy to each player. The recommendations are necessarily independent of the players' actual types but not necessarily of one another. The equilibrium condition is that it is always optimal for each player to take the action that the strategy recommended by the referee prescribes to his actual type, assuming that all the other players do the same. In the terminology of this work, in which a referee corresponds to a mechanism, the assumption that he does not know the players' types corresponds to the properties S and O, which together mean that the players' types do not affect the messages.

A *type correlated equilibrium* differs from a strategy correlated equilibrium in that each player is not told the whole strategy but only the action it prescribes to his actual type. However, it is still assumed that the referee does not know the players' types when he chooses his recommendations. Myerson (1991, p. 262) finds this assumption unnatural. Even though the referee does not know the player's type, he is able to prevent him from learning what he is not supposed to know, namely, the actions that the recommended strategy prescribes to each of the player's other possible types. However, as noted by Cotter (1994), such differentiation can be achieved if different types perceive the information presented to them differently. For example, different instructions may be issued to unilingual English and French readers simply by handing out a bilingual sheet with English and French texts that do not match. Alternatively, the mediator may simply learn the player's type after choosing his recommendation but before sending the message. Either way, the message that each player receives may depend on his own type, so that the mechanism only has property O.

Forges (1993) showed that some type correlated equilibria are not equivalent to any strategy correlated equilibrium. The next example provides another, simple demonstration of this result. Note that this example, and all the other examples in this paper that concern correlated strategy distributions or correlated equilibrium distributions, involve two-player games. This reflects (indeed, establishes) the fact that the Hasse diagrams in Figure 1 and Figure 2 apply to two-player Bayesian games as well as to the general, n-player case.

Example 5 **A correlated equilibrium distribution that is O- but not S, O-implementable**
Two players play a coordination game. They have the same action space, $\{L, R\}$, and the same payoff matrix,

$$\begin{matrix} & L & R \\ L & \multicolumn{2}{c}{\multirow{2}{*}{}} \end{matrix} \quad \begin{array}{c} \\ L \\ R \end{array}\begin{array}{c} \begin{matrix} L & R \end{matrix} \\ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \end{array}.$$

Player 1 can be of type $+1$ or type -1, which are both equally likely, and player 2 has the single type $+1$. A mechanism bases its messages to the players on the outcomes of two independent coin tosses, $\boldsymbol{m}^{+}$ and $\boldsymbol{m}^{-}$, each of which gives L or R with equal probabilities. A player of type $+1$ or -1 receives the message $\boldsymbol{m}^{+}$ or $\boldsymbol{m}^{-}$, respectively. This mechanism and the strategies that instruct the players to act according to the message the mechanism sends them together constitute a correlated equilibrium. At the equilibrium, types $+1$ and -1 of player 1 receive the expected payoffs 1 and 0.5, respectively, and player 2 gets 0.75. It is easy to check that, in all three cases, profitable deviations do not exist.

The above mechanism has properties $\tilde{S}$ and O. Player 1's type is not reflected by the message he receives (which has probability 0.5 of being L whether the type is $+1$ or -1), and has no effect whatsoever on player 2's message. Thus, the corresponding correlated equilibrium distribution is $\tilde{S}, O$-, and in particular O-, implementable. However, the CED is not implementable by any mechanism that satisfies S and O, i.e., one in which the messages to both players are unaffected by 1's type. The reason is that, in any correlated equilibrium that employs such a mechanism, the expected payoffs for the two types of player 1 must be equal. Otherwise, one of them could increase his payoff by mimicking the other type's reaction to the message he receives.

7 Attributes of Communication Equilibrium Distributions

A communication equilibrium in a Bayesian game is also a correlated equilibrium but the opposite is generally false. Whereas for correlated equilibrium the only incentive compatibility requirement is that a player cannot gain from taking a different action than that prescribed by the correlated strategy, communication equilibrium adds the requirement that reporting the type truthfully is also incentive compatible.

The second incentive compatibility requirement obviously has no bite if the mechanism ignores the players' type reports, that is, if it has properties S and O. Thus, if a correlated equilibrium distribution is implementable by a mechanism with these properties, it is automatically a communication equilibrium distribution. Among the attributes in Figure 2, S, O-implementability is in fact the weakest attribute of a CED that guarantees that it is also a MED.[23] Hence, for any given set of fundamental properties, the CEDs that are implementable by mechanisms with (all) these properties are either (i) all S, O-implementable, and hence MEDs, or (ii) not all MEDs. Significantly, in case (ii), those CEDs that *are* MEDs are not necessarily implementable as MEDs by a mechanism with the given properties. That is, a communication equilibrium with such a mechanism may not exist. This is demonstrated by Example 4, where the MED described is not D-implementable even though it has that attribute as a CED. Thus, MEDs and CEDs are not connected by a relation similar to that in Theorem 1, which concerns CEDs and CSDs.

Example 4 also shows that a result similar to that in Lemma 3 does not hold for communication equilibrium distributions: the default correlated strategy is not necessarily a communication equilibrium. The reason for this difference between correlated and communication equilibria is that, in the latter, the messages that the mechanism sends when it receives type reports that are patently not all truthful (since the profile of reported types lies outside the support of the common prior) cannot be chosen arbitrarily. The mechanism's reaction to such reports has to induce actions that ensure that the player who lied about his type (who may or may not be identifiable) is not rewarded. The feasibility of such a reaction may depend on the properties of the implementing mechanism.

Implementability by a mechanism with a particular set of fundamental properties is an attribute of MEDs, just as for CEDs and CSDs. As in Sections 5 and 6, a basic question, for each such attribute of MEDs or a conjunction of several attributes, is: which of the other attributes are implied by it? The answer is given by the Hasse diagram in Figure 3, which presents the implication relations among the various attributes of communication equilibrium distributions.

[23] This assertion is proved by Examples 4 and 5. The former presents three S, D-implementable CEDs which are not MEDs, and the latter presents an $\tilde{S}, O$-implementable CED where the otherwise identical two types of player 1 receive different payoffs, which is clearly impossible in a MED.

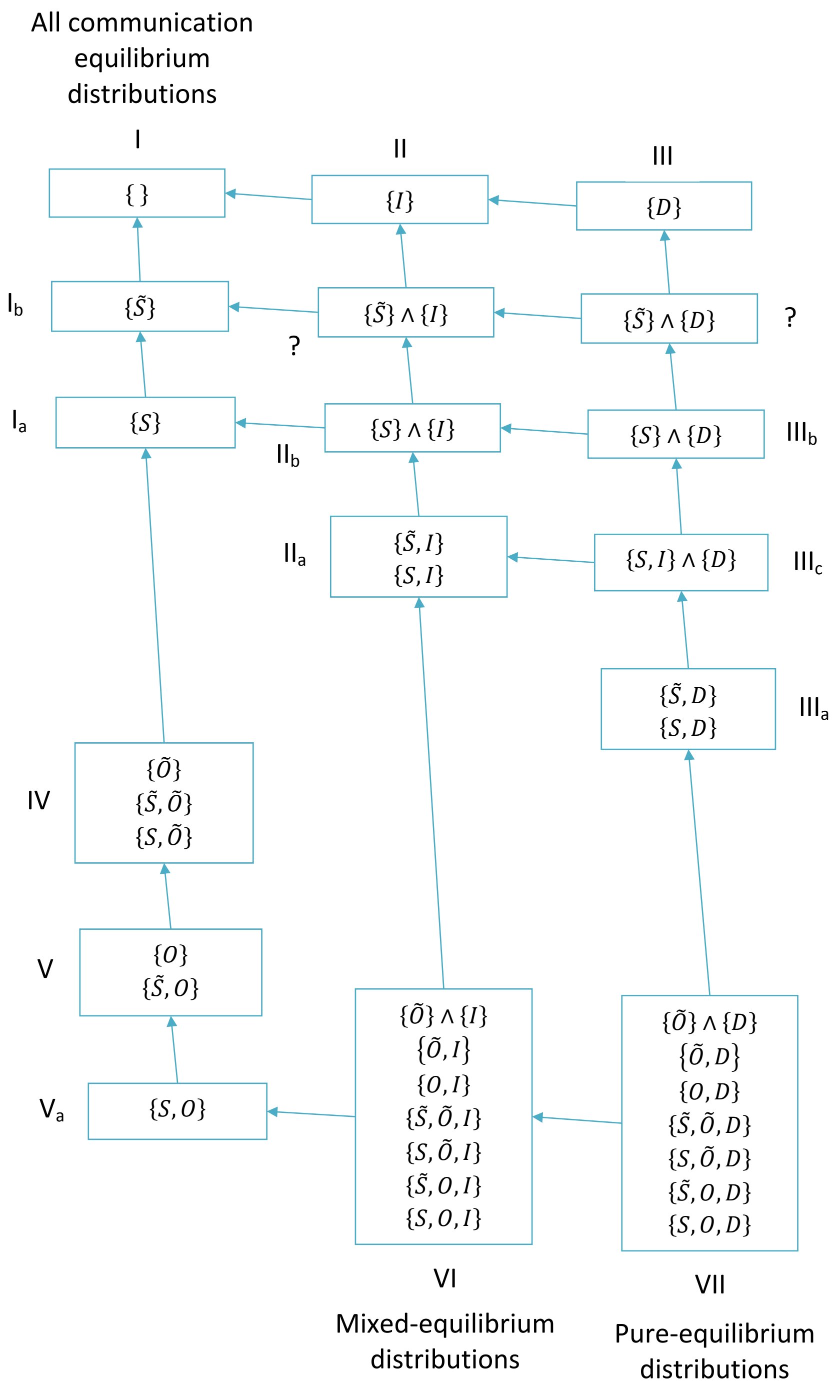

Figure 3. Hasse diagram of the different attributes of communication equilibrium distributions (MEDs) in Bayesian games, ordered by implication. Each attribute is represented by a box, and an implication relation is represented by an arrow. A conjunction symbol ∧ means that the MED is implementable both by a mechanism satisfying one property or pair of properties and by a mechanism satisfying the other property. A box marked with a question mark represents an attribute that may or may not be equivalent to the one below it. If the former holds, the two boxes need to be merged.

Comparison with the corresponding diagram for correlated equilibrium distributions (Figure 2) shows that, among the attributes that are defined by a single set of fundamental properties of mechanisms, the implication relations for MEDs and CEDs are identical. However, this is not so for attributes that are defined by conjunction. For example, for MEDs, unlike for CEDs, the conjunction of S- and D-implementability does not imply S, I-implementability. This is demonstrated by Example 8 in Appendix C, which presents an MED that is both S- and D-implementable but is not S, I-implementable. The appendix also includes the proofs all the other results that are expressed by Figure 3.

8 Synthesis

The three collections of attributes of distributions presented in the preceding three sections are complete in the sense of closedness under conjunctions. That is, each collection includes every attribute that can be defined as the conjunction of several of its elements, i.e., as the quality of possessing all of these attributes. This result is formally expressed by the following theorem, which is proved near the end of Appendix C.

Theorem 2. The conjunction of any number of the attributes of correlated strategy distributions in Figure 1, correlated equilibrium distributions in Figure 2 or communication equilibrium distributions in Figure 3 is equivalent to one of the attributes in the same figure.

Each of the three collections of attributes is a *two-dimensional lattice* with respect to the implication relation. The two-dimensionality is proved by the fact that, in each Hasse diagram, the different attributes are vertically and horizontally arranged in such a way that an attributes implies another if and only if the center of the box representing the latter is located higher and left of that of the former. The horizontal dimension corresponds to the attributes of equilibria in complete information games: pure-strategy, mixed-strategy and correlated equilibrium. The vertical dimension corresponds to the possible ways in which the players' types are connected with the messages that an implementing mechanism sends.

The assertion that the three collections of attributes are lattices means that every two attributes have a greatest lower bound (or infimum) and a least upper bound (or supremum). The greatest lower bound, also called the *meet* of the two attributes, is the unique attribute in the corresponding Hasse diagram that (i) implies both attributes and (ii) is implied by every other attribute in the diagram that implies them. The least upper bound, also called the *join* of the two attributes, is the unique attribute that (i) is implied by each of the two attributes and (ii) implies every other attribute in the diagram that is implied by each of them. The meet and join operations are customarily denoted by ∧ and ∨, respectively. This notation is consistent with the use of ∧ in Figure 2 and Figure 3 as the symbol for logical conjunction. An attribute that is defined as the conjunction of two other attributes is clearly their meet: it implies each of the two attributes and it is implied by any other attribute that does the same. Conversely, the meet of any two attributes must be equivalent to their conjunction, since by Theorem 2 an attribute equivalent to the latter appears in the diagram.

A CSD, CED or MED in a Bayesian game usually has more than one of the attributes of distributions described in this paper. In particular, if it has a certain attribute, then it also possesses every other attribute that is implied by it. For example, every type correlated distribution (attribute V in Figure 1) has the conditional independence property (attribute IV). However, it follows as a corollary from Theorem 2 that, among all the attributes in Figure 1, Figure 2 or Figure 3 that a given CSD, CED or MED has, there is always one that implies all the others; it is the distribution's *strongest* attribute. Clearly, specifying the strongest attribute is equivalent to specifying the whole collection of attributes that the distribution possesses.

Corollary. For every correlated strategy distribution η, the collection of all the attributes in Figure 1 that η possesses includes one attribute that implies all the others. The same is true for correlated equilibrium distributions and for communication equilibrium distributions, except that for them the relevant attributes are those in Figure 2 and Figure 3, respectively.

This result follows immediately from the closedness under conjunctions. By Theorem 2, the conjunction of all the attributes that a distribution η possesses is equivalent to one of the attributes in the relevant Hasse diagram. Since equivalence means two-way implication, it follows that (i) η has that attribute, and (ii) the attribute implies all the other attributes that η possesses. Parenthetically, the Corollary does not simply follow from the observation that each of the three Hasse diagrams is a lattice (or vice versa). Removing III_b, for example, from Figure 2 would invalidate the corollary, but the Hasse diagram would still be a lattice.

As an illustration of the Corollary, consider a CED with the conditional independence property that is not a pure-equilibrium distribution. Any mechanism that implements the CED must involve randomization. Otherwise, the CED's strongest attribute would imply both D- and $\tilde{O}$-implementability but not S, O, D-implementability. However, no attribute in Figure 2 has these properties.

Classification according to the strongest attribute partitions the collection of all CSDs into seven nonempty and mutually disjoint classes. The partition for CEDs, which is finer than (that inherited from) the former, has 14 or 15 elements, and for MEDs the number of classes is 15, 16

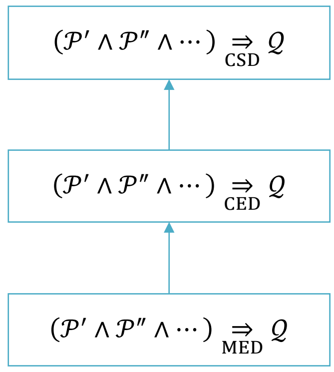

Figure 4. Hasse diagram of the implication relations between attributes of outcomes in Bayesian games: correlated strategy distributions (CSDs), correlated equilibrium distributions (CEDs), and communication equilibrium distributions (MEDs). A box represents a pair of equivalent implication relations: for all subsets $\mathcal{P}'$, $\mathcal{P}''$, … and $\mathcal{Q}$ of (the fundamental properties of mechanisms) $\{\boldsymbol{S}, \widetilde{\boldsymbol{S}}, \boldsymbol{O}, \widetilde{\boldsymbol{O}}, \boldsymbol{D}, \boldsymbol{I}\}$, one implication holds if and only if the other holds. An arrow represents only one-way implication (between implications): if the implication relations in the lower box hold, those in the higher box also hold.

or 17 (and, as proved below, is strictly greater than the previous number). Each of these classes can be designated by the same roman number and, if applicable, subscript letter as the corresponding attribute in Figure 1, Figure 2 or Figure 3. For example, class IV of CSDs consists of all the correlated strategy distributions with the conditional independence property that are not type correlated distributions.

8.1 Relations between solution concepts

The three Hasse diagrams presented in the preceding sections refer to different notions of solution concepts or outcomes in Bayesian games. Specifically, they present the implication relation between attributes of correlated strategy distributions (Figure 1), between attributes of correlated equilibrium distributions (Figure 2), and between attributes of communication equilibrium distributions (Figure 3). However, since in all cases the attributes are defined in terms of fundamental properties of implementing mechanisms, the implication relations are themselves potentially comparable. Indeed, by Propositions 11 and 22 in the appendices, each of the three relations implies or is implied by each of the others. This implication relation between implication relations is shown by the Hasse diagram in Figure 4. It follows, in particular, from this diagram and the previous ones that the number of attributes of CEDs (which is either 14 or 15, depending on the answer to the Open Question presented at the end of Appendix B) is *strictly* less than for MEDs.

