Sequential equilibria in Bayesian games with communication

Published version:

DOI:10.1016/j.geb.2006.09.006

Sequential Equilibria in Bayesian Games with Communication*

Dino Gerardi† Roger B. Myerson‡

December 2005

Abstract

We study the effects of communication in Bayesian games when the players are sequentially rational but some combinations of types have zero probability. Not all communication equilibria can be implemented as sequential equilibria. We define the set of strong sequential communication equilibria (SSCE) and characterize it. SSCE differs from the concept of sequential communication equilibrium (SCE) defined by Myerson (1986) in that SCE allows the possibility of trembles by the mediator. We show that these two concepts coincide when there are three or more players, but the set of SSCE may be strictly smaller than the set of SCE for two-player games.

KEYWORDS: Bayesian games, Communication, Communication equilibrium, Sequential communication equilibrium.

JEL CLASSIFICATION: C72, D82

*We thank Adam Brandenburger and Stephen Morris for helpful comments.

†Department of Economics, Yale University, 30 Hillhouse Avenue, New Haven, CT 06511, donato.gerardi@yale.edu.

‡Department of Economics, University of Chicago, 1126 E. 59th St., Chicago, IL 60637, myerson@uchicago.edu.

1 Introduction

Consider a static Bayesian game (Harsanyi 1967-68) and suppose that the players can communicate with one another and with a trustworthy mediator before choosing their actions. Communication gives to the players the opportunity to exchange their information and coordinate their actions. By taking advantage of these opportunities the players may implement a larger set of outcomes.

To study the effects of communication in static games we first need to specify the communication channels available to the players. In other words, we need to define an extensive-form game which consists of two stages. In the first one, the communication stage, the players send and receive messages. In the second stage, the action stage, the players choose their actions simultaneously. Once the game with communication is fully specified, we need a solution concept to analyze it.

The notion of communication equilibrium (Myerson 1982 and Forges 1986) characterizes the limits of what the players can achieve with communication when the solution concept is Bayesian Nash equilibrium (BNE). It follows from the revelation principle for Bayesian games that any communication equilibrium outcome can be implemented with a (canonical) game in which each player reports his type to the mediator and receives a recommendation to play a certain action. In equilibrium, the players are sincere and obedient.

It is well understood that the notions of Nash equilibrium and BNE are not suitable for dynamic games (such as the games with communication) because they permit the players to behave irrationally in events that have zero probability. Clearly, these events are not exogenous but depend crucially on the players' strategies. Therefore, in extensive-form games it is more appropriate to use stronger solution concepts that require rational behavior both on and off the equilibrium path. Two commonly used concepts are subgame perfect equilibrium (Selten 1975) and sequential equilibrium (SE) (Kreps and Wilson 1982).

The goal of this paper is to analyze the effects of communication in Bayesian games when the players are sequentially rational. We characterize the set of outcomes that the players can achieve with communication when the solution concept is SE. We call these outcomes strong sequential communication equilibria (SSCE).[1]

The first observation is that the stronger solution concept does not make a difference in games with full support. These are games in which all profiles of types have strictly positive probability. Intuitively, consider a communication equilibrium of a game with full support and the associated canonical game. Recall that this game admits a BNE in which each player is sincere and obedient. It is immediate to see that in games with full support sequential rationality does not impose any additional constraints. In fact, each

[1] The reason why we use the qualifier "strong" will become apparent shortly.

recommendation that a player receives is on path because there is a profile of types of the opponents that has positive probability and that may have generated the recommendation. In other words, the equilibrium in which the players announce their true types and follow the mediator's recommendations is sequential.

We then turn to games without full support. In these games, one or more type profiles are impossible in the sense that under no circumstances will a player believe that one of these profiles have occurred. Games without full support are very common. Consider, for example, the case in which a piece of information is common knowledge among a certain number of players. Clearly, the profile of types under which the players have different information is simply impossible.

We show that in games without full support not all communication equilibria can be implemented with the solution concept of SE. We also demonstrate that it is not enough to restrict attention to canonical games in which the players report their types and the mediator sends recommendations.

The solution concept of SE requires considering sequences of strategies in which the players tremble, i.e., make small mistakes. These sequence are used to construct consistent beliefs. In this paper, we treat the mediator very differently from the players. In particular, we do not allow the mediator to tremble. When a player receives an "unexpected" message he must believe that one or more of his opponents have deviated. Myerson (1986) considers multistage games with communication. He also requires the players to be sequentially rational and to hold consistent beliefs. However, in contrast to what we do in this paper, Myerson (1986) allows the mediator to tremble. This leads to the definition of sequential communication equilibrium (SCE).

We study the relationship between SCE and SSCE. As the names suggest, we show that any SSCE is also an SCE. The converse is true in games with three or more players. However, in two-person games the two concepts are not equivalent and we characterize the set of SSCE.

The results in this paper also have implications for the implementation of SCE when a trustworthy party is not available. Indeed, one might want to understand whether the players could implement all SCE without any outside mediator. Our analysis provides the first (and missing) step in answering this question. When there are at least three players, we show that the players are able to "replicate" the mediator's trembles. Thus, the mediator is simply needed to guarantee that the players can coordinate their actions without learning too much about their opponents. In Section 6 we explain that this problem can be solved using techniques similar to those developed by the literature on unmediated communication (see below). This implies that (under weak conditions) all SCE can be implemented without the help of an impartial mediator.

We have already mentioned a few important contributions to the literature on communication. This literature originated with Aumann (1974). In this article, Aumann introduced the notion of correlated equilibrium which characterizes the effects of communication in games with complete information.

Implicit in the definition of correlated and communication equilibrium is the assumption that the players can communicate with an impartial mediator. A number of authors have relaxed this assumption by restricting attention to unmediated communication (i.e., direct communication among the players). Barany (1992) and Forges (1990) show that (almost) all correlated and communication equilibria, respectively, can be implemented in Nash equilibrium without a mediator provided that there are at least four players.

Ben-Porath (1998,2003) and Gerardi (2004) consider games with full support and adopt the solution concept of SE. Ben-Porath (1998,2003) provides sufficient conditions for the implementation of correlated and communication equilibria in games with at least three players. Gerardi (2004) constructs a protocol of unmediated communication that permits to achieve (almost) all correlated and communication equilibria in games with five or more players.