8.2 Payoffs

In the context of correlated strategies, correlated equilibria or communication equilibria, an "outcome" may mean either the joint distribution of the players' types and actions or, perhaps more naturally, the corresponding players' payoffs. The analysis in the preceding sections is concerned with the former. However, as shown below, this analysis is also most relevant for the latter.

The expected payoffs of the players in a Bayesian game are completely determined by the joint distribution of their types and actions. However, the relation between distributions and payoff vectors is normally many-to-one. Hence, if a particular correlated equilibrium distribution, for example, cannot be implemented by a mechanism of a particular kind, it does not necessarily follow that the corresponding payoff vector is not thus implementable; it may be that a mechanism of that kind implements another CED with the same payoffs. Games with constant payoff functions provide a trivial example of this. In such games, a CED is implementable if and only if it is implementable as a CSD, so that the connection between distributions and the properties of the implementing mechanisms is as detailed in Section 5. By contrast, the single possible payoff vector is of course implementable by any mechanism.

A less trivial example is provided by Example 5. The original purpose of that example is to show that, in the two-player game considered, the joint distributions of type and action profiles achievable by type correlated equilibria are not identical to those achievable by strategy correlated equilibria. As shown, only the former include distributions for which the two types of player 1 receive the different payoffs 1 and 0.5. It is, however, not difficult to see that the expected payoff for player 1, which is (the mean) 0.75, is achievable also in a strategy correlated equilibrium. Thus, the example seemingly does not go as far as Forges' (1993) (considerably more involved) demonstration of the nonequivalence of type correlated and strategy correlated equilibria, which shows that even the players' payoffs, which combine those of all their types, may be different for the two kinds of equilibria. However, it follows from the next theorem that to study the effect of the properties of the implementing mechanism on the *correlated equilibrium payoffs* (CEPs), which are the n-tuples $v = (v_1, v_2, \ldots, v_n)$ specifying the players' (expected) payoffs in the correlated equilibria in an n-player Bayesian game, it is not necessary to actually examine these payoffs, as Forges (1993, 2006) did. Solving the (seemingly easier) problem of CED implementability (Section 6) may suffice. This is because any two kinds of mechanisms (of those considered in Figure 2) that do not implement the same CEDs *necessarily* also do not implement the same CEPs. A similar relation exists between correlated strategy distributions (Section 5) and *correlated strategy payoffs* (CSPs), and between communication equilibrium distributions (Section 7) and *communication equilibrium payoffs* (MEPs). Moreover, the proof of the theorem, which is given in Appendix D, is constructive, and thus provides a means of automatically transforming an example such as Example 5 into one where type correlated equilibria and strategy correlated equilibria have different payoff vectors, rather than just different joint distributions of types and actions or different payoffs for player types.

Theorem 3. For any two subsets $\mathcal{P}$ and $\mathcal{Q}$ of the six fundamental properties of mechanisms, the proposition

$$\mathcal{P}\text{-implementability implies } \mathcal{Q}\text{-implementability}$$

holds for correlated strategy payoffs, correlated equilibrium payoffs or communication equilibrium payoffs if and only if it holds for correlated strategy distributions, correlated equilibrium distributions or communication equilibrium distributions, respectively.

Note that Theorem 3 does not mean that the Hasse diagrams in Figure 1, Figure 2 and Figure 3 also apply to CSPs, CEPs and MEPs, respectively. This is because one direction of the result ("if") does not extend to the more general case in which the premise "$\mathcal{P}$-implementability" is replaced by the conjunction "$\mathcal{P}'$-implementability and $\mathcal{P}''$-implementability and …", for arbitrary list $\mathcal{P}', \mathcal{P}''$, … of subsets of the fundamental properties.

The last assertion is demonstrated by the simple game in which player 1 has two, equally probable types and only one action and player 2 has a single type and two actions. Player 2's action only affects player 1's payoff, which is 1 for one action and -1 for the other. A zero payoff for player 1 is implementable by a mechanism that reveals 1's type to player 2, who then chooses one action or the other depending on the type, or alternatively by a simple coin toss. The former has property D and the latter has property O. However, a zero payoff is not implementable by any mechanism with *both* properties, since with such a mechanism, player 2's action is necessarily always the same. Thus, unlike for correlated equilibrium distributions (see Figure 2, attribute VII), correlated equilibrium payoffs may not be O, D-implementable even if they are both O-implementable and D-implementable.

8.3 A unified framework – an outline

Correlated and communication equilibria may both be viewed as special cases of a model in which the messages that the mechanism sends to the players may depend on both their true and reported types. The dependence on the latter may be of little significance if there are no limitations on the mechanism's use of the former. However, the present setup is constructed specifically for facilitating the analysis of such limitations and their significance. Suppose, for example, that only certain aggregate data concerning the players' true types are available to the mechanism, e.g., a "checksum" of the types. The mechanism may then be able to detect unilateral deviations from truthful type reporting even if the profile of reported types lies inside the support of the common prior and is thus not a priori impossible, but may not be able to identify the player who lied about his type.

The meaning of correlated strategy in the general setting outlined above is the same as in the two special cases: it consists of a mechanism and instructions for translating the mechanism's messages into type-dependent actions for the players. A natural requirement, which generalizes both correlated and communication equilibrium, is that it is optimal for players to truthfully report their types and take the actions that are indicated by the messages they receive, if all the others do the same. The question that arises is: how do different limitations on the mechanism affect the outcomes of such generalized communication/correlated equilibria? From this perspective, the results reported in this work only concern two special kinds of limitations. In the first kind, the messages that the mechanism sends to the players are never affected by the reported types, and in the second, the true types do not affect the messages.

Appendix A Correlated Strategy Distributions

For a subset $\mathcal{P}$ of the six fundamental properties of mechanisms (see Section 4.2), a correlated strategy distribution is $\mathcal{P}$-implementable if there is some mechanism with all the properties in $\mathcal{P}$ that implements it. If $\mathcal{P}$ and $\mathcal{Q}$ are two subsets of properties, $\mathcal{P}$-implementability *implies* $\mathcal{Q}$-implementability if in every Bayesian game every $\mathcal{P}$-implementable CSD is also $\mathcal{Q}$-implementable. Shorthand for this relation is

$$\mathcal{P} \Rightarrow \mathcal{Q}.$$

A trivial sufficient condition for implication is reverse inclusion, $\mathcal{P} \supseteq \mathcal{Q}$. $\mathcal{P}$-implementability and $\mathcal{Q}$-implementability are *comparable* if the former implies the latter or vice versa, and *equivalent* if both implications hold. Shorthand for equivalence is

$$\mathcal{P} \Leftrightarrow \mathcal{Q}$$

The connection between properties of mechanisms and attributes of CSDs is extended by considering pairs of subsets of the fundamental properties. Each such pair, $\mathcal{P}$ and $\mathcal{P}'$, defines an attribute of CSDs , namely, the conjunction of $\mathcal{P}$-implementability and $\mathcal{P}'$-implementability, which is denoted by

$$\mathcal{P} \wedge \mathcal{P}'.$$

A CSD has this attribute if it is implementable both by a mechanism with the properties in $\mathcal{P}$ and by a (generally different) mechanism with the properties in $\mathcal{P}'$.[24] Lemma 5 at the end of this section shows that such a CSD is always implementable by a mechanism that has both the properties in $\mathcal{P}$ and those in $\mathcal{P}'$. It follows that conjunctions actually do not give rise to new attributes of CSDs.

If a particular implementing mechanism for a CSD does not have a certain fundamental property, it does not generally follow that the CSD lacks the corresponding attribute, for it still may be that some other implementing mechanism does have the property. An exception to this rule is presented by the following lemma. The lemma identifies several attributes of CSDs that can be determined by looking at a particular implementing mechanism, namely, the default one (see Section 4.2.1).

Lemma 4. A CSD is $\tilde{O}$-, D- or I- implementable if and only if its default mechanism has property $\tilde{O}$, D or I, respectively.

Proof. Consider a CSD η and its default mechanism $\boldsymbol{m}$. By definition, η is equal to the joint distribution of the random type profile $\boldsymbol{t}$ and the default random action profile $\boldsymbol{m}(\boldsymbol{t})$. Let $(\boldsymbol{m}', \sigma')$ be any correlated strategy other than the default one whose CSD is η, which means that η is equal to the joint distribution of $\boldsymbol{t}$ and the random action profile $\boldsymbol{a}'$ defined by

$$\boldsymbol{a}'_i = \sigma'_i(\boldsymbol{t}_i, \boldsymbol{m}'_i(\boldsymbol{t})), \qquad i \in N.$$

[24] The implication and other relations naturally extend to attributes that are defined by conjunction. The conjunction of three or more attributes is defined in the obvious way.

The two equalities together imply that

$$\boldsymbol{m}(t) \stackrel{d}{=} \left(\sigma_j'(t_j, \boldsymbol{m}_j'(t))\right)_{j\in N}, \qquad t \in \operatorname{supp}(\eta_T). \tag{15}$$

By property (5) of the default mechanism, for every $t \in T$ and i there is some t' with $(t_i, t_{-i}') \in \operatorname{supp}(\eta_T)$ such that (2) holds, which by (15) implies that

$$\boldsymbol{m}_i(t) \stackrel{d}{=} \sigma_i'(t_i, \boldsymbol{m}_i'(t_i, t_{-i}')).$$

If $\boldsymbol{m}'$ satisfies D, then the expression on the right-hand side has a degenerate distribution, which proves that $\boldsymbol{m}$ satisfies D. If $\boldsymbol{m}'$ satisfies $\tilde{O}$, then the distribution of the expression on the right-hand side is unaffected by replacing t' with any other type profile. The equality that results from this replacement proves that $\boldsymbol{m}$ also satisfies $\tilde{O}$. If $\boldsymbol{m}'$ satisfies I, then for every $t \in \operatorname{supp}(\eta_T)$ the entries on the right-hand side of the equality in (15) are independent, and therefore (3) holds. Since by property (6) of the default mechanism (3) holds also for $t \notin \operatorname{supp}(\eta_T)$, this means that $\boldsymbol{m}$ satisfies I. ∎

Unfortunately, Lemma 4 cannot be extended to all attributes of CSDs. In particular, as indicated at the end of Section 4.2.1, the default mechanism of an O-implementable CSD does not necessarily have property O.

A.1 Intrinsic characterizations

An intrinsic characterization of an attribute of CSDs presents an alternative to the attribute's original definition in terms of properties of implementing mechanisms. This subsection presents the proofs of the three propositions in Section 5.1 and adds to them intrinsic characterization for the remaining two attribute in Figure 1.

Proof of Proposition 1. To prove the sufficiency of the two conditions, consider a CSD η, with default mechanism $\boldsymbol{m}$ and default random action profile $\boldsymbol{a} = \boldsymbol{m}(\boldsymbol{t})$. Suppose that (i) or (the stronger condition) (ii) holds. For every type profile $t \in \operatorname{supp}(\eta_T)$ and action profile a,

$$\begin{aligned} &\Pr(\boldsymbol{a} = a \mid \boldsymbol{t} = t) = \Pr(\boldsymbol{a}_1 = a_1 \mid \boldsymbol{a}_{-1} = a_{-1}, \boldsymbol{t} = t)\Pr(\boldsymbol{a}_{-1} = a_{-1} \mid \boldsymbol{t} = t) \\ &\quad = \Pr(\boldsymbol{a}_1 = a_1 \mid \boldsymbol{t}_1 = t_1)\Pr(\boldsymbol{a}_{-1} = a_{-1} \mid \boldsymbol{t} = t) = \cdots \\ &\qquad\qquad = \Pr(\boldsymbol{a}_1 = a_1 \mid \boldsymbol{t}_1 = t_1)\Pr(\boldsymbol{a}_2 = a_2 \mid \boldsymbol{t}_2 = t_2)\cdots\Pr(\boldsymbol{a}_n = a_n \mid \boldsymbol{t}_n = t_n), \end{aligned} \tag{16}$$

where the second equality follows from (i) (with $i = 1$), and the subsequent equalities follow from using an identical trick for the other entries of a. Again by (i), for every player i, type t_i, action a_i, and type profile t' for which $(t_i, t_{-i}') \in \operatorname{supp}(\eta_T)$,

$$\Pr(\boldsymbol{a}_i = a_i \mid \boldsymbol{t}_i = t_i) = \Pr(\boldsymbol{a}_i = a_i \mid \boldsymbol{t} = (t_i, t_{-i}')) = \Pr(\boldsymbol{m}_i(t_i, t_{-i}') = a_i).$$

It follows from property (5) of the default mechanism that the equality between the left- and right-hand sides continues to hold even if $(t_i, t_{-i}') \notin \operatorname{supp}(\eta_T)$. This proves that, if (ii) holds, then $\boldsymbol{m}$ satisfies D. In addition, by (16), for $t \in \operatorname{supp}(\eta_T)$, $a \in A$ and any $t' \in T$,

$$\Pr(\boldsymbol{a} = a \mid \boldsymbol{t} = t) = \Pr(\boldsymbol{m}(t) = a) = \prod_{i=1}^{n} \Pr(\boldsymbol{m}_i(t_i, t_{-i}') = a_i). \tag{17}$$

Assume, without loss of generality, that the indexing of the (finite) type space $T_i = \{t_i^1, t_i^2, \dots\}$ of each player i (see Section 4.1) is such that $t^1 = (t_1^1, t_2^1, \dots, t_n^1) \in \operatorname{supp}(\eta_T)$. Define a mechanism $\overline{\boldsymbol{m}}$ by

$$\overline{\boldsymbol{m}}_i(t) = \left(\boldsymbol{m}_i(t_i^1, t_{-i}^1), \boldsymbol{m}_i(t_i^2, t_{-i}^1), \dots\right), \qquad i \in N. \tag{18}$$

Thus, the message that each player i receives from $\overline{\boldsymbol{m}}$ is a pure strategy of the form $(a_i^1, a_i^2, \dots) \in A_i^{T_i}$, where the jth action a_i^j ($j = 1,2, \dots$) coincides with the message player i would receive from the default mechanism $\boldsymbol{m}$ if his type were t_i^j and the other players' types were given by t_{-i}^1. Since this description does not involve in any way the actual, realized type profile, the mechanism $\overline{\boldsymbol{m}}$ satisfies S and O. Clearly, it also satisfies D if $\boldsymbol{m}$ does the same, which, as shown, is the case if (ii) holds. Even if only (i) holds, $\bar{\boldsymbol{m}}$ satisfies I. To prove this, it has to be shown that for any $a_1^1, a_1^2, \ldots \in A_1, a_2^1, a_2^2, \ldots \in A_2, \ldots, a_n^1, a_n^2, \ldots \in A_n$,

$$\Pr(\boldsymbol{m}_1(t_1^1, t_{-1}^1) = a_1^1, \boldsymbol{m}_1(t_1^2, t_{-1}^1) = a_1^2, \ldots; \boldsymbol{m}_2(t_2^1, t_{-2}^1) = a_2^1, \boldsymbol{m}_2(t_2^2, t_{-2}^1) = a_2^2, \ldots; \ldots) \\ = \prod_{i=1}^{n} \Pr(\boldsymbol{m}_i(t_i^1, t_{-i}^1) = a_i^1, \boldsymbol{m}_i(t_i^2, t_{-i}^1) = a_i^2, \ldots). \tag{19}$$

By (4), the right-hand is equal to

$$\prod_{i=1}^{n} \prod_{j} \Pr(\boldsymbol{m}_i(t_i^j, t_{-i}^1) = a_i^j) \tag{20}$$

and the left-hand side is equal to

$$\Pr(\boldsymbol{m}(t^1) = (a_1^1, a_2^1, \ldots, a_n^1)) \cdot \prod_{i=1}^{n} \prod_{j \neq 1} \Pr(\boldsymbol{m}_i(t_i^j, t_{-i}^1) = a_i^j).$$

By the second equality in (17), used with $t = t' = t^1$, the last expression is equal to (20). Thus, the mechanism $\bar{\boldsymbol{m}}$ satisfies I. To prove that it implements the CSD η, define a corresponding correlated strategy $(\bar{\boldsymbol{m}}, \bar{\sigma})$ by

$$\bar{\sigma}_i(t_i^j, (a_i^1, a_i^2, \ldots)) = a_i^j, \qquad i \in N, j = 1,2, \ldots. \tag{21}$$

According to $\bar{\sigma}_i$, of all the entries in the message, player i takes the one corresponding to his actual type. It has to be shown that for every $t \in \operatorname{supp}(\eta_T)$ and $a \in A$,

$$\Pr(\, \boldsymbol{a} = a \mid \boldsymbol{t} = t\,) \\ = \Pr(\, \bar{\sigma}_1(\boldsymbol{t}_1, \bar{\boldsymbol{m}}_1(\boldsymbol{t})) = a_1, \bar{\sigma}_2(\boldsymbol{t}_2, \bar{\boldsymbol{m}}_2(\boldsymbol{t})) = a_2, \ldots, \bar{\sigma}_n(\boldsymbol{t}_n, \bar{\boldsymbol{m}}_n(\boldsymbol{t})) = a_n \mid \boldsymbol{t} = t\,).$$

By (17), used with $t' = t^1$, this equality is equivalent to

$$\prod_{i=1}^{n} \Pr(\boldsymbol{m}_i(t_i, t_{-i}^1) = a_i) \\ = \Pr(\bar{\sigma}_1(t_1, \bar{\boldsymbol{m}}_1(t)) = a_1, \bar{\sigma}_2(t_2, \bar{\boldsymbol{m}}_2(t)) = a_2, \ldots, \bar{\sigma}_n(t_n, \bar{\boldsymbol{m}}_n(t)) = a_n).$$

By property I of the mechanism $\bar{\boldsymbol{m}}$, the right-hand side is equal to

$$\prod_{i=1}^{n} \Pr(\bar{\sigma}_i(t_i, \bar{\boldsymbol{m}}_i(t)) = a_i),$$

and by (18) and (21), the left-hand side is also equal to this product. Therefore, the above equality holds, which proves that $\bar{\boldsymbol{m}}$ implements η.

It remains to prove the necessity of the two conditions. For every CSD η there is a correlated strategy $(\boldsymbol{m}, \sigma)$ such that η is the joint distribution of $\boldsymbol{t}$ and the random action profile $\boldsymbol{a}$ defined by (7). By that definition,

$$\Pr(\boldsymbol{a}_i = a_i \mid \boldsymbol{t} = t) = \Pr(\sigma_i(t_i, \boldsymbol{m}_i(t)) = a_i)\,, \qquad i \in N, t \in \operatorname{supp}(\eta_T)\,, a_i \in A_i.$$

Therefore, if the mechanism $\boldsymbol{m}$ satisfies D, then the probability on the left-had side is either 0 or 1, and if $\boldsymbol{m}$ satisfies $\tilde{O}$, then

$$\Pr(\boldsymbol{a}_i = a_i \mid \boldsymbol{t} = t) = \Pr(\boldsymbol{a}_i = a_i \mid \boldsymbol{t}_i = t_i), \qquad i \in N, t \in \operatorname{supp}(\eta_T)\,, a_i \in A_i. \tag{22}$$

If $\boldsymbol{m}$ satisfies I, then for every $i \in N$, $t \in \operatorname{supp}(\eta_T)$ and $a \in A$

$$\Pr(\, \boldsymbol{a} = a \mid \boldsymbol{t} = t\,) = \prod_{j=1}^{n} \Pr\big(\sigma_j(t_j, \boldsymbol{m}_j(t)) = a_j\big) = \Pr(\, \boldsymbol{a}_i = a_i \mid \boldsymbol{t} = t\,) \Pr(\, \boldsymbol{a}_{-i} = a_{-i} \mid \boldsymbol{t} = t\,).$$

These equalities and the one in (22) together give

$$\begin{aligned}\Pr(\boldsymbol{t}_{-i} = t_{-i}, \boldsymbol{a} = a \mid \boldsymbol{t}_i = t_i) &= \Pr(\boldsymbol{a} = a \mid \boldsymbol{t} = t)\Pr(\boldsymbol{t}_{-i} = t_{-i} | \boldsymbol{t}_i = t_i) \\ &= \Pr(\boldsymbol{a}_i = a_i \mid \boldsymbol{t} = t)\Pr(\boldsymbol{a}_{-i} = a_{-i} \mid \boldsymbol{t} = t)\Pr(\boldsymbol{t}_{-i} = t_{-i} | \boldsymbol{t}_i = t_i) \\ &= \Pr(\boldsymbol{a}_i = a_i \mid \boldsymbol{t}_i = t_i)\Pr(\boldsymbol{t}_{-i} = t_{-i}, \boldsymbol{a}_{-i} = a_{-i} \mid \boldsymbol{t}_i = t_i).\end{aligned}$$

This proves that if η is both D- and $\tilde{O}$-implementable (and, a fortiori, if it is S, O, D-implementable), then (ii) holds, and if η is $\tilde{O}$- and I-implementable (and, a fortiori, if it is S, O, I-implementable), then (i) holds. ■

Proof of Proposition 2. To prove the sufficiency of the condition, suppose that a measure μ satisfying (14) exists for a CSD η. Restrict μ to its support R, and let the random variable $\boldsymbol{r}$ be the identity map on R. By construction, $\boldsymbol{r}$ is independent of the random type profile $\boldsymbol{t}$. Define a mechanism $\boldsymbol{m}$ by

$$\boldsymbol{m}_i(t) = \boldsymbol{r}, \qquad i \in N, t \in T. \tag{23}$$

This mechanism clearly has properties S and O. The message space of each player is R, each element r of which is a pure-strategy profile $(a_1^1, a_1^2, \dots; a_2^1, a_2^2, \dots; \dots; a_n^1, a_n^2, \dots)$ (where, for each i and j, a_i^j is the action prescribed to player i's jth type). Define a correlated strategy $(\boldsymbol{m}, \sigma)$ with this mechanism by

$$\sigma_i(t_i^j, r) = a_i^j, \qquad i \in N, j = 1,2, \dots. \tag{24}$$

Strategy σ_i simply instructs player i to take the action that his strategy prolife prescribes to his actual type. It has to be shown that the joint distribution of $\boldsymbol{t}$ and the random action profile $\boldsymbol{a}$ corresponding to this correlated strategy is equal to η. By (14), this equality is equivalent to

$$\Pr(\boldsymbol{a} = a \mid \boldsymbol{t} = t) = \mu^t(\{a\}), \qquad t \in \operatorname{supp}(\eta_T), a \in A. \tag{25}$$

By (7) and (23), for any type profile $t = (t_1^{j_1}, t_2^{j_2}, \dots, t_n^{j_n})$

$$\Pr(\boldsymbol{a} = a \mid \boldsymbol{t} = t) = \Pr((\sigma_1(t_1^{j_1}, \boldsymbol{r}), \sigma_2(t_2^{j_2}, \boldsymbol{r}), \dots, \sigma_n(t_n^{j_n}, \boldsymbol{r})) = a).$$

By (24) and the definition of $\boldsymbol{r}$, the right-hand side is the μ-measure of the set of all pure-strategy profiles $(a_1^1, a_1^2, \dots; a_2^1, a_2^2, \dots; \dots; a_n^1, a_n^2, \dots)$ such that $(a_1^{j_1}, a_2^{j_2}, \dots, a_n^{j_n}) = a$, which by definition is equal to $\mu^t(\{a\})$. Thus, (25) holds, so that $\boldsymbol{m}$ implements the CSD η.

To prove the necessity of the condition, consider a CSD η that is equal to the joint distribution of a pair of random variables $\boldsymbol{t}$ and $\boldsymbol{a}$ such that (7) holds for some correlated strategy $(\boldsymbol{m}, \sigma)$ with a mechanism that satisfies O. It has to be shows that there is a probability measure μ such that (14), or equivalently (25), holds. Fix a type profile t'. The random variable

$$(\sigma_1(t_1^1, \boldsymbol{m}_1(t_1^1, t'_{-1})), \sigma_1(t_1^2, \boldsymbol{m}_1(t_1^2, t'_{-1})), \dots; \sigma_2(t_2^1, \boldsymbol{m}_2(t_2^1, t'_{-2})), \sigma_2(t_2^2, \boldsymbol{m}_2(t_2^2, t'_{-2})), \dots; \dots)$$

returns values in $A_1^{T_1} \times A_2^{T_2} \times \cdots \times A_n^{T_n}$, i.e., pure-strategy profiles. Its distribution μ is given by

$$\begin{aligned}\mu(\{(a_1^1, a_1^2, \dots; a_2^1, a_2^2, \dots; \dots)\}) &= \Pr(\sigma_1(t_1^1, \boldsymbol{m}_1(t_1^1, t'_{-1})) = a_1^1, \sigma_1(t_1^2, \boldsymbol{m}_1(t_1^2, t'_{-1})) \\ &= a_1^2, \dots; \sigma_2(t_2^1, \boldsymbol{m}_2(t_2^1, t'_{-2})) = a_2^1, \sigma_2(t_2^2, \boldsymbol{m}_2(t_2^2, t'_{-2})) = a_2^2, \dots; \dots).\end{aligned}$$

For every type profile $t = (t_1^{j_1}, t_2^{j_2}, \dots, t_n^{j_n}) \in \operatorname{supp}(\eta_T)$ and action profile $a = (a_1, a_2, \dots a_n)$,

$$\begin{aligned}&\mu\left(\left\{ (a_1^1, a_1^2, \dots; a_2^1, a_2^2, \dots; \dots; a_n^1, a_n^2, \dots) \in A_1^{T_1} \times A_2^{T_2} \times \cdots \times A_n^{T_n} \;\middle|\; (a_1^{j_1}, a_2^{j_2}, \dots, a_n^{j_n}) = a \right\}\right) \\ &= \Pr(\sigma_1(t_1^{j_1}, \boldsymbol{m}_1(t_1^{j_1}, t'_{-1})) = a_1, \sigma_2(t_2^{j_2}, \boldsymbol{m}_2(t_2^{j_2}, t'_{-2})) = a_2, \dots) = \Pr(\boldsymbol{a} = a \mid \boldsymbol{t} = t),\end{aligned}$$

where the second equality uses (7) and the assumption that $\boldsymbol{m}$ has property O. Thus, (25) holds, as had to be shown. ■

Proof of Proposition 3. In view of Lemma 4, it suffices to show that the default mechanism $\boldsymbol{m}$ of a CSD η has property $\tilde{O}$ if and only if the condition that defines the conditional independence property (Definition 3) holds for the default random action profile $\boldsymbol{a} = \boldsymbol{m}(\boldsymbol{t})$. That condition can be presented as follows: for every player i and type t_i for that player and all type profiles t' and

t'' with $(t_i, t'_{-i}), (t_i, t''_{-i}) \in \operatorname{supp}(\eta_T)$,

$$\Pr(\boldsymbol{a}_i = a_i \mid \boldsymbol{t} = (t_i, t'_{-i})) = \Pr(\boldsymbol{a}_i = a_i \mid \boldsymbol{t} = (t_i, t''_{-i})), \qquad a_i \in A_i. \tag{26}$$

Since $\boldsymbol{a} = \boldsymbol{m}(\boldsymbol{t})$, (26) is equivalent to

$$\boldsymbol{m}_i(t_i, t'_{-i}) \stackrel{d}{=} \boldsymbol{m}_i(t_i, t''_{-i}). \tag{27}$$

It follows from property (5) of the default mechanism that (27) holds for all type profiles t' and t'' with $(t_i, t'_{-i}), (t_i, t''_{-i}) \in \operatorname{supp}(\eta_T)$ if and only if it holds for *all* $t', t'' \in T$. This is so for every player i and type t_i if and only if $\boldsymbol{m}$ has property $\tilde{O}$. ∎

Proposition 4. A CSD η is I-implementable if and only if the following holds for some (equivalently, every) random variable $\boldsymbol{a} = (\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_n)$ such that the joint distribution of $\boldsymbol{t}$ and $\boldsymbol{a}$ is equal to η:

(i) $\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_n$ are conditionally independent, given $\boldsymbol{t}$.

A CSD is D-implementable if and only if it satisfies the stronger condition in which (i) is replaced by:

(ii) The conditional distribution of $\boldsymbol{a}$ given $\boldsymbol{t}$ is degenerate.

Proof. In view of Lemma 4, it suffices to show that the default mechanism $\boldsymbol{m}$ of η has property I or D if and only if (i) or (ii), respectively, holds for the default random action profile $\boldsymbol{a} = \boldsymbol{m}(\boldsymbol{t})$. In other words, the condition that $\boldsymbol{m}(t)$ satisfies (3) or the condition that its distribution is degenerate holds for all type profiles t if and only if the condition holds for $t \in \operatorname{supp}(\eta_T)$. These equivalences are implied by properties (6) and (5), respectively, of the default mechanism. ∎

A.2 Equivalences

This subsection identifies a number of equivalent formulations for various attributes of CSDs. The first result establishes the irrelevance in the present context of the properties of mechanisms that concern the connections between the messages they send to a player and the player's own type, namely, properties S and $\tilde{S}$.

Proposition 5. For CSDs, $\{S\} \Leftrightarrow \{\tilde{S}\} \Leftrightarrow \{\,\}$, $\{S, \tilde{O}\} \Leftrightarrow \{\tilde{S}, \tilde{O}\} \Leftrightarrow \{\tilde{O}\}$, $\{S, O\} \Leftrightarrow \{\tilde{S}, O\} \Leftrightarrow \{O\}$, $\{S, I\} \Leftrightarrow \{\tilde{S}, I\} \Leftrightarrow \{I\}$ and $\{S, D\} \Leftrightarrow \{\tilde{S}, D\} \Leftrightarrow \{D\}$.

Proof. Property S of mechanisms implies $\tilde{S}$, and therefore $\{S\} \Rightarrow \{\tilde{S}\} \Rightarrow \{\,\}$. Hence, to prove that the three attributes are equivalent it suffices to show that every CSD is S-implementable.

Consider the default mechanism $\boldsymbol{m}$ of a CSD η. That mechanism does not necessarily have property S (see Section 4.2.1). A mechanism $\vec{\boldsymbol{m}}$ that does have that property is defined by

$$\vec{\boldsymbol{m}}_i(t) = \left(\boldsymbol{m}_i(t_i^1, t_{-i}), \boldsymbol{m}_i(t_i^2, t_{-i}), \dots\right), \qquad i \in N. \tag{28}$$

This mechanism is similar to that defined in (18) in that the message that each player receives is a pure strategy, but differs from it in that the partial type profile on the right-hand side is t_{-i} rather than the constant one t_{-i}^1. Therefore, the mechanism $\vec{\boldsymbol{m}}$ satisfies S but not O. To prove that it implements η, it only needs to be noted that the random action profile of the correlated strategy $(\vec{\boldsymbol{m}}, \bar{\sigma})$, with $\bar{\sigma}$ defined by (21), coincides with the default one:

$$\bar{\sigma}_i(\boldsymbol{t}_i, \vec{\boldsymbol{m}}_i(\boldsymbol{t})) = \boldsymbol{m}_i(\boldsymbol{t}), \qquad i \in N. \tag{29}$$

The proofs that $\{S, \tilde{O}\} \Leftrightarrow \{\tilde{S}, \tilde{O}\} \Leftrightarrow \{\tilde{O}\}$, $\{S, I\} \Leftrightarrow \{\tilde{S}, I\} \Leftrightarrow \{I\}$ and $\{S, D\} \Leftrightarrow \{\tilde{S}, D\} \Leftrightarrow \{D\}$ are very similar, and only require the following additions to the above proof.

If the CSD η is $\tilde{O}$-, I- or D-implementable, then by Lemma 4 the default mechanism $\boldsymbol{m}$ has property $\tilde{O}$, I or D, respectively. It has to be shown that the mechanism $\vec{\boldsymbol{m}}$ also has the same property, in addition to S. When $\boldsymbol{m}$ satisfies $\tilde{O}$, for every player i and type profiles t' and t'' the equality (27) holds for all types t_i, which by (4) implies that

$$\left(\boldsymbol{m}_i(t_i^1, t'_{-i}), \boldsymbol{m}_i(t_i^2, t'_{-i}), \dots\right) \stackrel{d}{=} \left(\boldsymbol{m}_i(t_i^1, t''_{-i}), \boldsymbol{m}_i(t_i^2, t''_{-i}), \dots\right).$$

Thus, $\vec{\boldsymbol{m}}$ has property $\tilde{O}$. When $\boldsymbol{m}$ satisfies I or D, (4) implies that the random variables

$$\{\boldsymbol{m}_j(t)\}_{j\in N, t\in T} \tag{30}$$

are independent or have degenerate distributions, respectively. It follows that the same is true, for every type profile t, for the n random variables

$$\{\vec{\boldsymbol{m}}_j(t)\}_{j\in N},$$

each of which is a vector whose entries are a subset of the random variables in (30), such that these n subsets are disjoint. Thus, the mechanism $\vec{\boldsymbol{m}}$ satisfies I or D, respectively.

To prove that $\{S, O\} \Leftrightarrow \{\tilde{S}, O\} \Leftrightarrow \{O\}$, it suffices to show that $\{O\} \Rightarrow \{S, O\}$. This is shown in the proof of Proposition 2, where it is proved that the existence of a measure μ as in that proposition, which is implied by O-implementability, in turn implies S, O-implementability. ■

Proposition 6. For CSDs, $\{S, O, I\} \Leftrightarrow \{\tilde{S}, O, I\} \Leftrightarrow \{S, \tilde{O}, I\} \Leftrightarrow \{\tilde{S}, \tilde{O}, I\} \Leftrightarrow \{O, I\} \Leftrightarrow \{\tilde{O}, I\} \Leftrightarrow \left(\{\tilde{O}\} \wedge \{I\}\right)$ and $\{S, O, D\} \Leftrightarrow \{\tilde{S}, O, D\} \Leftrightarrow \{S, \tilde{O}, D\} \Leftrightarrow \{\tilde{S}, \tilde{O}, D\} \Leftrightarrow \{O, D\} \Leftrightarrow \{\tilde{O}, D\} \Leftrightarrow \left(\{\tilde{O}\} \wedge \{D\}\right)$.