In two-person games with complete information it is possible to implement only convex combinations of Nash equilibria when communication is unmediated (and the players are fully rational). Urbano and Vila (2002) assume that the two players have bounded rationality and can solve only problems of limited computational complexity. They show that the two players are able to generate all correlated equilibria without the help of the mediator. Finally, Aumann and Hart (2003) study unmediated communication in two-person games with incomplete information. They characterize the set of outcomes that can be implemented in BNE.

The paper is organized as follows. In Section 2, we provide a formal description of games with communication and define the notion of SSCE. In Section 3, we review the notion of SCE and Myerson's characterization in terms of codominated actions. In Section 4, we study the relationship between the concepts of SCE and SSCE. In Section 5, we analyze two-person games. Section 6 concludes. All the proofs are relegated to the Appendix.

2 Bayesian Games with Communication

Let $G = (T_1, \ldots, T_n, A_1, \ldots, A_n, u_1, \ldots, u_n, p)$ be a finite Bayesian game. The set of players is $N = \{1, \ldots, n\}$. T_i is the set of types of player i and $\hat{T} = \prod_{i \in N} T_i$ is the set of type profiles.

We let $T \subseteq \hat{T}$ denote the set of profiles of types that are possible, i.e., that occur with strictly positive probability. Any profile of types that does not belong to T occurs with probability zero and is considered impossible by the players. In other words, under no circumstances will the players have beliefs that assign positive probability to the type profiles outside the set T.[2] We let $p \in \Delta^0(T)$ denote the probability distribution of the possible profiles of types.[3] Without loss of generality, we assume that every type of every player has strictly positive probability. That is, for every $i \in N$ and every $t_i \in T_i$,

$$p(t_i) \equiv \sum_{t_{-i} \in \hat{T}_{-i}} p(t_{-i}, t_i) > 0,$$

where $\hat{T}_{-i} = \prod_{j \neq i} T_j$ denotes the set of type profiles of the players different from i.

The set of actions available to player i is A_i. We let $A = \prod_{i \in N} A_i$ denote the set of action profiles, and $A_{-i} = \prod_{j \neq i} A_j$ be the set of action profiles of player i's opponents. The payoff function of player i is $u_i : T \times A \to \mathbb{R}$.

We say that G is a game with full support if $T = \hat{T}$, i.e., if all profiles of types have strictly positive probability. For notational convenience, when the game G does not have full support we extend the probability distribution p to $\hat{T}$ by setting $p(t) = 0$ for every $t \in \hat{T} \backslash T$. Similarly, for each player $i \in N$, we extend the payoff function u_i to the set $\hat{T} \times A$ by setting $u_i(t, a) = 0$ for every $t \in \hat{T} \backslash T$, and every $a \in A$.

An outcome of the game G is a mapping from $\hat{T}$ into $\Delta(A)$ that assigns a probability distribution over action profiles to every profile of types. Of course, we are only interested in the outcomes associated with the type profiles that are possible. However, it is convenient to extend the domain of an outcome to the whole set of type profiles $\hat{T}$.

It is well known that pre-play communication expands the set of outcomes that the players can implement. This is possible because communication allows the players to exchange their information and to coordinate their actions. The goal of this paper is provide a better understanding of the benefits and the limits of communication.

We therefore assume that the players can communicate with one another and with an impartial mediator before playing the game G. We consider extensive-form games in which the players first exchange messages and then choose their actions (messages are "cheap" in the sense that they do not affect directly the players' payoffs). Of course, there are different solution concepts to analyze these extensive-form games with communication.

[2]For example, suppose that the types of player 1 and 2 denote the realization of a certain random variable. Suppose also that both players observe the random variable (i.e., the realization is common knowledge between the two players). Then player $i = 1, \ldots, n$ will never believe that the first two players have different types.

[3]Given any finite set Z, we let $\Delta(Z)$ denote the set of probability distributions over Z. We also let $\Delta^0(Z)$ denote the interior of $\Delta(Z)$. That is, $\Delta^0(Z)$ is the set of probability distributions that assign strictly positive probability to every element of Z.

The notion of *communication equilibrium* characterizes the set of outcomes of G that can be implemented with communication when the solution concept is BNE.

Definition 1 (Communication Equilibrium) *A mapping $\mu : \hat{T} \to \Delta(A)$ is a communication equilibrium of G if and only if:*

$$\sum_{t_{-i} \in \widehat{T}_{-i}} \sum_{a \in A} p(t_{-i}|t_i)\, \mu(a|t_{-i}, t_i)\, u_i(t_{-i}, t_i, a) \geqslant$$

$$\sum_{t_{-i} \in \widehat{T}_{-i}} \sum_{a \in A} p(t_{-i}|t_i)\, \mu(a|t_{-i}, t_i')\, u_i(t_{-i}, t_i, a_{-i}, \psi_i(a_i)) \tag{1}$$

$$\forall i \in N,\ \forall (t_i, t_i') \in T_i^2,\ \forall \psi_i : A_i \to A_i.$$

Consider the following game with communication. The players announce their types to the mediator. For each profile of announcements t, the mediator randomly selects an action profile according to the probability distribution $\mu(t)$ and informs each player of his own action. This game is usually called the *canonical* game. Inequality (1) guarantees that the canonical game admits a BNE in which the players announce their types truthfully and follow the mediator's recommendation. Clearly, this equilibrium implements the outcome μ.

On the other hand, it follows from the revelation principle for Bayesian games that if an outcome can be implemented in BNE with communication then the same outcome can also be implemented with a canonical game and with a BNE in which the players are sincere and obedient (see Myerson (1982) and Forges (1986)). Thus, the set of outcomes of a Bayesian game G that can be implemented in BNE with communication is equal to the set of communication equilibria of G. We denote this set by $CE(G)$.

Since games with communication are extensive-form games it is natural to consider solution concepts stronger than BNE and require the players to be rational both on and off the equilibrium path. In this paper, we use the solution concept of SE introduced by Kreps and Wilson (1982).

It is obvious that the set of outcomes of G that can be implemented in SE with communication cannot be larger than $CE(G)$. It is also easy to see that if G is a game with full support and μ is a communication equilibrium of G, then the corresponding canonical game admits an SE in which the players are sincere and, after being sincere, they are also obedient. Intuitively, if G has full support all the recommendations that a sincere player can possibly receive are on path. Upon receiving a recommendation, the sincere player cannot have an incentive to deviate since it is ex-ante optimal for him to be obedient.[4] We

[4] Of course, a player who does not report his true type may prefer to disobey the mediator's recommen-

conclude that in games with full support there is no difference between implementation in BNE and implementation in SE. If G is a game with full support, then the set of outcomes that can be implemented in SE with communication is equal to $CE(G)$.