Proof. It clearly suffices to show that $\left(\{\tilde{O}\} \wedge \{I\}\right) \Rightarrow \{S, O, I\}$ and $\left(\{\tilde{O}\} \wedge \{D\}\right) \Rightarrow \{S, O, D\}$. As shown in the last part of the proof of Proposition 1, every CSD η that is both $\tilde{O}$- and I-implementable, or both $\tilde{O}$- and D-implementable, is the joint distribution of $\boldsymbol{t}$ and some random variable $\boldsymbol{a}$ that satisfies condition (i) or (ii), respectively, in Proposition 1. Therefore, by that proposition, in the first case η is also S, O, I-implementable, and in the second, it is S, O, D-implementable. ■

A.3 Implications

Propositions 5 and 6 identify seven attributes of correlated strategy distributions that are defined (in several equivalent ways) by subsets of the six fundamental properties of mechanisms. Figure 1 shows these attributes as well as certain trivial implications between them, which all follow immediately from relations between properties of mechanisms. To prove that the figure presents a complete picture of the implication relations between attributes of CSDs, it remains to prove that implications additional to those shown do not hold, so that, in particular, none of the seven attributes is equivalent to another. For this, the following four propositions are required.

Proposition 7. For CSDs, $\{S, O, I\} \nRightarrow \{D\}$.

Proof. It suffices to consider any complete information game (that is, a game where every player has only one type) with a mixed-strategy profile that is not pure. ■

Proposition 8. For CSDs, $\{S, D\} \nRightarrow \{\tilde{O}\}$.

Proof. In a two-player Bayesian game in which player 1 has a single type and two actions and player 2 has a single action and two types, consider a correlated strategy distribution in which player 1 takes his first or second action if player 2 is of the first or second type, respectively. This CSD is implementable by mechanism that simply tells player 1 the type of player 2, and thus satisfies S and D. However, the CSD is not $\tilde{O}$-implementable, since a mechanism with property $\tilde{O}$ cannot possibly provide player 1 with any information about player 2's type. ■

Proposition 9. For CSDs, $\{\tilde{O}\} \nRightarrow \{O\}$.

Proof. By Example 1, there exists a CSD that has the conditional independence property but is not O-implementable. By Proposition 3, that CSD is $\tilde{O}$-implementable. ■

Proposition 10. For CSDs, $\{S, O\} \nRightarrow \{I\}$.

Proof. In a complete information game, properties S and O automatically hold for every mechanism, but a CSD is I-implementable only if the players' actions are independent. ■

Proposition 7 proves that attribute VI only implies the other attributes in Figure 1 that the diagram indicates it implies (in other words, it does not imply III or VII). Proposition 8 proves the same for attribute III. These two results prove that attribute II (which is implied by both III and VI) only implies attribute I, and therefore the latter does not imply IV. Proposition 9 proves that IV does not imply V. Proposition 10 proves that attribute V only implies the (two) attributes that the diagram indicates it implies, which establishes the same for attribute IV and for attribute I.

Since, for mechanisms, property S implies $\tilde{S}$, property O implies $\tilde{O}$, and D implies I, there are only 27 relevant subsets of $\{S, \tilde{S}, O, \tilde{O}, D, I\}$, which all appear in Figure 1. Therefore, there are no additional attributes of CSDs that can be described by single subsets of the six fundamental properties of mechanisms. The following lemma shows that the same is true for pairs (hence also triplets, etc.) of sets of properties of mechanisms: no additional attributes of CSDs can be defined by them. This is because, if a CSD is implementable both by a mechanism with one set of properties and by a mechanism with a second set of properties, then it is implementable by a single mechanism that has all the properties of the other two.

Lemma 5. For CSDs, for every two subsets $\mathcal{P}, \mathcal{Q} \subseteq \{S, \tilde{S}, O, \tilde{O}, D, I\}$,

$$(\mathcal{P} \wedge \mathcal{Q}) \Leftrightarrow \mathcal{P} \cup \mathcal{Q}. \tag{31}$$

Proof (an outline). Proposition 6 proves the two cases of (31) in which $\mathcal{P} = \{\tilde{O}\}$ and $\mathcal{Q}$ is $\{I\}$ or $\{D\}$. By inspection of Figure 1, every other case follows from one of these two. ■

It is shown in Section B.3 below that for correlated equilibrium distributions a similar result to Lemma 5 does not hold. In other words, the requirement of incentive compatibility may invalidate the equivalence (31).

Appendix B Correlated Equilibrium Distributions

The analysis of correlated strategy distributions in the previous appendix is a first step in the analysis of correlated equilibrium distributions. The former concerns qualitative differences between distributions that reflect the limitations of the implementing mechanisms. The latter also incorporates the constraints inherent in the incentive compatibility requirement. Whereas in the case of CSDs the limiting factor is the mechanism's ability to transmit information to players (about the other players' types and outside random events), in the case of CEDs its ability to do so selectively also comes into play.

As for CSDs, each subset $\mathcal{P}$ of the six fundamental properties of mechanisms defines an attribute of correlated strategy distributions, namely, $\mathcal{P}$-implementability. A CED has this attribute if there is some mechanism with all the properties in $\mathcal{P}$ that implements it. However, in the present context, implementability has a different meaning than for CSDs. Namely, the correlated strategy involved is required to be a correlated equilibrium. Thus, an expression like $\mathcal{P} \Rightarrow \mathcal{Q}$ has a different meaning for CSDs and CEDs. Wherever confusion is possible, the generic implication sign may be replaced by the more explicit one $\underset{\text{CSD}}{\Rightarrow}$ or $\underset{\text{CED}}{\Rightarrow}$. As the following shows, the second relation is in a sense stronger than the first one.

Proposition 11. For every two subsets $\mathcal{P}, \mathcal{Q} \subseteq \{S, \tilde{S}, O, \tilde{O}, D, I\}$,

$$\mathcal{P} \underset{\text{CED}}{\Rightarrow} \mathcal{Q} \text{ implies } \mathcal{P} \underset{\text{CSD}}{\Rightarrow} \mathcal{Q}. \tag{32}$$

The same is moreover true with $\mathcal{P}$ replaced by by $\mathcal{P}' \wedge \mathcal{P}'' \wedge \cdots$, for any list $\mathcal{P}', \mathcal{P}''$, … of subsets of $\{S, \tilde{S}, O, \tilde{O}, V, I\}$.

Proof. It has to be shown that (i) $\mathcal{P} \underset{\text{CED}}{\Rightarrow} \mathcal{Q}$ and (ii) $\mathcal{P} \underset{\text{CSD}}{\not\Rightarrow} \mathcal{Q}$ are contradictory. Condition (i) means that, in every Bayesian game, every $\mathcal{P}$-implementable CED is also $\mathcal{Q}$-implementable. Condition (ii) means that there is some CSD in some Bayesian game that is $\mathcal{P}$- but not $\mathcal{Q}$-implementable.

Without loss of generality, the payoff functions in that game (which are irrelevant for CSD implementability) are identically zero. However, this means that every correlated strategy in the game is a correlated equilibrium and vice versa, which contradicts (i).

The same argument applies virtually unchanged also to the more general version in which $\mathcal{P}$ is replaced by $\mathcal{P}' \wedge \mathcal{P}'' \wedge \cdots$. ■

The converse of (32) does not generally hold. Consequently, not all attributes of CEDs that can be described in terms of the six fundamental properties of mechanisms correspond to attributes of CSDs. In other words, the former are not simply the restrictions of the latter to correlated equilibrium distributions. Rather, restriction is followed by refinement, which gives rise to additional attributes.

Some of the attributes of CEDs, including the majority of those inherited from CSDs, are presented in the following subsection. The subsequent subsection describes additional attributes, by specifically identifying instances in which the converse of (32) does not hold. The last subsection completes the description of the implication relation $\underset{\text{CED}}{\Rightarrow}$ (henceforth written simply as $\Rightarrow$) by considering implications involving conjunctions of attributes of CEDs.

B.1 Equivalences

The following propositions identify equivalent formulations for several attributes of CEDs.

Proposition 12. For CEDs, $\{S, \tilde{O}\} \Leftrightarrow \{\tilde{S}, \tilde{O}\} \Leftrightarrow \{\tilde{O}\}$.

Proof. In view of Lemma 4, it suffices to show that if the default mechanism $\boldsymbol{m}$ of a CED η has property $\tilde{O}$, then the correlated strategy $(\vec{\boldsymbol{m}}, \bar{\sigma})$ defined in the proof of Proposition 5 is a correlated equilibrium. As shown in that proof, if the default mechanism satisfies $\tilde{O}$ (or I or D), then $\vec{\boldsymbol{m}}$ satisfies both S and $\tilde{O}$ (or I or D, respectively).

By Lemma 3, the default random action profile $\boldsymbol{a} = \boldsymbol{m}(\boldsymbol{t})$ satisfies (11). By (29) and the definition of correlated equilibrium (Definition 1), $(\vec{\boldsymbol{m}}, \bar{\sigma})$ is a correlated equilibrium if and only if

$$E(u_i(\boldsymbol{t}, \boldsymbol{a}) - u_i(\boldsymbol{t}, (a_i', \boldsymbol{a}_{-i})) \mid \boldsymbol{t}_i, \vec{\boldsymbol{m}}_i(\boldsymbol{t})) \geq 0, \qquad i \in N, a_i' \in A_i. \tag{33}$$

It follows that the latter condition holds if the conditional expectations in (11) and (33) are equal. The formal difference between the former and the latter is that player i's action $\boldsymbol{a}_i$, which coincides with the message $\boldsymbol{m}_i(\boldsymbol{t})$ he receives from the default mechanism, is replaced by $\vec{\boldsymbol{m}}_i(\boldsymbol{t})$, which by (28) also specifies the messages the player would receive if his type were different. Therefore, the meaning of the above equality is that these messages do not provide player i with any information that he could use for choosing a better action.

Since $\boldsymbol{a} = \boldsymbol{m}(\boldsymbol{t})$, if the conditional expectations in (11) and (33) are not equal, then by (28) there is some type of player i, say the first one t_i^1, and some messages $m_i^1, m_i^2, \ldots$ such that

$$E(u_i(\boldsymbol{t}, \boldsymbol{m}(t_i^1, \boldsymbol{t}_{-i})) - u_i(\boldsymbol{t}, (a_i', \boldsymbol{m}_{-i}(t_i^1, \boldsymbol{t}_{-i}))) \mid \boldsymbol{t}_i = t_i^1, \boldsymbol{m}_i(t_i^1, \boldsymbol{t}_{-i}) = m_i^1)$$
$$\neq E(u_i(\boldsymbol{t}, \boldsymbol{m}(t_i^1, \boldsymbol{t}_{-i})) - u_i(\boldsymbol{t}, (a_i', \boldsymbol{m}_{-i}(t_i^1, \boldsymbol{t}_{-i}))) \mid \boldsymbol{t}_i = t_i^1, \boldsymbol{m}_i(t_i^1, \boldsymbol{t}_{-i}) = m_i^1, \boldsymbol{m}_i(t_i^2, \boldsymbol{t}_{-i}) = m_i^2, \ldots).$$

The inequality implies that the pair of random variables $\boldsymbol{t}$ and $\boldsymbol{m}(t_i^1, \boldsymbol{t}_{-i})$ is not independent of $\boldsymbol{m}_i(t_i^2, \boldsymbol{t}_{-i}), \boldsymbol{m}_i(t_i^3, \boldsymbol{t}_{-i}), \ldots$. However, if the default mechanism $\boldsymbol{m}$ has property $\tilde{O}$, then it follows from (4) that such independence does in fact hold, so that the above inequality cannot hold, which proves that $(\vec{\boldsymbol{m}}, \bar{\sigma})$ is a correlated equilibrium. ■

Proposition 13. For CEDs, $\{\tilde{S}, O\} \Leftrightarrow \{O\}$.

Proof. Let $(\boldsymbol{m}, \sigma)$ be a correlated equilibrium with a mechanism that satisfies O. It has to be shown that there is some correlated equilibrium $(\widehat{\boldsymbol{m}}, \hat{\sigma})$ with a mechanism that satisfies $\tilde{S}$ and O such that the two correlated equilibria have identical CEDs.

The construction of $(\widehat{\boldsymbol{m}}, \hat{\sigma})$ is based on the idea of encoding the messages that the mechanism $\boldsymbol{m}$ sends to the players in a particular manner. Suppose, without loss of generality, that these messages are integers, more specifically, that the message space M_i of each player i consists of all integers between 1 and some number K_i. Since property O of the mechanism implies $\tilde{O}$, the (random) message $\boldsymbol{m}_i(t)$ that player i receives has a distribution function F_{i,t_i} that only depends on the player's own type t_i. That is, for any t_{-i},

$$F_{i,t_i}(s) = \Pr(\boldsymbol{m}_i(t_i, t_{-i}) \le s), \qquad -\infty < s < \infty.$$

The mechanism $\widehat{\boldsymbol{m}}$ combines the message $\boldsymbol{m}_i(t)$ with a random variable $\boldsymbol{r}$ that is uniformly distributed on the half-open interval $(0,1]$ and is independent of $\boldsymbol{m}$. (The assumption of uniform distribution is actually inconsistent with the definition of random variable in footnote 4, which requires a finite probability space. However, this assumption is only temporary; it is removed in the last part of the proof.) Specifically, for every player i and type profile t,

$$\widehat{\boldsymbol{m}}_i(t) = \boldsymbol{r}\, F_{i,t_i}(\boldsymbol{m}_i(t)) + (1-\boldsymbol{r})F_{i,t_i}(\boldsymbol{m}_i(t) - 1). \tag{34}$$

It is not difficult to see that $\widehat{\boldsymbol{m}}_i(t)$ is uniformly distributed on the unit interval. Therefore, the mechanism $\widehat{\boldsymbol{m}}$ satisfies $\tilde{S}$ as well as O.

The next step is to define a strategy $\hat{\sigma}_i$ for each player i by $\hat{\sigma}_i(t_i, \widehat{m}_i) = \sigma_i(t_i, \psi_i(t_i, \widehat{m}_i))$, where ψ_i is a function that "decodes" the message $\widehat{\boldsymbol{m}}_i(t)$ and recovers the original message $\boldsymbol{m}_i(t)$:

$$\psi_i(t_i, x) = \min\{m_i \in M_i \mid F_{i,t_i}(m_i) \ge x\}.$$

By virtue of this decoding, $\hat{\sigma}_i$ always specifies the same action as σ_i. Since, in addition, the messages that the players receive from the mechanism $\widehat{\boldsymbol{m}}$ convey precisely the same information about the other players' types and actions as those from $\boldsymbol{m}$, this proves that $(\widehat{\boldsymbol{m}}, \hat{\sigma})$, like $(\boldsymbol{m}, \sigma)$, is a correlated equilibrium.

The above construction does not strictly conforms to the definition of mechanism since the message spaces in $\widehat{\boldsymbol{m}}$ are infinite. This problem may be overcome by replacing the uniformly-distributed random variable $\boldsymbol{r}$ with one that has only finitely many possible values, specifically, a random variable of the form $\mu(\boldsymbol{r})$, where μ is a real-valued function with a finite range. The first step is to consider the (finite) set

$$X = \{F_{i,t_i}(m_i)\}_{i,t_i,m_i}$$

of all values that may appear in the first term in (34). The next step is to modify the definition of the mechanism $\widehat{\boldsymbol{m}}$ by changing the message that it sends to each player i from $\widehat{\boldsymbol{m}}_i(t)$ (which is defined by (34)) to $\lambda(\widehat{\boldsymbol{m}}_i(t))$, where $\lambda\colon [0,1] \to [0,1]$ is the non-decreasing left continuous function defined by $\lambda(x) = \min(\{x' \in X \mid x' \ge x\})$. This change is inconsequential. Since always $\widehat{\boldsymbol{m}}_i(t) \le \lambda(\widehat{\boldsymbol{m}}_i(t)) \le F_{i,t_i}(\boldsymbol{m}_i(t))$, applying the decoder ψ_i to the modified message $\lambda(\widehat{\boldsymbol{m}}_i(t))$ still recovers $\boldsymbol{m}_i(t)$. Let the function $\mu\colon (0,1] \to (0,1]$ be defined by

$$\mu(r) = \max\left\{ 0 < r' \le 1 \;\middle|\; \begin{array}{c} \lambda(r'\, F_{i,t_i}(m_i) + (1-r')F_{i,t_i}(m_i - 1)) \\ = \lambda(r\, F_{i,t_i}(m_i) + (1-r)F_{i,t_i}(m_i - 1)) \\ \text{for all } i, t_i, m_i \end{array} \right\}.$$

It is not difficult to see that the function μ is well defined and, as required, has only finitely many possible values. The final step is to replace $\boldsymbol{r}$ in (34) with $\mu(\boldsymbol{r})$. By definition of μ, this replacement does not change the message $\lambda(\widehat{\boldsymbol{m}}_i(t))$. ■

Proposition 14. For CEDs, $\{S, I\} \Leftrightarrow \{\tilde{S}, I\}$ and $\{S, D\} \Leftrightarrow \{\tilde{S}, D\}$.