This result does not extend to games that do not have full support. The following example (taken from Gerardi (2004)) demonstrates that not all communication equilibria can be implemented in SE when some profiles of types have zero probability.

Example 1 *A communication equilibrium that cannot be implemented in SE.*

Consider the following two-person game G'. The sets of types of player 1 and 2 are $T_1 = \{t_1, v_1\}$ and $T_2 = \{t_2, v_2\}$, respectively. The probability of the type profile (v_1, v_2) is zero. All the other type profiles are equally likely. Player 1 chooses an action from the set $A_1 = \{a, b, c\}$ and player 2 has no action available.

Table 1 describes the players' payoffs. For each combination of type profile and action, we report the corresponding vector of payoffs (the first entry denotes the payoff of player 1).

		t_2	v_2
t_1	a	$(1,0)$	$(0,0)$
	b	$(-1,0)$	$(-1,0)$
	c	$(0,2)$	$(1,1)$
v_1	a	$(-1,-1)$	$(0,0)$
	b	$(1,1)$	$(0,0)$
	c	$(0,1)$	$(0,0)$

Table 1: Payoffs of the game G'

Notice that in any BNE of G', type v_1 of player 1 plays action b, and type t_1 chooses either a, or c, or a randomization between the two actions.

It is easy to show that $CE(G')$, the set of communication equilibria of G', is strictly larger than the set of BNE outcomes of G'. A communication equilibrium that does not belong to the set of BNE outcomes of G' is, for example, μ' defined by:

$$\mu'(t_1, t_2) = \mu'(v_1, v_2) = a, \quad \mu'(t_1, v_2) = c, \quad \mu'(v_1, t_2) = b,$$

dation. This, however, does not affect the incentives of a sincere player. In fact, in the SE that we are considering a sincere player assigns probability zero to the event that his opponents lie about their types.

where we adopt the convention of writing, for example, $\mu'(t_1, t_2) = a$ to denote $\mu'(a|t_1, t_2) = 1$. Notice that μ' is the (unique) communication equilibrium that maximizes the expected payoff of player 1.

We now show that when the players are sequentially rational the set of outcomes that can be implemented with communication is strictly smaller than $CE(G')$. Consider an arbitrary game with communication. Suppose that each player sends a message to the mediator. Let M_i denote the set of messages available to player $i = 1, 2$. After receiving a vector of messages (m_1, m_2), the mediator chooses a message from the set S_1 (at random, according to some probability distribution) and sends it to player 1. Finally, player 1 chooses his action.

In this game with communication, type v_1 will play action b after sending any message m_1 and receiving any message s_1 (type v_1 knows that his opponent is of type t_2). Furthermore, type t_1 of player 1 will never play b (independent of player 2's type, b is strictly dominated by a and c).

We now show that if $\mu : T_1 \times T_2 \to \Delta(A_1)$ is an outcome of G' that can be implemented in SE with communication, then $\mu(t_1, t_2) = \mu(t_1, v_2)$. In fact, both types of player 2 prefer action c to any other action when the type of player 1 is t_1. Suppose, by contradiction, that there exists an SE of the game with communication that induces an outcome μ with $\mu(c|t_1, t_2) > \mu(c|t_1, v_2)$. Then type v_2 would have an incentive to deviate and mimic the behavior of type t_2 to increase the probability of c. Similarly, suppose that an SE induces an outcome μ with $\mu(c|t_1, t_2) < \mu(c|t_1, v_2)$. Type t_2 knows that, independent of his strategy, type v_1 of player 1 will choose action b. Thus, type t_2 has an incentive to mimic the behavior of type v_2 : if player 1 has type t_1 the probability of action c will increase.

We therefore conclude that the set of outcomes of G' that can be implemented in SE with communication coincides with the set of BNE outcomes of G'. If the players are sequentially rational, communication cannot expand the set of outcomes that can be implemented.

As Example 1 points out, the reason why a communication equilibrium may not be implementable in SE is that obedience fails to be sequentially rational after recommendations that have zero probability. The notion of BNE allows a player to obey a recommendation to play even a dominated action, provided that in equilibrium this recommendation occurs with probability zero. Clearly, obedience to dominated actions can never be sequentially rational. As this seems to suggest, the notion of SE will put some restrictions on the actions that the mediator can possibly recommend.

Although not all communication equilibria of the game G' can be implemented in SE, the revelation principle is still valid in Example 1. Clearly, any BNE of G' can be implemented in SE with a (trivial) canonical game in which the mediator's recommendation depends only on the type announced by player 1. However, it turns out that when some profiles of types have zero probability the class of canonical games may be too restrictive. Our next example illustrates this point. We show that by considering games with communication different from the canonical ones it is possible expand the set of outcomes that are implementable in SE.

Example 2 *A communication equilibrium that cannot be implemented with the canonical game.*

G'' is a two-person Bayesian game. The sets of types of player 1 and 2 are $T_1 = \{t_1, v_1\}$ and $T_2 = \{t_2, v_2, w_2\}$, respectively. The probability of the type profile (v_1, w_2) is zero. All the other type profiles are equally likely. Player 1 chooses an action from the set $A_1 = \{a, b, c, d, e\}$ and player 2 has no action available.

The players' payoffs are described in Table 2 (the first entry denotes the payoff of player 1).

	t_2	v_2	w_2
t_1	a $(1,0)$ b $(0,0)$ c $(0,1)$ d $(-1,0)$ e $(-1,0)$	a $(0,0)$ b $(1,0)$ c $(0,1)$ d $(-1,0)$ e $(-1,0)$	a $(0,0)$ b $(0,0)$ c $(1,1)$ d $(-1,0)$ e $(-1,0)$
v_1	a $(1,1)$ b $(-2,-1)$ c $(-10,1)$ d $(2,1)$ e $(0,1)$	a $(-2,-1)$ b $(1,1)$ c $(2,1)$ d $(-10,1)$ e $(0,1)$	a $(0,0)$ b $(0,0)$ c $(0,0)$ e $(0,0)$ f $(0,0)$

Table 2: Payoffs of the game G''

The communication equilibrium μ that maximizes the expected utility of player 1 is unique and equal to:

$$\begin{array}{lll} \mu(t_1, t_2) = a & \mu(t_1, v_2) = b & \mu(t_1, w_2) = c \\ \mu(v_1, t_2) = d & \mu(v_1, v_2) = c & \mu(v_1, w_2) = \frac{1}{2}a, \frac{1}{2}b \end{array}$$

Consider the canonical game in which the mediator selects the recommendations according to the function μ. It is easy to check that there is no SE that induces μ. In fact, the outcome μ can be implemented only if player 2 uses a fully revealing strategy. Thus, suppose that player 2 reports his type truthfully. Consider type v_1 of player 1. The beliefs of type v_1 over his opponent's types when he receives recommendation a must be the same as the beliefs when he receives recommendation b. Notice however that action a is optimal for type v_1 only if the probability that player 2 has type t_2 belongs to the interval $[2/3, 8/9]$. On the other hand, action b is optimal for type v_1 only if the probability that player 2 has type t_2 belongs to the interval $[1/9, 1/3]$. It follows that it cannot be sequentially rational for type v_1 to obey both the recommendation a and the recommendation b.

Consider now the following game with communication. Each player sends a message to the mediator. The set of messages available to player 1 is $M_1 = T_1$ (i.e., player 1 announces his type). The set of messages available to player 2 is $M_2 = \{t_2, v_2, w_2, \hat{m}_2, \tilde{m}_2\}$. For each vector of messages $m \in M_1 \times M_2$, the mediator randomly selects an action according to the probability distribution $\gamma(m) \in \Delta(A_1)$ (see below). Then the mediator recommends the chosen action to player 1. Finally, player 1 chooses an action from the set A_1. The mapping $\gamma : M_1 \times M_2 \to \Delta(A_1)$ used by the mediator is given by:

$$\begin{array}{lllll} \gamma(t_1, t_2) = a & \gamma(t_1, v_2) = b & \gamma(t_1, w_2) = c & \gamma(t_1, \hat{m}_2) = a & \gamma(t_1, \tilde{m}_2) = b \\ \gamma(v_1, t_2) = d & \gamma(v_1, v_2) = c & \gamma(v_1, w_2) = \frac{1}{2}a, \frac{1}{2}b & \gamma(v_1, \hat{m}_2) = a & \gamma(v_1, \tilde{m}_2) = b \end{array}$$

This game with communication admits an SE with the following features. First, both players announce their types truthfully. Second, after he announces his type truthfully, player 1 obeys any recommendation.

We now construct consistent beliefs that make it sequentially rational for type v_1 to obey the recommendations a and b (it is trivial to check that all the other constraints are satisfied). Suppose that along the sequence of completely mixed strategies type t_2 sends message v_2 and $\tilde{m}_2$ with probability ε^2 each, and message w_2 and $\hat{m}_2$ with probability 2ε each. Type v_2 sends messages t_2 and $\hat{m}_2$ with probability ε^2 each, and message w_2 and $\tilde{m}_2$ with probability 2ε each. Then upon receiving recommendation a (b) type v_1 believes that with probability $3/4$ ($1/4$) player 2 has type t_2. It follows that it is optimal for type v_1 to follow the recommendation.

Clearly, the SE described above implements the outcome μ (notice that $\gamma(t) = \mu(t)$ for every profile of types $t \in T_1 \times T_2$).

Example 2 implies that there can be loss of generality in restricting attention to canonical games when the solution concept is SE. In particular, the example emphasizes the importance of endowing the players with sets of messages that are larger than their sets of types. In this paper, we consider the following class of games with communication. Consider a Bayesian game G. A game with communication is denoted by $\left(G, (M_i)_{i \in N}, (S_i)_{i \in N}, \gamma\right)$. For every $i \in N$, M_i and S_i are two arbitrary finite sets of messages. Moreover:

$$\gamma : M \to \Delta(S)$$

where $M = \prod_{i \in N} M_i$ and $S = \prod_{i \in N} S_i$.

The game $\left(G, (M_i)_{i \in N}, (S_i)_{i \in N}, \gamma\right)$ proceeds as follows. Player i sends a message $m_i \in M_i$ to the mediator. The players send their messages simultaneously. After receiving the vector of messages $m \in M$, the mediator randomly chooses an element $s \in S$ with probability $\gamma(s|m)$ and sends message s_i to every player $i \in N$. Then the players choose their actions.[5]

Given a game with communication $\left(G, (M_i)_{i \in N}, (S_i)_{i \in N}, \gamma\right)$, we define the set $M_i \otimes S_i$ as follows:

$$M_i \otimes S_i = \left\{(m_i, s_i) \in M_i \times S_i : \gamma(s_i, s_{-i}|m_i, m_{-i}) > 0 \text{ for some } (m_{-i}, s_{-i}) \in M_{-i} \times S_{-i}\right\}.$$

In words, a pair (m_i, s_i) belongs to the set $M_i \otimes S_i$ if and only if player i can receive message s_i after reporting m_i to the mediator. Thus, the set $M_i \otimes S_i$ denotes the collection of information sets at the action stage of each type of player i.

In the game $\left(G, (M_i)_{i \in N}, (S_i)_{i \in N}, \gamma\right)$, the strategy of player i consists of the following two functions:

$$\sigma_i^M : T_i \to \Delta(M_i),$$

$$\sigma_i^A : T_i \times (M_i \otimes S_i) \to \Delta(A_i).$$

The function σ_i^M specifies how player i chooses his message while the function σ_i^A describes how he chooses his action. We denote by $\sigma_i = \left(\sigma_i^M, \sigma_i^A\right)$ player i's strategy and by $\sigma = (\sigma_i)_{i \in N}$ a profile of strategies. We also let $\sigma^M = \left(\sigma_i^M\right)_{i \in N}$ denote the profile of message strategies.

At the action stage player $i = 1, \ldots, n$ has beliefs over the types of his opponents and the messages they sent. Then he can use his opponents' action strategies to generate a system of beliefs over their types and the actions that they will play. Notice that these are the relevant beliefs for player i. This is because messages do not affect the players' payoffs (recall that messages are cheap talk). Thus, player i's beliefs are described by the function

$$\phi_i : T_i \times (M_i \otimes S_i) \to \Delta(\widehat{T}_{-i} \times A_{-i}).$$

[5] Our approach does not depend on there being only one round of communication. If there are multiple rounds, let M_i be player i's set of all possible strategies for sending messages over the various rounds (conditioned at each round on the messages that i has received earlier), and let S_i be the set of all possible sequence of messages that i could receive. Also, the assumption that the sets M_i and S_i are finite does not affect any of our results and avoids a number of unnecessary technicalities.