Proof. To prove that $\{\tilde{S}, I\} \Rightarrow \{S, I\}$, it has to be shown that every CED implementable by a mechanism $\boldsymbol{m}$ with properties $\tilde{S}$ and I is also implementable by a mechanism with properties S and I.

Property $\tilde{S}$ of $\boldsymbol{m}$ means that for every type profile t and player i the distribution of $\boldsymbol{m}_i(t)$ does not change when only player i's type t_i changes. In other words, the distribution only depends on i and on the partial type profile t_{-i}. Therefore, it is possible to construct a family $\{\boldsymbol{r}^{i,t_{-i}}\}_{i,t_{-i}}$

of independent random variables, indexed by the players and partial type profiles, such that each entry $\boldsymbol{r}^{i,t_{-i}}$ has the distribution described above. Define a mechanism $\widetilde{\boldsymbol{m}}$ by

$$\widetilde{\boldsymbol{m}}(t) = (\boldsymbol{r}^{1,t_{-1}}, \boldsymbol{r}^{2,t_{-2}}, \dots, \boldsymbol{r}^{n,t_{-n}}), \qquad t \in T.$$

Thus,

$$\widetilde{\boldsymbol{m}}_i(t) \stackrel{d}{=} \boldsymbol{m}_i(t), \qquad i \in N, t \in T. \tag{35}$$

The mechanism $\widetilde{\boldsymbol{m}}$ has properties S and I by construction. Since $\boldsymbol{m}$ also has property I, it follows from (35) that

$$\widetilde{\boldsymbol{m}}(t) \stackrel{d}{=} \boldsymbol{m}(t), \qquad t \in T. \tag{36}$$

It is not difficult to see that any correlated strategy of the form $(\boldsymbol{m}, \sigma)$ is a correlated equilibrium if and only if this is so for $(\widetilde{\boldsymbol{m}}, \sigma)$. Therefore, the two mechanisms implement precisely the same CEDs.

An almost identical proof shows that $\{\tilde{S}, D\} \Rightarrow \{S, D\}$. The only required change it to assume that the mechanism $\boldsymbol{m}$ has properties $\tilde{S}$ and D. This assumption implies that for every t the distribution of $\boldsymbol{m}(t)$ is degenerate, which by (36) implies the same for $\widetilde{\boldsymbol{m}}(t)$. Thus, $\widetilde{\boldsymbol{m}}$ satisfies D. ∎

Proposition 15. For CEDs, $\{S, O, I\} \Leftrightarrow \{\tilde{S}, O, I\} \Leftrightarrow \{S, \tilde{O}, I\} \Leftrightarrow \{\tilde{S}, \tilde{O}, I\} \Leftrightarrow \{O, I\} \Leftrightarrow \{\tilde{O}, I\} \Leftrightarrow (\{\tilde{O}\} \wedge \{I\})$ and $\{S, O, D\} \Leftrightarrow \{\tilde{S}, O, D\} \Leftrightarrow \{S, \tilde{O}, D\} \Leftrightarrow \{\tilde{S}, \tilde{O}, D\} \Leftrightarrow \{O, D\} \Leftrightarrow \{\tilde{O}, D\} \Leftrightarrow (\{\tilde{O}\} \wedge \{D\})$.

Proof. In view of Lemma 4, it suffices to show that every CED η whose default mechanism $\boldsymbol{m}$ has property $\tilde{O}$ and, in addition, property I or D is S, O, I- or S, O, D-implementable, respectively. It is shown by Proposition 6 that such a CED is indeed thus implementable *as a CSD*. The proof of that proposition refers to the proof of Proposition 1, where it is shown that the mechanism $\overline{\boldsymbol{m}}$ defined by (18) has the required three properties, and the correlated strategy $(\overline{\boldsymbol{m}}, \bar{\sigma})$, where $\bar{\sigma}$ is defined in (21), has the distribution η. Therefore, it only remains to show that $(\overline{\boldsymbol{m}}, \bar{\sigma})$ is in fact a correlated equilibrium.

As shown in proof of Proposition 12, the mechanism $\overrightarrow{\boldsymbol{m}}$ defined by (28) satisfies $\tilde{O}$ and I, and the correlated strategy $(\overrightarrow{\boldsymbol{m}}, \bar{\sigma})$ is a correlated equilibrium. Therefore, it suffices to prove that

$$\overline{\boldsymbol{m}}(t) \stackrel{d}{=} \overrightarrow{\boldsymbol{m}}(t), \qquad t \in T. \tag{37}$$

Since both mechanisms satisfy I and $\tilde{O}$, (37) is equivalent to

$$\overline{\boldsymbol{m}}_i(t_i, t_{-i}^1) \stackrel{d}{=} \overrightarrow{\boldsymbol{m}}_i(t_i, t_{-i}^1), \qquad i \in N, t_i \in T_i.$$

This condition holds by definition of the two mechanisms. ∎

B.2 Implications

By Proposition 11, an implication relation that does not hold for CSDs also does not hold for CEDs. Therefore, an immediate corollary of Propositions 7, 8, 9 and 10 is the following result.

Proposition 16. For CEDs, $\{S, O, I\} \nRightarrow \{D\}$, $\{S, D\} \nRightarrow \{\tilde{O}\}$, $\{\tilde{O}\} \nRightarrow \{O\}$ and $\{S, O\} \nRightarrow \{I\}$.

The next three propositions identify implication relations that do hold for CSDs but do not hold for CEDs.

Proposition 17. For CEDs, $\{D\} \nRightarrow \{\tilde{S}\}$.

Proof. This is demonstrated by Example 2. ∎

Proposition 18. For CEDs, $\{\tilde{S}\} \nRightarrow \{S\}$.

Proof. This is demonstrated by Example 3. ∎

Proposition 19. For CEDs, $\{\tilde{S}, O\} \nRightarrow \{S, O\}$.

Proof. This is demonstrated by Example 5. ■

Propositions 12, 13, 14 and 15 identify six attributes of correlated equilibrium distributions that are defined by subsets of the fundamental properties of mechanisms. These attributes plus S-implementability are shown in Figure 2 as attributes I_a, II_a, III_a, IV, V, VI and VII. The implication relations that are specified by the Hasse diagram among these attributes all hold trivially since they follow immediately from relations between properties of mechanisms. It follows from Proposition 16 that additional implications among the seven attributes do not hold, and in particular, none of them is equivalent to any of the others. (The more detailed argument given in Section A.3 also applies here, mutatis mutandis.) Three more attributes are defined by $\{\ \}$, $\{\tilde{S}\}$ and $\{D\}$ (I, I_b and III in Figure 2). The implication relations shown in Figure 2 among these attributes and between them and the other seven all hold trivially. It follows from Propositions 17 and 18, and from $\{S, O, I\} \nRightarrow \{D\}$ in Proposition 16, that additional such implications do not hold. Two more attributes are defined by $\{I\}$ and $\{S, O\}$ (II and V_a). It follows from Proposition 19, and from $\{S, O\} \nRightarrow \{I\}$ in Proposition 16, that the implication relations shown in Figure 2 between each of these attributes and each of the other ten are the only ones holding. This proves that there are precisely twelve distinct attributes of CEDs that can be defined by single subsets of the fundamental properties.

Theorem 1 in Section 6.1 can now be proved. It follows immediately from it that attributes I, II, III, IV, V, VI and VII of CEDs are obtained from the similarly numbered attributes of CSDs by restriction. That is, a CED has any of these attributes if and only if it has the corresponding attribute as a CSD.

Proof of Theorem 1. Suppose first that $\mathcal{P} \subseteq \{\tilde{O}, D, I\}$, and let η be a CED that is $\mathcal{P}$-implementable as a CSD. By Lemma 4, the default mechanism of η has all the properties in $\mathcal{P}$. By Lemma 3, the default correlated strategy is a correlated equilibrium. Therefore, η is $\mathcal{P}$-implementable also as a CED.

Next, consider the case $\mathcal{P} = \{O\}$. Let η be a CED that is O-implementable as a CSD. By Lemma 3, the default correlated strategy $(\boldsymbol{m}, \sigma)$ of η is a correlated equilibrium. However, as remarked, the default mechanism $\boldsymbol{m}$ may not satisfy O. On the other hand, by assumption, η is the CSD of some correlated strategy $(\boldsymbol{m}', \sigma')$ with a mechanism that satisfies O, but that correlated strategy may not be a correlated equilibrium. Consider the mechanism $\widetilde{\boldsymbol{m}}$ defined by

$$\widetilde{\boldsymbol{m}}_i(t) = \sigma_i'(t_i, \boldsymbol{m}_i'(t)), \qquad i \in N, t \in T.$$

It clearly satisfies O. In addition, by (15),

$$\widetilde{\boldsymbol{m}}(t) \overset{d}{=} \boldsymbol{m}(t), \qquad t \in \operatorname{supp}(\eta_T).$$

These equalities imply that the correlated strategy obtained from the default one $(\boldsymbol{m}, \sigma)$ by replacing $\boldsymbol{m}$ with a mechanism $\widetilde{\boldsymbol{m}}$ is also a correlated equilibrium, with the same CED. This proves that η is O-implementable also as a CED.

To complete the proof of the theorem it remains to note that, by Propositions 6 and 15, for both CSDs and CEDs, O, I-implementability is equivalent to $\tilde{O}, I$-implementability, and the same is true for O, D- and $\tilde{O}, D$-implementability. ■

B.3 Conjunction of attributes

The next step is to consider attributes of CEDs that are defined by pairs (or possibly triplets, etc.) of subsets of fundamental properties of mechanisms, that is, by conjunction of two (or more) of the twelve attributes identified in the previous subsection. Unlike for CSDs (see Lemma 5), genuinely new attributes can be defined this way. For example, it follows from the next example that the conjunction of S-implementability and D-implementability is a new attribute.

Example 6 A correlated equilibrium distribution that is S- as well as D-implementable but not S, D-implementable

In a two-player Bayesian game, player 1 has two types, t_1' and t_1'', and two actions, L and R. Player 2 has three types, t_2', t_2'' and t_2''', and only one action, L. All type profiles except (t_1', t_2') may occur, and they have the same probability (1/5). If player 1 plays R, the payoff to both players is 0.5. If he plays L, the payoff vector is determined by the type profile according to the following table:

	t_2'	t_2''	t_2'''
t_1'	N/A	(1,0)	(0,1)
t_1''	(3,0)	(0,0)	(0,0)

The lowest possible expected payoff for player 2 in this game is 0.1. For this payoff to be reached, player 1 should play R if and only if the type profile is (t_1', t_2'''). This is in fact a correlated equilibrium distribution, which is implementable by a mechanism that sends to player 1 the message R if the type profile is (t_1', t_2''') and otherwise sends L. Acting according to the message is incentive compatible for player 1 since it always gives maximum payoff to type t_1' and gives t_1'' (who is always instructed to play L) an expected payoff of 1, which is greater than the 0.5 he would receive from playing R. An alternative implementing mechanism sends to player 1 the message L or R if player 2 has type t_2'' or t_2''' respectively, and sends either message with probability 1/2 if the type is t_2'. The correlated strategy with this mechanism that instructs player 1 to follow the mechanism's instructions if his type is t_1' but play L if the type is t_1'' is a correlated equilibrium. This is because the message that type t_1'' of player 1 receives never changes the probability he assigns to player 2's type being t_2', which remains 1/3 regardless of the received message.

There is no implementing mechanism that, like the first mechanism above, does not randomize and, like the second mechanism, sends to player 1 a message that only depends on 2's type, say m_1', m_1'' or m_1''' if the type is t_2', t_2'' and t_2''', respectively. To see this, note that since the action that the given correlated equilibrium distribution specifies for type t_1' of player 1 depends on whether 2's type is t_2'' or t_2''', the message m_1'' must be different from m_1'''. Therefore, one of them, say m_1'', must also be different from m_1'. However, this means that the mechanism effectively tells player 1 whether or not 2's type is t_2''. It follows that, for player 1, choosing L when the type profile is (t_1'', t_2'') is not incentive compatible: R would yield a higher payoff.

Proposition 20. For CEDs, $(\{S\} \wedge \{D\}) \nRightarrow \{S, D\}$ but $(\{S\} \wedge \{D\}) \Leftrightarrow (\{\tilde{S}\} \wedge \{D\}) \Rightarrow \{S, I\}$.[25]

Proof. The first part of the proposition is proved by Example 6. To prove the second part, it suffices to show that $(\{\tilde{S}\} \wedge \{D\}) \Rightarrow \{S, I\}$; the equivalence then follows immediately from the trivial implications $\{S, I\} \Rightarrow \{S\} \Rightarrow \{\tilde{S}\}$.

Consider a CED η that is both $\tilde{S}$- and D-implementable. It has to be shown that η is also S, I-implementable. By the assumption of D-implementability and the second part of Proposition 4, there is a mapping $\phi = (\phi_1, \phi_2, \dots, \phi_n): T \to A$ such that $\eta(\{(t, \phi(t))\}) = \eta_T(\{t\})$ for all type profiles t. By the assumption of $\tilde{S}$-implementability, there is a correlated strategy $(\boldsymbol{m}, \sigma)$ with a mechanism that satisfies $\tilde{S}$ such that η is equal to the joint distribution of the random type profile $\boldsymbol{t}$ and the random action profile $\boldsymbol{a}$ defined by (7). In particular, for every $t \in \operatorname{supp}(\eta_T)$,

$$\Pr(\sigma_i(t_i, \boldsymbol{m}_i(t)) = \phi_i(t) \text{ for all } i) = \Pr(\boldsymbol{a} = \phi(t) \mid \boldsymbol{t} = t) = \frac{\eta\big(\{(t, \phi(t))\}\big)}{\eta_T(\{t\})} = 1.$$

Therefore,

$$\boldsymbol{a}_i = \sigma_i(\boldsymbol{t}_i, \boldsymbol{m}_i(\boldsymbol{t})) = \phi_i(\boldsymbol{t}), \qquad i \in N. \tag{38}$$

Let the mechanism $\widetilde{\boldsymbol{m}}$ be as in the proof of Proposition 14. By (35) and (38),

[25] An argument broadly similar to that used in the proof of the proposition shows that $(\{S\} \wedge \{D\}) \Rightarrow \{S, D\}$ *would* hold if it were assumed that the type distribution η_T has full support (i.e., $\operatorname{supp}(\eta_T) = T$), so that every type profile has positive probability.

			Player 2						
			Type $+1$				**Type -1**		
			L	R			L	R	
Player 1	**Type $+1$**	L	2	0	2/3	L	0	0	2/3
		R	0	1	1/3	R	1	0	1/3
			1/2	1/2			1/2	1/2	
			L	R			L	R	
	Type -1	L	0	0	1/3	L	0	0	1/2
		R	0	0	2/3	R	0	0	1/2
			1/2	1/2			1/2	1/2	

Table 2. A correlated equilibrium distribution for Example 7. The four type profiles are equally probable. For each of them, the actions that players 1 and 2 take are independent. The probabilities that these actions are L or R are given at the margins of the corresponding box. The numbers inside the box are player 1's payoffs. The payoff of player 2 is always 0.

$$\sigma_i(\boldsymbol{t}_i, \widetilde{\boldsymbol{m}}_i(\boldsymbol{t})) = \phi_i(\boldsymbol{t}), \qquad i \in N.$$

Therefore, using the mechanism $\widetilde{\boldsymbol{m}}$ instead of $\boldsymbol{m}$ gives a correlated strategy $(\widetilde{\boldsymbol{m}}, \sigma)$ whose CSD is also η. Moreover, if $\widetilde{\boldsymbol{m}}$ is used and a single player i changes his strategy from σ_i to some other strategy σ_i', the player's expected payoff changes to $E\big(u_i(\boldsymbol{t}, (\sigma_i'(\boldsymbol{t}_i, \widetilde{\boldsymbol{m}}_i(\boldsymbol{t})), \phi_{-i}(\boldsymbol{t})))\big)$. By (35), this new payoff is equal to

$$E\big(u_i(\boldsymbol{t}, (\sigma_i'(\boldsymbol{t}_i, \boldsymbol{m}_i(\boldsymbol{t})), \phi_{-i}(\boldsymbol{t})))\big),$$

which by (38) is also i's expected payoff if he unilaterally changes his strategy from σ_i to σ_i' in the correlated strategy $(\boldsymbol{m}, \sigma)$ (rather than $(\widetilde{\boldsymbol{m}}, \sigma)$). Since the latter is a correlated equilibrium, i's change of strategy cannot increase his expected payoff. This proves that $(\widetilde{\boldsymbol{m}}, \sigma)$ is also a correlated equilibrium, with a mechanism that by construction satisfies S and I. ■

It follows from the next example that the conjunction of S-implementability and I-implementability is also a new attribute of CEDs.