After announcing message m_i and receiving message s_i, type t_i assigns probability $\phi_i(t_{-i}, a_{-i}|t_i, m_i, s_i)$ to the event that his opponents have the profile of types t_{-i} and will choose the profile of actions a_{-i}. We let $\phi = (\phi_i)_{i\in N}$ denote a profile of beliefs.

An assessment (σ, ϕ) constitutes an SE of the game $\left(G, (M_i)_{i\in N}, (S_i)_{i\in N}, \gamma\right)$ if and only if: (i) for every player $i \in N$, the message strategy σ_i^M is a best response to the strategy profile σ_{-i} (given σ_i^A); (ii) for every $i \in N$, the action strategy σ_i^A is sequentially rational given the beliefs ϕ_i; and (iii) the system of beliefs ϕ is consistent with σ in the sense that ϕ is the limit of the beliefs computed using a sequence of completely mixed messages strategy profiles $\left\{\sigma^{M,k}\right\}_{k=1}^{\infty}$ that converges to σ^M.

Every SE (σ, ϕ) of the game $\left(G, (M_i)_{i\in N}, (S_i)_{i\in N}, \gamma\right)$ induces an outcome of G, i.e. a mapping $\mu : \widehat{T} \to \Delta(A)$, in the following way:

$$\mu(a|t) = \sum_{m\in M}\left(\prod_{i\in N}\sigma_i^M(m_i|t_i)\right)\sum_{s\in S:\gamma(s|m)>0}\gamma(s|m)\left(\prod_{i\in N}\sigma_i^A(a_i|t_i, m_i, s_i)\right).$$

Definition 2 (Strong Sequential Communication Equilibrium) *A mapping $\mu : \widehat{T} \to \Delta(A)$ is a strong sequential communication equilibrium (SSCE) if and only if there exists a game with communication $\left(G, (M_i)_{i\in N}, (S_i)_{i\in N}, \gamma\right)$ with an SE (σ, ϕ) that induces μ.*

As mentioned above, in games with full support the set of SSCE and the set of communication equilibria coincide. In the rest of the paper, we characterize the set of SSCE for Bayesian games without full support.

3 Sequential Communication Equilibria

The fact that we require the players to be sequentially rational forces us to specify the players' beliefs after zero probability events. The notion of SSCE is based on the original definition of SE of Kreps and Wilson (1982). This means that only the players, but not the mediator, are allowed to tremble. Thus, if player i follows his equilibrium message strategy σ_i^M and receives a zero probability message s_i, then he must believe that one or more of his opponents deviated at the message stage. This, in turn, implies that we need to specify the actions that the players choose after they deviate at the message stage. Without specifying these actions it is not possible to check whether the action strategies are sequentially rational.

A different approach is to allow also the mediator to tremble. Although the mediator is supposed to select the messages according to some probability distribution, he can make small mistakes. Thus, a player who receives a zero probability message can now believe that either his opponents deviated or that they did follow their equilibrium strategies but the mediator made a mistake. This approach leads to the notion of SCE introduced by Myerson (1986).

It is convenient to start the analysis with the formal definition of SCE. Then we characterize the set of SCE of a Bayesian game. Finally, in the next two sections, we analyze the relationship between SCE and SSCE.

At this point we need to introduce some additional concepts. We shall provide an informal description of these concepts right after Definition 3.

A *mediation range* $Q = (Q_i(t_i))_{i\in N, t_i\in T_i}$ is a function that assigns a set of actions $Q_i(t_i) \subseteq A_i$ to every type t_i of every player i.

Let a mediation range $Q = (Q_i(t_i))_{i\in N, t_i\in T_i}$ be given. For every type profile $t \in \hat{T}$, we let $Q(t) = \prod_i Q_i(t_i)$. We also let the set $T \otimes Q$ be equal to

$$T \otimes Q = \{(t, a) : t \in T \text{ and } a \in Q(t)\}.$$

Further, for any type t_i we let the set $(T \otimes Q)_{-i}(t_i)$ be defined by:

$$(T \otimes Q)_{-i}(t_i) = \{(t_{-i}, a_{-i}) : (t_{-i}, t_i) \in T \text{ and } a_j \in Q_j(t_j) \text{ for } j = 1, \ldots, i-1, i+1, \ldots, n\}.$$

Finally, given an outcome $\mu : \widehat{T} \to \Delta(A)$, we construct the mediation range $R^\mu = (R_i^\mu(t_i))_{i\in N, t_i\in T_i}$ as follows. For every i and every t_i we let

$$R_i^\mu(t_i) = \left\{a_i \in A_i : \exists t_{-i} \in \widehat{T}_{-i} \text{ and } a_{-i} \in A_{-i} \text{ such that } \mu(a_{-i}, a_i | t_{-i}, t_i) > 0\right\}.$$

Intuitively, suppose that the players announce their types to the mediator and that the mediator uses the function μ to select a profile of recommendations. The set $R_i^\mu(t_i)$ contains all the recommendations that player i could receive when he announces t_i and the other players announce types that may be different from their true types.

We are now ready to define the notion of SCE.