Example 7 A correlated equilibrium distribution that is S- as well as I-implementable but not S, I-implementable

The game structure and common prior are as in Examples 1, 2 and 3. The payoff matrices of player 1 – one for each type profile – are shown in Table 2. Player 2 receives a constant payoff of zero. A mechanism randomly chooses for each type profile an action for each player according to the (marginal) probabilities shown in Table 2 in such a way that these eight choices are independent. It then informs each player of the action that was chosen for him for the actual type profile. The players' strategy of always taking the indicated action is a correlated equilibrium. This is because a change of action by player 1 may affect his payoff only if his type is $+1$ and, in addition, (i) player 2 has type $+1$ and he plays L, (ii) player 2 has type $+1$ and he plays R, or (iii) player 2 has type -1 and he plays L. The effect in case (i) has the opposite sign and twice the magnitude of that in the other two cases. Since (i), (ii) and (iii) always have equal conditional probabilities, given that the type of player 1 is $+1$ and given his action, this relation between the effects means that the conditional expectation of the gain from changing action is always zero. Thus, the probabilities in Table 2 define a correlated equilibrium distribution, which the mechanism described above implements. By construction, the mechanism has property I.The same correlated strategy distribution is also implementable by the following mechanism, which has property S. The mechanism first chooses two pairs of actions, $a^+ = (a^{++}, a^{-+})$ and $a^- = (a^{+-}, a^{--})$, independently of one another. The probability that the pair a^+ equals (L, L), (L, R), (R, L) or (R, R) is $1/6$, $1/2$, $1/6$ and $1/6$, respectively, and for a^- the corresponding probabilities are $1/3$, $1/3$, $1/6$ and $1/6$. Then, for each type profile $t = (t_1, t_2)$, the mechanism chooses an action a_2^t for player 2 with probabilities (for L and R) that depend on (both t and) a^- (that was chosen in the first stage). Specifically, the probability that $a_2^t = L$ is $1/2$ *unless* $t = (+1, -1)$ and, in addition, (i) $a^- = (L, L)$, in which case the probability is $1/4$, or (ii) $a^- = (L, R)$, in which case the probability is $3/4$. Finally, the mechanism sends messages to the players, which depend on the choices made in the first two stages as well as the players' actual type profile $t = (t_1, t_2)$. The message to player 1 is a^+ or a^- if 2's type is $+1$ or -1, respectively, and the message to player 2 is the pair of actions $(a_2^{(t_1,+1)}, a_2^{(t_1,-1)})$. Thus, neither message is affected by the player's own type. It is not very difficult to check that the correlated strategy with this mechanism that instructs each player to choose the first or second action in his message if his type is $+1$ or -1, respectively, has the distribution in Table 2. For example, if $t = (+1, +1)$, the action profile is $(a^{++}, a_2^{(+1,+1)})$, which is (L, L), (L, R), (R, L) or (R, R) with probability $1/3$, $1/3$, $1/6$ and $1/6$, respectively. Thus, the players' actions are independent and are distributed as specified at the margins of the top-left box in Table 2.

To show that the above correlated strategy is a correlated equilibrium, it suffices to prove that, given that the type of player 1 is $+1$ and given the message he receives (which can be (L, L), (L, R), (R, L) or (R, R)), the conditional probabilities of the following three events are equal: (i) player 2 has type $+1$ and he plays L, (ii) player 2 has type $+1$ and he plays R, and (iii) player 2 has type -1 and he plays L. As indicated above, such equality means that player 1 is indifferent between his two actions. The equality can be viewed as the conjunction of two equalities: (a) events (i) and (ii) have equal conditional probabilities, which are necessarily one-half the conditional probability that $t_2 = +1$, and (b) the latter is also equal to twice the conditional probability of (iii). To prove (a), it suffices to note that, given that $t = (+1, +1)$, the message a^+ that player 1 receives and the action $a_2^{(+1,+1)}$ that player 2 takes are conditionally independent, and the probability that the latter is L is $1/2$. To prove (b), note, first, that by the specification of the mechanism and Bayes' rule, the conditional probability that $t_2 = +1$, given that player 1's type is $+1$ and that he receives the message (L, L), (L, R), (R, L) or (R, R), is equal to $1/3$, $3/5$, $1/2$ or $1/2$, respectively. It is therefore sufficient to show that the conditional probability, given the same information, that $t_2 = -1$ and $a_2^{(+1,-1)} = L$ is $1/6$, $3/10$, $1/4$ or $1/4$, respectively. This conditional probability is equal to the product of two terms: the condition probability that $t_2 = -1$, given that $t_1 = +1$ and player 1's message has the specified value, and the condition probability that $a_2^{(+1,-1)} = L$, given that $t = (+1, -1)$ and a^- has that value. The first term is the complement of the conditional probability that $t_2 = +1$, and is hence $2/3$, $2/5$, $1/2$ or $1/2$ if the message is (L, L), (L, R), (R, L) or (R, R), respectively; and by the specification of the mechanism, the second term is $1/4$, $3/4$, $1/2$ or $1/2$, respectively. Therefore, the product of the two terms is $1/6$, $3/10$, $1/4$ or $1/4$, respectively, as had to be shown.

There does not exist any implementing mechanism for the CED specified by Table 2 that has both properties S and I. Note that this is so despite the fact that, according to Table 2, for each type profile, the two players' *actions* are independent. To see this, suppose that such a mechanism does exist. Consider a correlated equilibrium with that mechanism that has the above CED. Partition all the messages that player 1 may receive from the mechanism into four groups, (L, L), (L, R), (R, L) and (R, R), according to the actions that player 1's strategy prescribes to him when he receives the message and his type is $+1$ (first entry) or -1 (second entry). Since the mechanism satisfies S, and hence also $\tilde{S}$, the probability of receiving a message that belongs to a particular group when player 2 has type $+1$ is the same for both types of player 1. Denote these probabilities by p_{LL}^+, p_{LR}^+, p_{RL}^+ and p_{RR}^+. Let p_{LL}^-, p_{LR}^-, p_{RL}^- and p_{RR}^- be the corresponding probabilities for the case in which player 2's type is -1. Since for each type profile the probability that player 1 plays L is as specified by Table 2, the following equalities hold:

$$p_{LL}^+ + p_{LR}^+ = \frac{2}{3}, \qquad p_{LL}^- + p_{LR}^- = \frac{2}{3}, \tag{L_+}$$

$$p_{LL}^+ + p_{RL}^+ = \frac{1}{3}, \qquad p_{LL}^- + p_{RL}^- = \frac{1}{2}. \tag{L_-}$$

By definition, in a correlated equilibrium, taking the prescribed action is always incentive compatible. In particular, type $+1$ of player 1 cannot increase the conditional expectation of his payoff when the message he receives belongs to group (L, L), (L, R), (R, L) or (R, R) by playing R, R, L or L, respectively (instead of the opposite action he is supposed to take). Since the mechanism satisfies I, for every type profile the message that player 1 receives is independent of player 2's message, and hence of his action: player 2 always plays L with probability $1/2$. The above incentive compatibility condition is therefore expressed by the following four inequalities:

$$p_{LL}^{+}\left(-\frac{1}{2}\cdot 2+\frac{1}{2}\cdot 1\right)+p_{LL}^{-}\left(\frac{1}{2}\cdot 1+\frac{1}{2}\cdot 0\right)\leq 0,$$
$$p_{LR}^{+}\left(-\frac{1}{2}\cdot 2+\frac{1}{2}\cdot 1\right)+p_{LR}^{-}\left(\frac{1}{2}\cdot 1+\frac{1}{2}\cdot 0\right)\leq 0,$$
$$p_{RL}^{+}\left(\frac{1}{2}\cdot 2-\frac{1}{2}\cdot 1\right)+p_{RL}^{-}\left(-\frac{1}{2}\cdot 1-\frac{1}{2}\cdot 0\right)\leq 0,$$
$$p_{RR}^{+}\left(\frac{1}{2}\cdot 2-\frac{1}{2}\cdot 1\right)+p_{RR}^{-}\left(-\frac{1}{2}\cdot 1-\frac{1}{2}\cdot 0\right)\leq 0.$$

All inequalities must in fact hold as equalities. If any of the first two inequalities or any of the last two were strict, then $-(p_{LL}^{+}+p_{LR}^{+})+(p_{LL}^{-}+p_{LR}^{-})<0$ or $(p_{RL}^{+}+p_{RR}^{+})-(p_{RL}^{-}+p_{LR}^{-})<0$ would hold. These two inequalities are equivalent (since the probabilities in each quartet sum up to 1), and they contradict (L_{+}). Therefore, in particular, the first and third equalities above hold as equalities, which implies that $-(p_{LL}^{+}+p_{RL}^{+})+(p_{LL}^{-}+p_{RL}^{-})=0$. This equation contradicts (L_{-}). The contradiction proves that no implementing mechanism for the CED specified in Table 2 has both property S (or even only $\tilde{S}$) and I.

Proposition 21. For CEDs, $(\{S\}\wedge\{I\})\nRightarrow\{S,I\}$.

Proof. This is demonstrated by Example 7. ■

Whether the conjunction of $\tilde{S}$-implementability and I-implementability is also a new attribute of CEDs is not known. It depends on the answer to the following question.

Open Question. For CEDs, does $\left(\{\tilde{S}\}\wedge\{I\}\right)\Rightarrow\{S\}$?

This question corresponds to the question mark in Figure 2. The marked attribute is different from the one below it (attribute II_b) if and only if the answer is negative, that is, if there exists a CED in some Bayesian game that is both $\tilde{S}$- and I-implementable but not S-implementable. If the answer is affirmative, the two attributes of CEDs are actually one and the same, that is, they are equivalent.

Depending on the answer to the Open Question, there are two or three attributes of CEDs that can be defined as the conjunction of a pair of incomparable attributes of the twelve ones presented in Section B.2. Thus, there are in total 14 or 15 attributes of CEDs, which are related to one another as in Figure 2. The following lemma shows that this list is complete in that there are no additional, nonequivalent attributes that can be defined as the conjunction of two or more of those in Figure 2. This result holds regardless of the answer to the Open Question.

Lemma 6. The conjunction of any number of the attributes of CEDs in Figure 2 is equivalent to one of the attributes in the same figure.

Proof (an outline). It has to be shown that for every list $\mathcal{P}',\mathcal{P}'',\ldots\subseteq\{S,\tilde{S},O,\tilde{O},D,I\}$ the conjunction $\mathcal{P}'\wedge\mathcal{P}''\wedge\cdots$ is equivalent to one of the attributes in Figure 2. It suffices to consider lists with three or fewer entries, since in any longer list at least two elements represent comparable attributes. Proposition 15 proves the two cases of the conjunction of $\tilde{O}$-implementability and either D- or I-implementability. All the other cases follow quite easily from these two. ■

Appendix C Communication Equilibrium Distributions

As for correlated strategy distributions and correlated equilibrium distributions, different kinds of mechanisms implement different kinds of communication equilibrium distributions. Specifically, for each subset $\mathcal{P}$ of the six fundamental properties of mechanisms, a MED is $\mathcal{P}$-implementable if it is the CSD of some communication equilibrium with a mechanism that has all the properties in $\mathcal{P}$. This appendix, like the previous two, is mainly concerned with the implication relation between these attributes, and conjunctions of several attributes. Implication is denoted by the generic symbol ⇒ when it is clear from the context that it refers to attributes of MEDs. Otherwise, the more explicit symbol $\underset{\text{MED}}{\Rightarrow}$ is used.

The following proposition shows that a necessary condition for the last relation to hold is that a similar relation holds for CEDs. The propositions in Section C.1 prove that this condition is also sufficient, as long as only attributes that are defined by single sets of properties of mechanisms are involved. Thus, the reverse of implication (39) below holds too. However, this result does not extend to attributes that are defined by conjunction (compare Proposition 33 below with the second part of Proposition 20). Hence the differences between the Hasse diagram of the implications relations between attributes of MEDs (Figure 3) and the corresponding diagram for CEDs (Figure 2).

Proposition 22. For every two subsets $\mathcal{P}, \mathcal{Q} \subseteq \{S, \tilde{S}, O, \tilde{O}, D, I\}$,

$$\mathcal{P} \underset{\text{MED}}{\Rightarrow} \mathcal{Q} \text{ implies } \mathcal{P} \underset{\text{CED}}{\Rightarrow} \mathcal{Q}. \tag{39}$$

Moreover, the same is true with $\mathcal{P}$ replaced by by $\mathcal{P}' \wedge \mathcal{P}'' \wedge \cdots$, for any list $\mathcal{P}', \mathcal{P}''$, ... of subsets of $\{S, \tilde{S}, O, \tilde{O}, V, I\}$.

The proof of Proposition 22, which is given at the end of this appendix, uses the results in the following two subsections.

C.1 Equivalences

The following propositions parallel those in Section B.1.

Proposition 23. For MEDs, $\{S, \tilde{O}\} \Leftrightarrow \{\tilde{S}, \tilde{O}\} \Leftrightarrow \{\tilde{O}\}$.

Proof. It has to be shown that the MED η of any communication equilibrium $(\boldsymbol{m}, \sigma)$ with a mechanism that has property $\tilde{O}$ is $S, \tilde{O}$-implementable. Unlike in the proof of Proposition 12, it cannot be assumed that the mechanism $\boldsymbol{m}$ is the default one. Nevertheless, without loss of generality, it may be assumed that it satisfies (4). Otherwise, $\boldsymbol{m}$ could be replaced by any mechanism $\widetilde{\boldsymbol{m}}$ satisfying (4) such that

$$\widetilde{\boldsymbol{m}}(t) \overset{d}{=} \boldsymbol{m}(t), \qquad t \in T.$$

These equalities mean that, for any profile of reported types t, the messages that $\widetilde{\boldsymbol{m}}$ sends are indistinguishable from those of $\boldsymbol{m}$. They implies that $\widetilde{\boldsymbol{m}}$ also has property $\tilde{O}$, and $(\widetilde{\boldsymbol{m}}, \sigma)$ is also a communication equilibrium.

Consider the correlated strategy $(\overrightarrow{\boldsymbol{m}}, \bar{\sigma})$ defined by (21) and the following generalization of (28):

$$\overrightarrow{\boldsymbol{m}}_i(t) = \left(\sigma_i(t_i^1, \boldsymbol{m}_i(t_i^1, t_{-i})), \sigma_i(t_i^2, \boldsymbol{m}_i(t_i^2, t_{-i})), \dots\right), \qquad i \in N, t \in T.$$

Arguments similar to those used in the proof of Proposition 5 show that $\overrightarrow{\boldsymbol{m}}$ has properties S and $\tilde{O}$. To prove that $(\overrightarrow{\boldsymbol{m}}, \bar{\sigma})$ is a communication equilibrium, is has to be shown that, for every player i, type t_i' for that player and strategy $\bar{\sigma}_i': T_i \times A_i^{T_i} \to A_i$,

$$E(\, u_i(\boldsymbol{t}, \boldsymbol{a}) - u_i(\boldsymbol{t}, \boldsymbol{a}') \mid \boldsymbol{t}_i \,) \geq 0, \tag{40}$$

where $\boldsymbol{a}$ is the random action profile corresponding to $\bar{\sigma}$, that is,

$$\boldsymbol{a}_j = \bar{\sigma}_j(\boldsymbol{t}_j, \overrightarrow{\boldsymbol{m}}_j(\boldsymbol{t})) = \sigma_j(\boldsymbol{t}_j, \boldsymbol{m}_j(\boldsymbol{t})), \qquad j \in N, \tag{41}$$

and $\boldsymbol{a}' = (\boldsymbol{a}_1', \boldsymbol{a}_2', \dots, \boldsymbol{a}_n')$ is defined by

$$\begin{aligned} \boldsymbol{a}_i' &= \bar{\sigma}_i'(\boldsymbol{t}_i, \overrightarrow{\boldsymbol{m}}_i(t_i', \boldsymbol{t}_{-i})) = \bar{\sigma}_i'(\boldsymbol{t}_i, (\sigma_i(t_i^1, \boldsymbol{m}_i(t_i^1, \boldsymbol{t}_{-i})), \sigma_i(t_i^2, \boldsymbol{m}_i(t_i^2, \boldsymbol{t}_{-i})), \dots)), \\ \boldsymbol{a}_j' &= \bar{\sigma}_j(\boldsymbol{t}_j, \overrightarrow{\boldsymbol{m}}_j(t_i', \boldsymbol{t}_{-i})) = \sigma_j(\boldsymbol{t}_j, \boldsymbol{m}_j(t_i', \boldsymbol{t}_{-i})), \qquad j \neq i. \end{aligned}$$

Suppose that, for example with $i = 1$ and $t_i' = t_1^1$, (40) does not hold. This implies that there are some t_1'' and $m_1^2, m_1^3, \dots$ such that

$$E(\, u_1(\boldsymbol{t}, \boldsymbol{a}) \mid \boldsymbol{t}_1 = t_1'' \,) < E(\, u_1(\boldsymbol{t}, \boldsymbol{a}') \mid \boldsymbol{t}_1 = t_1'', \boldsymbol{m}_1(t_1^2, \boldsymbol{t}_{-1}) = m_1^2, \boldsymbol{m}_1(t_1^3, \boldsymbol{t}_{-1}) = m_1^3, \dots).$$

It follows from properties $\tilde{O}$ and (4) of $\boldsymbol{m}$ that the pair of random variables $\boldsymbol{t}$ and $\boldsymbol{m}(t_1^1, \boldsymbol{t}_{-1})$ is independent of $\boldsymbol{m}_1(t_1^2, \boldsymbol{t}_{-1}), \boldsymbol{m}_1(t_1^3, \boldsymbol{t}_{-1}), \dots$. Therefore, the last inequality is equivalent to

$$E(u_1(\boldsymbol{t}, \boldsymbol{a}) \mid \boldsymbol{t}_1 = t_1'') \\ < E(u_1(\boldsymbol{t}, (\sigma_1'(\boldsymbol{t}_1, \boldsymbol{m}_1(t_1^1, \boldsymbol{t}_{-1})), \sigma_2(\boldsymbol{t}_2, \boldsymbol{m}_2(t_1^1, \boldsymbol{t}_{-1})), \dots, \sigma_n(\boldsymbol{t}_n, \boldsymbol{m}_n(t_1^1, \boldsymbol{t}_{-1})))) \mid \boldsymbol{t}_1 = t_1''), \tag{42}$$

where $\sigma_1' : T_1 \times M_1 \to A_1$ is the function defined by

$$\sigma_1'(t_1, m_1) = \bar{\sigma}_1'(t_1, (\sigma_1(t_1^1, m_1), \sigma_1(t_1^2, m_1^2), \sigma_1(t_1^3, m_1^3), \dots)).$$

However, in conjunction with (41), inequality (42) contradicts the assumption that $(\boldsymbol{m}, \sigma)$ is a communication equilibrium, since it shows that when player 1's type is t_1'', he can gain from misreporting it as t_1^1 and switching from σ_1 to σ_1'. The contradiction proves that (40) must in fact hold, so that $(\vec{\boldsymbol{m}}, \bar{\sigma})$ is a communication equilibrium. ■

Proposition 24. For MEDs, $\{\tilde{S}, O\} \Leftrightarrow \{O\}$.

Proof. Identical to the proof of Proposition 13. ■

Proposition 25. For MEDs, $\{S, I\} \Leftrightarrow \{\tilde{S}, I\}$ and $\{S, D\} \Leftrightarrow \{\tilde{S}, D\}$.

Proof. Identical to the proof of Proposition 14. ■

Proposition 26. For MEDs, $\{S, O, I\} \Leftrightarrow \{\tilde{S}, O, I\} \Leftrightarrow \{S, \tilde{O}, I\} \Leftrightarrow \{\tilde{S}, \tilde{O}, I\} \Leftrightarrow \{O, I\} \Leftrightarrow \{\tilde{O}, I\} \Leftrightarrow (\{\tilde{O}\} \wedge \{I\})$ and $\{S, O, D\} \Leftrightarrow \{\tilde{S}, O, D\} \Leftrightarrow \{S, \tilde{O}, D\} \Leftrightarrow \{\tilde{S}, \tilde{O}, D\} \Leftrightarrow \{O, D\} \Leftrightarrow \{\tilde{O}, D\} \Leftrightarrow (\{\tilde{O}\} \wedge \{D\})$.

Proof. It suffices to show that $(\{\tilde{O}\} \wedge \{I\}) \Rightarrow \{S, O, I\}$ and similarly with I replaced by D. By Proposition 15, both implications hold of CEDs. The result that they also hold for MEDs follows immediately from the fact that an S, O-implementable CED is automatically a MED. ■

C.2 Implications

Implication (39) in Proposition 22 can equivalently be expressed by its counterpositive: if an example of a $\mathcal{P}$-implementable CED that is not $\mathcal{Q}$-implementable exists, then a similar counterexample can be found for MEDs. Finding the latter may or may not be easy. The former holds if the CED example employs a correlated equilibrium (with a mechanism with the properties in $\mathcal{P}$) that is also a communication equilibrium, that is, players have no incentive to lie about their types. In this case, the same example can be used for MEDs, since a CED that is not $\mathcal{Q}$-implementable a fortiori does not have that attribute as a MED. The proofs of the following two propositions use this simple observation.

Proposition 27. For MEDs, $\{S, O, I\} \nRightarrow \{D\}$, $\{S, D\} \nRightarrow \{\tilde{O}\}$, $\{\tilde{O}\} \nRightarrow \{O\}$ and $\{S, O\} \nRightarrow \{I\}$.

Proof. Proposition 16, which establishes the same for CEDs, relies on Proposition 11. Therefore, it suffices to show that a result similar to the latter holds with communication equilibrium (distribution) replacing correlated equilibrium (distribution). This can be shown by simply making this replacement throughout the proof of Proposition 11. ■

Proposition 28. For MEDs, $\{D\} \nRightarrow \{\tilde{S}\}$.

Proof. The correlated equilibrium with the mechanism with property D that is described in Example 2 is in fact a communication equilibrium. If player 1 lies about his type, player 2 gets as a message an incorrect type profile and will consequentially choose an action with which a positive payoff for 1 is impossible. For a similar reason, player 2 cannot gain from lying; in this case, the lie will only affect type $+1$ of player 2. The MED of this communication equilibrium in not $\tilde{S}$-implementable since, as shown, it does not have that attribute even as a CED. ■

Even if the correlated equilibrium that is used for demonstrating that a certain implication does not hold for CEDs is not a communication equilibrium, not all hope is necessarily lost. It may be possible to make truthful type reports incentive compatible by augmenting the original game with a suitable auxiliary game and modifying the correlated strategy accordingly.

Suppose, for example, that each of the two players in a Bayesian game can have type $+1$ or -1, and all four type profiles are equally probable. The game can then be modified by adding to it an auxiliary game that requires each player to push one of three buttons, B_1, B_2 or B_3. Depending on both players' choice of button and on their types, a very large number $K > 0$ is either added to or subtracted from their payoffs in the original game. Specifically, the change in payoffs ($+K$ or $-K$) is determined according to the following table, where the rows and columns correspond to the choices of player 1 and 2, respectively, and $\tau = t_1 t_2$ is the product of their types:

	B_1	B_2	B_3
B_1	τK	$-\tau K$	$-K$
B_2	$-K$	τK	$-\tau K$
B_3	$-\tau K$	$-K$	τK

Thus, for both $\tau = +1$ and -1, three cells in the table represent reward and six cells represent punishment. Any mechanism in the original game can be turned into a mechanism in the augmented game by appending to the message it sends to each player, which pertains to the original game, a recommendation of button in the auxiliary game. The recommendation is determined in the following way. The mechanism attempts to identify the "rewarding" cells by calculating the product of the players' *reported* types. It then randomly selects one of these cells, each with probability $1/3$, and recommends its row and column to player 1 and 2, respectively. As detailed below, the feature of the mechanism that encourages truth telling is that misreporting will result in misidentification of the rewarding cells. Note that, for any pair of reported types, the recommendation to each player is equally likely to be B_1, B_2 or B_3. Therefore, the modified mechanism has property $\tilde{S}$ or $\tilde{O}$ if the original mechanism has the same property. It cannot, however, have any of the other four fundamental properties. (However, with a somewhat more complicated auxiliary game, it is possible to also retain property S.)

To any correlated equilibrium in the original game, there corresponds a communication equilibrium in the augmented game. In that equilibrium, the mechanism appends recommendations as described above, and then each player pushes the recommended button and plays in the original game according to the original correlated equilibrium. To see that truthful type reports are incentive compatible, suppose that, for example, button B_1 is recommended to type $+1$ of player 1. The player can infer from the recommendation that, if both players reported their types truthfully and they will follow the mechanism's recommendations, player 2 will choose B_1 or B_2 if his type is $+1$ or -1, respectively, and in both cases, the players will get the reward of K. However, if (only) he, player 1, misreported his type as -1, then player 2's choice of button will have the opposite relation with his type. Consequently, player 1 can get a reward rather than a penalty of K only by choosing B_2 or B_3 if 2's type is $+1$ or -1, respectively. However, since the players' types are independent, this means that player 1 cannot get more than zero in expectation. Hence, misreporting the type does not pay.

Proposition 29. For MEDs, $\{\tilde{S}\} \nRightarrow \{S\}$.

Proof. Consider the Bayesian game and the $\tilde{S}$- but not S-implementable CED presented in Example 3. That CED is not a MED. However, a communication equilibrium with a mechanism that has property $\tilde{S}$ can be obtained by modifying the game and the correlated equilibrium described in the original example by adding an auxiliary game as above. The corresponding MED is not S-implementable even as CED. It is not difficult to see that, if it were S-implementable, the same would be true for the original CED. ■

The proofs of the next two propositions involve more special modifications of the original counterexamples, i.e., those pertaining to CEDs.

Proposition 30. For MEDs, $\{\tilde{S}, O\} \nRightarrow \{S, O\}$.

Proof. Consider the following modification of the game and CED in Example 5. Both players can have type $+1$ or -1, and all four type profiles are equally probable. If the players' types are different or identical, respectively, they both receive the payoff specified by the matrix

$$\begin{array}{c} \\ L \\ R \end{array}\begin{array}{c} L \quad R \\ \begin{pmatrix} 0 & 0 \\ 0 & 1.5 \end{pmatrix} \end{array} \text{ or } \begin{array}{c} \\ L \\ R \end{array}\begin{array}{c} L \quad R \\ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \end{array}.$$

The correlated strategy described in the original example is a correlated equilibrium also in the modified game. For a player of any type who receives the message L and takes that action, the expected payoff is $1/2 \cdot 0 + 1/2 \cdot 1 = 0.5$, whereas playing R instead would only yield $1/2 \cdot 1/2 \cdot 1.5 + 1/2 \cdot 0 = 0.375$. If the message is R, taking that action yields 0.875, while playing L would yield 0. This correlated equilibrium is moreover a communication equilibrium. If a player misreports his type, he will maximize his payoff by taking the recommended action, since this is also the action that the other player will take if the (real) types differ (and if the types are identical, then the expected payoff from any action is 0.5). Thus, a dishonest player cannot get more than $1/2 \cdot 1/2 \cdot 1.5 + 1/2 \cdot 0.5 = 0.625$, which is less than the $1/2 \cdot 0.5 + 1/2 \cdot 0.875 = 0.6875$ that a truthful report would yield him.

It remains to show that the corresponding MED is different from the distribution of any communication (or even correlated) equilibrium $(\boldsymbol{m}, \sigma)$ with a mechanism that has properties S and O. The messages that $\boldsymbol{m}$ sends to the players can be written as $\boldsymbol{m}_1(t')$ and $\boldsymbol{m}_2(t')$, for arbitrary type profile t'. Since the players' actions are identical if their types are identical, necessarily

$$\big(\sigma_1(+1, \boldsymbol{m}_1(t')), \sigma_1(-1, \boldsymbol{m}_1(t'))\big) = \big(\sigma_2(+1, \boldsymbol{m}_2(t')), \sigma_2(-1, \boldsymbol{m}_2(t'))\big). \tag{43}$$

If $\boldsymbol{m}_1(t')$ is such that the left- (and, hence, also the right-) hand side equals (L, R) or (R, L), respectively, then type $+1$ or -1 of player 1 will get $1/2 \cdot 1$ from taking the action L he is supposed to take but $1/2 \cdot 1.5$ from playing R. Therefore, the probability that the four actions in (43) are not all the same must be zero, which shows that the above MED, in which the players' actions may differ, cannot be attained. ■

Proposition 31. For MEDs, $(\{S, I\} \wedge \{D\}) \not\Rightarrow \{S, D\}$.

Proof. Consider the following modification of the game and CED in Example 6. Unlike in the original example, player 2 has the constant payoff 0, and he is allowed to choose action R as well as L. Choosing R rather than L reduces by 3 the payoff of type t_1'' of player 1 but has no effect on the other payoffs. The two mechanisms considered in the original example and the corresponding correlated equilibria are modified as follows. Both mechanisms instruct player 2 to play R if player 1 reports the type t_1' and to play L otherwise, and player 2 obeys. Clearly, this means that type t_1'' of player 1 has an incentive to report his type truthfully. The same is true for type t_1', for whom the CED gives the highest possible payoff. ■

An alternative proof for the last proposition can be obtained by using the following simple and generally applicable modification of the game and correlated equilibria in the original example. Instead of changing the players' action spaces, a new player is added to the game. This "player 0" has a single type, and his action space is the collection T of all type profiles of the original players. If the action that player 0 chooses coincides with the other players' actual type profile, everyone gets a huge bonus. Any correlated equilibrium in the original game can be modified as follows. The mechanism sends to player 0 the type reports of the other players, and he chooses the corresponding action. This obviously creates an incentive for the players to report their types truthfully, and thus turns the correlated equilibrium into a communication equilibrium (in the modified game). If property S, $\tilde{S}$, I or D holds for the original mechanism, the modified mechanism also has the same property.

The following proposition uses this construction to show that if the answer to the Open Question presented in Section B.3 is negative, then the same is true for MEDs. Note that if the answer will turn out to be affirmative, the proposition is technically correct but uninformative, since its assertion holds vacuously.

Proposition 32. If, for MEDs, $\big(\{\tilde{S}\} \wedge \{I\}\big) \Rightarrow \{S\}$, then the same is true for CEDs.

Proof. Suppose that, in some Bayesian game, there is CED η that is $\tilde{S}$-implementable and I-implementable but not S-implementable. It has to be shown that a MED with similar properties also exists.

Consider two correlated equilibria whose CED is η, one with a mechanism that satisfies $\tilde{S}$ and the other with a mechanism that satisfies I, and modify the game and the two correlated equilibria by adding a new player, player 0, as detailed above. The modification turns the two correlated equilibria into communication equilibria, whose common MED assigns nonzero probability only to pairs of type and action profiles in which the former coincides with the action of player 0, and the probability in this case is equal to that assigned by η to the pair obtained by omitting player 0's action. Any communication, or even just correlated, equilibrium whose distribution is this MED can be turned into a correlated equilibrium in the original game (in which η is a CED) simply by omitting the message to player 0 and that player's strategy. If the mechanism in that equilibrium had property S, the same would be true after the omission. It therefore follows from the assumption concerning the CED η that a mechanism implementing the MED cannot in fact have property S. ■

Proposition 32 precludes the possibility that the implication under consideration holds for MEDs but not for CEDs. However, it is mute about the opposite possibility. The following example, by contrast, considers a (different) implication for which the latter definitely holds, as comparison with the second part of Proposition 20 shows.

Example 8 A communication equilibrium distribution that is *S*- as well as *D*-implementable but not *S*, *I*-implementable

In a three-player Bayesian game, each player has two types: t_1' and t_1'' for player 1, and t' and t'' for players 2 and 3. All type profiles except (t_1', t', t') may occur, and they have the same probability (1/7). Each player can choose to play L or R. Player 1's payoff depends only on the type profile t and on the other players' actions. Specifically, the payoff is 0 if $t \neq (t_1'', t', t')$, and if an equality holds, it is given by the following symmetric matrix, where the rows and columns correspond to the actions of player 2 and 3:

$$\begin{matrix} & L & R \\ L & \multirow{2}{*}{} 0 & -1 \\ R & -1 & 6 \end{matrix} \quad \begin{pmatrix} 0 & -1 \\ -1 & 6 \end{pmatrix}.$$

For players 2 and 3 the payoff is the sum of two numbers. The first number, which is the same for both players, is: (i) 4 if player 1 plays L and players 2 and 3 have identical types, (ii) 4 also if player 1 plays R and players 2 and 3 have different types, and (iii) 0 otherwise. The second number depends on whether the player's own action is R or L. In the first case, it is equal to $1/2$, and in the second, it is given by the following table, in which the rows correspond to the player's type and the columns correspond to the types of the other two players:

	t_1', t'	t_1'', t'	t_1', t''	t_1'', t''
t'	N/A	1	0	0
t''	3	0	0	0

Consider the mechanism with property D that, for each type profile t, instructs player 1 to play L or R if the (reported) types of players 2 and 3 are identical or different, respectively, and instructs players 2 and 3 to take the actions specified by the following table, in which the rows and columns correspond to the player's own type and to that of the other player, respectively:

	t'	t''
t'	L	R
t''	L	L

This mechanism and the strategies of following the instructions together constitute a communication equilibrium. Player 1 cannot increase his payoff of 0 since there is no way he can make players 2 and 3 play R when they both have type t'. And for these players, a truthful type report is incentive compatible, since if (only) one of them lies, they both lose the 4 they would get from a match between their types (either identical or different) and player 1's action (L or R, respectively). In addition, for players 2 and 3, acting according to their strategy is incentive compatible. For a player of type t', doing so always guarantees maximum payoff, and for type t'', playing R instead of the action L he is instructed to take would decrease the expected payoff by $(1/4 \times 3 - 1/2 =)\ 1/4$.