Definition 3 (Sequential Communication Equilibrium) *A mapping $\mu : \widehat{T} \to \Delta(A)$ is an SCE of G if and only if (i) μ is a communication equilibrium of G; and (ii) there exist a mediation range $Q = (Q_i(t_i))_{i\in N, t_i\in T_i}$ and a sequence of mappings $\{\mu_k\}_{k=1}^\infty$ from T into $\Delta(A)$ satisfying the following conditions:*

$$R_i^\mu(t_i) \subseteq Q_i(t_i), \quad \forall i \in N, \quad \forall t_i \in T_i, \tag{2}$$

$$\mu_k(t) \in \Delta^0(Q(t)), \quad \forall t \in T, \quad k = 1, 2, \ldots, \tag{3}$$

$$\lim_{k\to\infty} \mu_k(t) = \mu(t), \quad \forall t \in T, \tag{4}$$

and

$$\sum_{(t_{-i},a_{-i})\in(T\otimes Q)_{-i}(t_i)} \beta_i\left(t_{-i},a_{-i}|t_i,a_i\right)\left(u_i\left(t_{-i},t_i,a_{-i},a_i\right)-u_i\left(t_{-i},t_i,a_{-i},a_i'\right)\right)\geqslant 0$$
$$\forall i\in N,\quad \forall t_i\in T_i,\quad \forall a_i\in Q_i\left(t_i\right),\quad \forall a_i'\in A_i, \tag{5}$$

where $\beta_i\left(t_{-i},a_{-i}|t_i,a_i\right)$ *is given by:*

$$\beta_i\left(t_{-i},a_{-i}|t_i,a_i\right)=\lim_{k\to\infty}\frac{p\left(t_{-i},t_i\right)\mu_k\left(a_{-i},a_i|t_{-i},t_i\right)}{\sum\limits_{\left(t'_{-i},a'_{-i}\right)\in(T\otimes Q)_{-i}(t_i)} p\left(t'_{-i},t_i\right)\mu_k\left(a'_{-i},a_i|t'_{-i},t_i\right)}. \tag{6}$$

Fix a communication equilibrium μ and consider the associated canonical game. We know that the game admits a BNE in which the players are sincere and obedient. Clearly, the notion of communication equilibrium does not require that a sincere player have an incentive to obey recommendations that are off the equilibrium path. This is the additional requirement of the notion of SCE.

Let us reconsider the canonical game. $Q_i\left(t_i\right)$ denotes the set of possible recommendations that player i can receive when he announces type t_i. Given a profile of reports t, the mediator should select an action profile randomly according to the probability distribution $\mu\left(t\right)$. However, the mediator makes mistakes and instead uses the probability distributions $\mu_k\left(t\right)$ $(k=1,2,\ldots)$. Notice, however, that when player i announces t_i, the mediator never recommends an action outside the set $Q_i\left(t_i\right)$. In the limit the probability of every mistake goes to zero.

Suppose that player i reveals his true type t_i. Upon receiving a recommendation $a_i\in Q_i\left(t_i\right)$, the player can use the sequence $\{\mu_k\}_{k=1}^{\infty}$ to compute his beliefs over his opponents' types and recommendations. Constraint (5) guarantees that it is optimal for player i to obey the recommendation a_i provided that his opponents are sincere and obedient (upon being sincere).

The definition of SCE does not specify the actions that the players choose after they lie to the mediator. We can ignore those actions because implicit in the definition of SCE is the idea that a player assigns probability zero to the event that his opponents have lied. Notice, in fact, that any recommendation that a sincere player may receive can be explained by actions of the trembling mediator without any deviations by the other players (recall that $R_i^{\mu}\left(t_i\right)\subseteq Q_i\left(t_i\right)$).[6]

[6] Notice also that it is without loss of generality to assume that a sincere player assigns probability zero to the event that his opponents have lied to the mediator. When the players are sequentially rational and the mediator is allowed to tremble, the set of SCE represents the largest set of outcomes that can be implemented with communication. The basic idea is that the mediator can always replicate the players' tremble. See Myerson (1986) for details. See also Theorem 2 below.

We now turn to the characterization of the set of SCE of a Bayesian game. First, this set is never empty. Clearly, the outcome induced by a BNE of any Bayesian game G is an SCE of G. Second, any communication equilibrium μ of a game with full support is also an SCE. In fact, when $T = \hat{T}$ we can take the set of possible recommendations $Q_i(t_i)$ to be equal to $R_i^{\mu}(t_i)$ because all recommendations in $R_i^{\mu}(t_i)$ will indeed have positive probability. Then constraint (5) is satisfied for any sequence $\{\mu_k\}_{k=1}^{\infty}$ that converges to μ.

In games without full support, however, communication equilibria and SCE are not equivalent concepts. For example, the communication equilibrium described in Example 1 fails to be an SCE. Obviously, type v_1 of player 1 will never obey the recommendation to play the dominated action a.

The fact that in an SCE it has to be sequentially rational for a sincere player to obey all recommendations clearly puts some restrictions on the messages that the mediator can possibly send. For example, a player will never obey a recommendation to play a dominated action. Once we rule out dominated actions, we can go one step further. A player will never play an action that his optimal if and only if at least one of his opponents plays a dominated action. And this process goes on. Therefore, to characterize the set of SCE it is crucial to determine the actions that can be possible recommendations. To do this, we need to introduce the concept of codomination (Myerson 1986).

Consider a Bayesian game G. Recall that T denotes the set of type profiles of G that have strictly positive probability.

Definition 4 (Codominated Actions) *A mediation range $B = (B_i(t_i))_{i \in N, t_i \in T_i}$ is codominated if there does not exist a probability distribution $\pi \in \Delta(T \times A)$ such that:*

$$\sum_{t_{-i}:(t_{-i},t_i)\in T} \sum_{a_{-i}\in A_{-i}} \pi(t_{-i}, t_i, a_{-i}, a_i)\left(u_i(t_{-i}, t_i, a_{-i}, a_i) - u_i(t_{-i}, t_i, a_{-i}, a_i')\right) \geqslant 0,$$

$$\forall i \in N, \quad \forall t_i \in T_i, \quad \forall a_i \in B_i(t_i), \quad \forall a_i' \in A_i,$$

and

$$\sum_{t\in T} \sum_{i\in N} \sum_{a_i \in B_i(t_i)} \sum_{a_{-i}\in A_{-i}} \pi(t, a_i, a_{-i}) > 0.$$

Suppose that $B = (B_i(t_i))_{i \in N, t_i \in T_i}$ is a codominated system of actions. Let us assume that the players report their types truthfully to the mediator and that the mediator randomly selects a profile of recommendations according to some probability distribution. This process generates a probability distribution π over the set $T \times A$. Suppose that there is a positive probability that a player receives a recommendation to play a codominated action. Then it is impossible that all obedience constraints are satisfied. In other words, there is at least one player who has a strictly incentive to disobey at least one recommendation.

Although Definition 4 is rather intuitive, occasionally it will be convenient to work with the following equivalent definition.

Definition 5 (Codominated Actions - Dual Definition) *A mediation range* $B = (B_i(t_i))_{i \in N, t_i \in T_i}$ *is codominated if there exists a vector* $\alpha = (\alpha_i(a'_i|t_i, a_i))_{i \in N, t_i \in T_i, (a_i, a'_i) \in A_i^2}$ *with nonnegative components and such that:*

(i) $\alpha_i(a'_i|t_i, a_i) = 0$ *if* $a_i \notin B_i(t_i)$;

(ii) For every $t \in T$ *and every* $a \in A$, *if* $\{i : a_i \in B_i(t_i)\} \neq \emptyset$ *then*

$$\sum_{i \in N} \sum_{a'_i \in A_i} \alpha_i(a'_i|t_i, a_i)(u_i(t, a) - u_i(t, a_{-i}, a'_i)) < 0.$$

The nonnegative number $\alpha_i(a'_i|t_i, a_i)$ may be interpreted as the shadow price for the incentive constraint that type t_i of player i should not expect to gain by using action a'_i when he is told to choose action a_i. If the mediator recommends a codominated action at least to one player, then the aggregate value of the incentive constraints is negative.

As already mentioned, the two definitions of codominated systems of actions are equivalent. For completeness, we state this result formally. The proof is a simple application of the duality theorem of linear programming and is therefore omitted.

Fact 1 *A mediation range* $B = (B_i(t_i))_{i \in N, t_i \in T_i}$ *is codominated according to Definition 4 if and only if it is codominated according to Definition 5.*

Let B and B' be two mediation ranges. We say that the mediation range B'' is the union of B and B' if for every player $i \in N$ and every type $t_i \in T_i$, $B''_i(t_i) = B_i(t_i) \cup B'_i(t_i)$. It is easy to show that if two mediation ranges are codominated, then their union is also a codominated system of actions.[7]

Since the game G is finite there are finitely many codominated systems of actions. We let $E = (E_i(t_i))_{i \in N, t_i \in T_i}$ denote the union of all codominated systems. In other words, E is the maximal codominated system. We also let $\bar{Q} = \left(\bar{Q}_i(t_i)\right)_{i \in N, t_i \in T_i}$ denote the mediation range that remains after eliminating all codominated actions for all types of all players. Formally, for every $i \in N$ and every $t_i \in T_i$:

$$\bar{Q}_i(t_i) = A_i \backslash E_i(t_i).$$

[7] Suppose B and B' are two codominated mediation ranges. Let α (α') denote the vector used in Definition 5 to show that B (B') is codominated. Then the vector $\alpha + \alpha'$ and the union of B and B' satisfy all the conditions of Definition 5.

It should be emphasized that the mediation range $\bar{Q}$ is nonempty because $\bar{Q}_i(t_i)$ includes all the actions that type t_i of player i uses with positive probability in any BNE of the original game G (for details, see the proof of Lemma 2 in the Appendix).

We are now ready to provide a complete characterization of the set of SCE. The following theorem is a special case of Theorem 2 in Myerson (1986).

Theorem 1 (Myerson 1986) *A communication equilibrium μ is an SCE if and only if for every $i \in N$ and every $t_i \in T_i$:*

$$R_i^{\mu}(t_i) \cap E_i(t_i) = \emptyset.$$

Proof. In the appendix. ■

Theorem 1 reduces the problem of finding all SCE of a Bayesian game to the simpler problem of determining the largest codominated system of actions E. Once we know E it is easy to check whether a mapping $\mu : \widehat{T} \to \Delta(A)$ constitutes an SCE. In fact, μ is an SCE if and only if it satisfies two types of linear constraints. The first type of constraints guarantees that μ is a communication equilibrium. The second type of constraints requires that μ assigns probability zero to any codominated action. It follows that the set of SCE of a Bayesian game is a convex polyhedron.

4 Relationship Between the Sets of SCE and SSCE

In this section we compare the solution concepts SCE and SSCE. As their names suggest, SSCE is a stronger solution concept than SCE. More precisely, we shall show that the two concepts coincide in games with three or more players. However, in two-person games without full support the set of SSCE may be strictly smaller than the set of SCE.

Our first result considers a Bayesian game with an arbitrary number of players. It shows that if an outcome can be implemented with the players' trembles then it can also be implemented with the mediator's trembles.

Theorem 2 *Consider a Bayesian game G. If μ is an SSCE of G then μ is also an SCE of G.*

Proof. In the appendix. ■

The proof of Theorem 2 extends the logic of the revelation principle to the trembles. Consider a game with communication $\left(G, (M_i)_{i\in N}, (S_i)_{i\in N}, \gamma\right)$. Suppose that (σ, ϕ) is an

SE that induces the SSCE μ. Let $\left\{\sigma^{M,k}\right\}_{k=1}^{\infty}$ denote the corresponding sequence of completely mixed message strategies. We construct a canonical game in which the players announce their types to the mediator. Then the mediator uses the strategies σ^M to determine the messages that the players would send in equilibrium. In this stage of the game, however, the mediator can make small mistakes and choose the wrong messages. We use the sequence $\left\{\sigma^{M,k}\right\}_{k=1}^{\infty}$ to determine the probability of these mistakes. Moreover, the mediator uses the function γ and the equilibrium strategies σ^A to determine the actions to recommend to the players. In this stage of the game the mediator does not tremble.

The fact that the mediator trembles according to the probability distributions $\left\{\sigma^{M,k}\right\}_{k=1}^{\infty}$ and the strategies $\sigma_i^A, \ldots, \sigma_n^A$ are sequentially rational implies that it is in the best interest of every player to obey all possible recommendations (even those that have zero probability).

We now investigate whether the opposite of Theorem 2 holds. Is it possible to implement an SCE when only the players but not the mediator tremble? The answer is affirmative provided that there are at least three players.

Theorem 3 *Consider a Bayesian game G with $n \geqslant 3$ players. If μ is an SCE of G then μ is also an SSCE of G.*

Proof. In the appendix. ■

We now provide an informal description of the proof and explain why it requires at least three players. Given an SCE μ, we construct a game with communication and an SE that induces μ.