The MED of the above communication equilibrium is also implementable by a mechanism with property S. That mechanism sends to player 1 the same messages as the mechanism described above, and sends to each of the other two players i $(= 2,3)$ a message that depends on the others' types according to the table

t_1', t'	t_1'', t'	t_1', t''	t_1'', t''
$\boldsymbol{Z}_i$	L	R	R

,

where $(\boldsymbol{Z}_2, \boldsymbol{Z}_3)$ is a pair of (dependent) random variables that equals (L, R) with probability 0.5 and (R, L) with probability 0.5. A communication equilibrium with this mechanism that has the same MED as the previous one is defined as follows. The strategy of each player is to play according to the message he receives, unless he is of type t'', in which case he plays L. For a player of type t'', playing R would not increase the conditional expectation of the payoff, regardless of the message he receives. This is because, given that the received message is L or R, the conditional probability that the other players have types t_1' and t' is $1/3$ or $1/5$, respectively. Since both $1/3 \times 3$ and $1/5 \times 3$ are greater than $1/2$, deviations to R are unprofitable. The incentive compatibility of truthful type reports is proved by arguments similar to those used for the previous equilibrium.

There is no communication equilibrium with a mechanism with properties S and I that has the above MED. To see this, suppose that such a communication equilibrium exists. Since property S implies $\tilde{S}$, the distribution of the mechanism's messages to player 3 only depends on the other players' types, so that it can be described by the table

t_1', t'	t_1'', t'	t_1', t''	t_1'', t''
P^1	P^2	P^3	P^4

,

where P^1, P^2, P^3, P^4 are four probability measures on player 3's message space M_3. If the type of player 3 is t', he is supposed to play L or R if he receives any message in $\mathrm{supp}(P^2)$ or in $\mathrm{supp}(P^3) \cup \mathrm{supp}(P^4)$, respectively. Therefore, these two subsets of M_3 must be disjoint. If the type of player 3 is t'', he is supposed to play L regardless of the message m_3 he receives. Deviation to R should not increase the conditional expectation of the player's payoff, which means that

$$\left(\frac{1}{2} - 3\right) P^1(\{m_3\}) + \frac{1}{2} P^2(\{m_3\}) + \frac{1}{2} P^3(\{m_3\}) + \frac{1}{2} P^4(\{m_3\}) \leq 0.$$

Summing over all $m_3 \in \mathrm{supp}(P^3) \cup \mathrm{supp}(P^4)$ gives

$$-\frac{5}{2} P^1(\mathrm{supp}(P^3) \cup \mathrm{supp}(P^4)) + 0 + \frac{1}{2} + \frac{1}{2} \leq 0.$$

It follows that if the type profile is (t_1', t', t'), the probability that player 3 plays R is at least $2/5$. The same is true for player 2. Therefore, by the assumed independence of the messages (property I), the probability that both 2 and 3 play R when the type profile is (t_1', t', t') is at least $4/25$. Since $4/25 \times 6 + 12/25 \times (-1) > 0$, this means that player 1 has an incentive to misreport his type as t_1' when it is really t_1'', which contradicts the equilibrium assumption.

Proposition 33. For MEDs, $(\{S\} \wedge \{D\}) \nRightarrow \{S, I\}$.

Proof. This is demonstrated by Example 8. ■

Propositions 23, 24, 25 and 26 identify six attributes of communication equilibrium distributions that are defined by subsets of the fundamental properties of mechanisms. Figure 3 shows these attributes, marked $\mathrm{II_a}$, $\mathrm{III_a}$, IV, V, VI and VII, as well as eleven additional ones. The implication relations that are specified by the Hasse diagram among these 17 attributes all hold trivially, since they follow immediately from relations between properties of mechanisms. For two of the implications, it is not known whether the reverse implication also holds. The uncertainty is indicated in Figure 3 by a question mark. If the reverse implication does hold, then the marked box and the one below it should be coalesced, as they represent equivalent attributes. The following arguments show that none of the other attributes in Figure 3 are equivalent, and more generally, that the diagram shows *all* the implication relations between the attributes.

If attributes that involve conjunctions were removed from Figure 2 and Figure 3, the two Hasse diagrams would become identical. In Appendix B it is shown that, among the remaining twelve attributes of CEDs, the implications shown in the diagram are the only ones holding. Essentially the same arguments prove the same for MEDs, except that Propositions 27, 28, 29 and 30 replace 16, 17, 18 and 19, respectively. For each of the attributes in Figure 3 that does involve conjunction, it follows from Propositions 31 and 33 that the only other attributes that imply or are implied by it are those indicated as such by the Hasse diagram. This proves that the diagram is complete in terms of implication relations.

Like the Hasse diagram for CEDs (Figure 2), that for MEDs (Figure 3) is complete also in that it is closed under conjunctions. The proof is similar to that of Lemma 6 except that is uses Proposition 26 instead of 15. Since it follows from Lemma 5 that closedness under conjunctions also holds for CSDs (Figure 1), this gives Theorem 2. Thus, for any given collection of attributes in one of the three Hasse diagrams, there is an attribute in the same diagram that that is equivalent to their conjunction. That attribute can easily be identified. Since it clearly implies each of the attributes and it is implied by every other attribute with the same property, it must be the meet, or greatest lower bound, of the given attributes (see Section 8). For example, Figure 3 shows that the conjunction of O-implentability and I-implementabiltiy is equivalent to S, O, I-implentability. In other words, the only MEDs with both attributes are the mixed-equilibrium distributions.

It is now possible to give the proof of the result presented at the beginning of this section, namely, that the implication relation between attributes of MEDs is in a sense stronger than that for CEDs.

Proof of Proposition 22. As indicated, if attributes that involve conjunctions were removed from Figure 2 and Figure 3, they would become identical. This means that (39) as well as the reverse implication hold for all $\mathcal{P}$ and $\mathcal{Q}$ that belong to the collection of 27 subsets shown in these diagrams, which clearly implies the same for *all* subsets of the six fundamental properties of mechanisms.

To prove the more general implication in which $\mathcal{P}$ is replaced by $\mathcal{P}' \wedge \mathcal{P}''$, it suffices to consider the case in which $\mathcal{P}'$- and $\mathcal{P}''$-implementability (of CEDs, or equivalently MEDs) are incomparable; the case in which they are comparable reduces to the version just analyzed. A straightforward examination shows that, with a single possible exception, for all $\mathcal{Q} \subseteq \{S, \tilde{S}, O, \tilde{O}, D, I\}$, if the meet of $\mathcal{P}'$-implementability and $\mathcal{P}''$-implementability implies $\mathcal{Q}$-implementability in Figure 3, this is so also in Figure 2 (but not conversely). The possible exception is the case where $\mathcal{P}' = \{\tilde{S}\}$, $\mathcal{P}'' = \{I\}$ and $\mathcal{Q} = \{S\}$, which is however covered by Proposition 32. This proves the generalization of (39) in which $\mathcal{P}$ is replaced by $\mathcal{P}' \wedge \mathcal{P}''$.

To prove the further generalization in which the list $\mathcal{P}', \mathcal{P}''$, … has three or more elements it again suffices to consider the case in which no two elements describe comparable attributes. However, it is not difficult to check that this means that, in both diagrams, $\mathcal{P}' \wedge \mathcal{P}'' \wedge \cdots$ is (equivalent to) attribute VII. Therefore, the version of (39) in which $\mathcal{P}$ is replaced by $\mathcal{P}' \wedge \mathcal{P}'' \wedge \cdots$ holds trivially. ■

Appendix D Correlated Strategy, Correlated Equilibrium and Communication Equilibrium Payoffs

Correlated strategy payoffs (CSPs), correlated equilibrium payoffs (CEPs) and communication equilibrium payoffs (MEPs) in Bayesian games can be classified in a manner similar to the classification of CSDs (Figure 1), CEDs (Figure 2) and MEDs (Figure 3). Each subset $\mathcal{P}$ of the fundamental properties of mechanisms defines an attribute of CSPs, CEPs and MEPs, namely, $\mathcal{P}$-implementability. A payoff vector $v = (v_1, v_2, \dots, v_n) \in \mathbb{R}^n$ in a specified n-player Bayesian game is $\mathcal{P}$-implementable if it coincides with the players' (expected) payoffs in some correlated strategy, correlated equilibrium or communication equilibrium with a mechanism that has all the properties in $\mathcal{P}$ (equivalently, if v is obtained in some CSD, CED or MED, respectively, that is implementable by such a mechanism). For two subsets $\mathcal{P}, \mathcal{Q} \subseteq \{S, \tilde{S}, O, \tilde{O}, D, I\}$, $\mathcal{P}$-implementability of CSPs implies $\mathcal{Q}$-implementability if in every Bayesian game every CSP that is implementable by some mechanism with the properties in $\mathcal{P}$ is also implementable by a mechanism with the properties in $\mathcal{Q}$. This relation is written as $\mathcal{P} \underset{\text{CSP}}{\Rightarrow} \mathcal{Q}$. For CEPs and MEPs, the relations $\underset{\text{CEP}}{\Rightarrow}$ and $\underset{\text{MEP}}{\Rightarrow}$ are defined similarly.

The main result concerning implementability of payoff vectors is Theorem 3 in Section 8.2, which asserts that the effects of the properties of the implementing mechanisms on the payoffs mirror the effects on the joint distributions of types and actions. Thus, there is similarity in this respect between these two possible notions of "outcome" in a Bayesian game.

Proof of Theorem 3. The proofs for correlated equilibria and for communication equilibria are nearly identical. Only the former is presented below; the latter can be obtained from it essentially by replacing 'correlated' with 'communication' throughout. The proof for correlated strategies can also be easily obtained from the proof below by simplifying it in the obvious manner.

It has to be shown that, for every $\mathcal{P}, \mathcal{Q} \subseteq \{S, \tilde{S}, O, \tilde{O}, D, I\}$,

$$\mathcal{P} \underset{\text{CEP}}{\Rightarrow} \mathcal{Q} \text{ if and only if } \mathcal{P} \underset{\text{CED}}{\Rightarrow} \mathcal{Q}. \tag{44}$$

One direction of (44) ("if") is easy. $\mathcal{P} \underset{\text{CEP}}{\not\Rightarrow} \mathcal{Q}$ and $\mathcal{P} \underset{\text{CED}}{\Rightarrow} \mathcal{Q}$ are contradictory, since the former means that, in some Bayesian game, there is a $\mathcal{P}$-implementable CED η with a payoff vector that is different from that of every $\mathcal{Q}$-implementable CED in the same game, whereas the latter implies that η itself is $\mathcal{Q}$-implementable.

To prove the nontrivial direction of (44) ("only if"), define the *extension* of a Bayesian game as the game obtained by the addition of *dummy players* — one for each element of $T \times A$. A dummy player has only one possible type and one action, which are therefore insignificant in that they cannot affect the payoff of any player. In the following, the types and actions of the dummy players are ignored, and the collections of type profiles and action profiles in the extended game are thus identified with those in the original game, namely, T and A, respectively. The significance of the dummy players lies in their payoff functions. The payoff function $u_{t,a}: T \times A \to \mathbb{R}$ of the dummy player representing the types-actions pair $(t, a) \in T \times A$ is defined as the indicator function $1_{\{(t,a)\}}$. It returns 1 if the argument is equal to (t, a) and 0 otherwise. Thus, the dummy players' payoffs indicate the types and actions of the original, real players. In particular, for every correlated equilibrium distribution η and every element (t, a) of $T \times A$, the expected payoff of the corresponding dummy player is equal to $\eta(\{(t, a)\})$. It follows that two CEDs in the extended game, η and $\tilde{\eta}$, give the same CEP if and only if they are equal, $\eta = \tilde{\eta}$.

Every mechanism in the original game can be extended in a natural way to a mechanism in the extended game by sending arbitrary constant messages to the dummy players. The original and the extended mechanisms have the exact same fundamental properties, and in the following, they are identified. Using this identification, every correlated strategy in the original game can be extended in a natural way to a correlated strategy with the same mechanism in the extended game by assigning to each of the dummy players his single possible strategy. Observe that:

1. the original correlated strategy has the same distribution as the extended one (recall the above comment regarding the identification of profiles in the original and the extended games), and
2. one of them is a correlated equilibrium if and only if this is so for the other.

Moreover, every CED in the extended game can be obtained in the above manner from some CED in the original game. The former may be the distribution of a correlated strategy with a mechanism that sends variable messages to some dummy players. However, these messages are inconsequential (since a dummy player has only one possible action) and hence can be replaced by constant messages. Such replacement preserves each of the fundamental properties.

Suppose now that $\mathcal{P} \underset{\text{CEP}}{\Rightarrow} \mathcal{Q}$. Then, for every $\mathcal{P}$-implementable CED η in the extended game there is a $\mathcal{Q}$-implementable CED $\tilde{\eta}$ in the same game with an identical payoff vector. As indicated, necessarily $\tilde{\eta} = \eta$, so that η is also $\mathcal{Q}$-implementable. It follows, by Observations 1 and 2 above, that every $\mathcal{P}$-implementable CED in the original game is also $\mathcal{Q}$-implementable. This proves that $\mathcal{P} \underset{\text{CED}}{\Rightarrow} \mathcal{Q}$. ■

Note that the proof of the "only if" direction of (44) applies virtually unchanged also to the more general version in which $\mathcal{P}$ is replaced by $\mathcal{P}' \wedge \mathcal{P}'' \wedge \cdots$, for any list $\mathcal{P}', \mathcal{P}''$, ... of subsets of $\{S, \tilde{S}, O, \tilde{O}, V, I\}$.

References

Aumann RJ (1987) Correlated equilibrium as an expression of Bayesian rationality. Econometrica 55:1–18

Ben-Porath E (2003) Cheap talk in games with incomplete information. J Econ Theory 108:45–71

Bergemann D, Morris SE (2007) Belief free incomplete information games. Cowles Foundation Discussion Paper No. 1629, Yale University

Bergemann D, Morris SE (2011) Correlated equilibrium in games with incomplete information. Cowles Foundation Discussion Paper No. 1822, Yale University

Cotter KD (1991) Correlated equilibrium in games with type-dependent strategies. J Econ Theory 54:48–68

Cotter KD (1994) Type correlated equilibria for games with payoff uncertainty. Econ Theory 4:617–627

Forges F (1990) Universal mechanisms. Econometrica 58:1341–1364

Forges F (1993) Five legitimate definitions of correlated equilibrium in games with incomplete information. Theory Dec 35:277–310

Forges F (2006) Correlated equilibrium in games with incomplete information revisited. Theory Dec 61:329–344

Gerardi D (2004) Unmediated communication in games with complete and incomplete information. J Econ Theory 114:104–131

Kar A, Ray I, Serrano R (2010) A difficulty in implementing correlated equilibrium distributions. Games Econ Behav 69:189–193

Lehrer E, Rosenberg D, Shmaya E (2006) Signaling and mediation in Bayesian games. Unpublished manuscript

Lehrer E, Rosenberg D, Shmaya E (2010) Signaling and mediation in games with common interests. Games Econ Behav 68:670–682

Milchtaich I (2004) Random-player games. Games Econ Behav 47:353–388

Myerson RB (1991) Game theory: Analysis of conflict. Harvard University Press, Cambridge, MA

Myerson RB (1994) Communication, correlated equilibria and incentive compatibility. In: Aumann RJ, Hart S (eds) Handbook of game theory, Vol. 2. North-Holland, Amsterdam pp 827–847

Osborne MJ, Rubinstein R (1994) A course in game theory. MIT Press, Cambridge, MA

Samuelson L, Zhang J (1989) Correlated and mediated equilibria in games with incomplete information. Unpublished manuscript

Shiryaev AN (1996) Probability, 2nd Edition. Springer, New York