We require that the game with communication is finite in the sense that the sets of messages available to the players and to the mediator contain finitely many elements. This requirement forces us to establish a preliminary result which we now describe. Given an SCE μ, the players' beliefs are derived from the infinite sequence $\{\mu_k\}_{k=1}^{\infty}$ converging to μ. We show that it is possible to generate the same beliefs by specifying only finitely many mappings from T into $\Delta(A)$. More precisely, given $\{\mu_k\}_{k=1}^{\infty}$ we construct a finite sequence of mappings $\{\tilde{\mu}_1, \ldots, \tilde{\mu}_L\}$ such that: (i) $\tilde{\mu}_1(t) = \mu(t)$ for every type profile $t \in T$. (ii) Suppose that the mediator uses the function $\tilde{\mu}_\ell$, $\ell = 2, \ldots, L$, with probability $\varepsilon^{\ell-1}$ and the function $\tilde{\mu}_1 = \mu$ with the remaining probability. Then as ε goes to zero these trembles generate exactly the same beliefs as the original sequence $\{\mu_k\}_{k=1}^{\infty}$ (see the proof for details).

Our game with communication is as follows. The message that each player sends to the mediator has two components. In particular, each player i announces his type and a number in $\{1, \ldots, L\}$, where L is the length of the sequence $\{\tilde{\mu}_1, \ldots, \tilde{\mu}_L\}$ defined above. Suppose that the players report the profile of types t. If t is not an element of T then the mediator selects a profile of recommendations (i.e., a profile of actions) according to the probability distribution $\mu(t)$. Suppose now that t belongs to T, and let $\tilde{\ell} = 1, \ldots, L$ denote the second largest integer announced by the players (or the largest integer if this is announced by two or more players). In this case, the mediator selects a profile of recommendation according to the probability distribution $\tilde{\mu}_{\tilde{\ell}}(t)$ defined above.

In equilibrium, every player is sincere (i.e., announces his type truthfully) and reports the number one. Moreover, a sincere player obeys every recommendation (both when he reports the number one and when he announces a number larger than one). The fact that μ is a communication equilibrium and that it takes at least two players to induce the mediator to ignore μ implies that no player has an incentive to deviate at the message stage. It remains to show that it is sequentially rational for a sincere player to obey all recommendations.

We consider a sequence of completely mixed message strategies that converge to the equilibrium strategies and satisfy the following two properties: (i) if $z_i' > z_i \geqslant 2$, then it is much more likely that player i is sincere and announces z_i than he is sincere and announces z_i'. In other words, sincere players are more likely to announce small integers. (ii) The deviations in which a player lies about his type are much less likely than the deviations in which the player is sincere. In particular, we assume that it is more likely that two players are sincere and announce two arbitrary numbers than a single player lies about his type.

The two properties mentioned above have the following important implication. The beliefs of a sincere player are identical to the beliefs derived from the mediator's trembles. Then it follows from the definition of the SCE μ that it is optimal for the player to obey every recommendation he receives. It is important to emphasize that this holds even when the sincere player announces a number different from one.

At this point it should also be clear why our proof requires three or more players. Suppose that $n = 2$. Suppose also that player 1 is sincere, announces the number one and receives a recommendation that has zero probability when player 2 follows the equilibrium strategy. What should player 1 believe? It must be the case that player 2 has lied about his type. The beliefs derived from the mediator's trembles are of no use in this case. Clearly, this problem does not arise with three or more players. In fact, each player can believe that two of his opponents were sincere but announced a number different from one.

5 Two-Person Games

Theorem 3 does not cover the case $n = 2$. It is clear that the logic of our proof does not apply when there are only two players. However, at this point it is still an open question whether there is any difference between SCE and SSCE for $n = 2$. Our next example answers this question. We construct a two-person Bayesian game with the set of SSCE strictly included in the set of SCE.

Example 3 *Theorem 3 does not hold for $n = 2$.*

$\tilde{G}$ is a two-person Bayesian game. The sets of types of player 1 and 2 are $T_1 = \{t_1, v_1\}$ and $T_2 = \{t_2, v_2, w_2\}$, respectively. The probability of the type profile (v_1, w_2) is zero. All the other type profiles are equally likely. The sets of actions of player 1 and 2 are $A_1 = \{a, b, c\}$ and $A_2 = \{d, e\}$, respectively. The players' payoffs are described in Table 3 (the first entry denotes the payoff of player 1).

		t_2		v_2		w_2	
		d	e	d	e	d	e
t_1	a	$2, 2$	$-3, 1$	$0, -10$	$0, -11$	$-2, 0$	$1, -1$
	b	$-2, 0$	$1, 1$	$0, 1$	$2, 1$	$0, 1$	$0, 0$
	c	$0, 1$	$0, 0$	$0, 3$	$0, 3$	$2, 2$	$-3, 0$
		d	e	d	e	d	e
v_1	a	$0, 0$	$-1, 1$	$-1, -1$	$2, -3$	$0, 0$	$0, 0$
	b	$2, 2$	$0, 1$	$1, 0$	$1, 0$	$0, 0$	$0, 0$
	c	$1, 1$	$1, 0$	$0, 1$	$2, 1$	$0, 0$	$0, 0$

Table 3: Payoffs of the game $\tilde{G}$

We claim that the SCE μ that maximizes the expected payoff of player 1 is unique and equal to:

$$\mu(t_1, t_2) = (a, d) \quad \mu(t_1, v_2) = (b, e) \quad \mu(t_1, w_2) = (c, d)$$
$$\mu(v_1, t_2) = (b, d) \quad \mu(v_1, v_2) = (c, e) \quad \mu(v_1, w_2) = (a, d)$$

where we write, for example, $\mu(t_1, t_2) = (a, d)$ for $\mu((a, d) \mid (t_1, t_2)) = 1$.

Obviously it is impossible to do better than μ since player 1 obtains the highest possible payoff in every single state. It is also easy to verify that μ is a communication equilibrium. To prove that it is also sequential, we need to demonstrate that type v_1 of player 1 has an incentive to obey the recommendation to play action a. Consider the following mediation range Q:

$$Q_1(t_1) = Q_1(v_1) = \{a, b, c\},$$
$$Q_2(t_2) = Q_2(w_2) = \{d\}, \; Q_2(v_2) = \{e\}.$$

The mediator's trembles are described by a (converging) sequence $\{\mu_k\}_{k=1}^{\infty}$ that satisfies:

$$\mu_k\left((a,d)\,|\,(v_1,t_2)\right)=\tfrac{1}{k^2}\quad \mu_k\left((a,e)\,|\,(v_1,v_2)\right)=\tfrac{1}{k}.$$

for $k=1,2,\ldots.$

It follows that $\beta_1(v_1,a)$, the beliefs of type v_1 when he is sincere and receives a, assign probability one to the event that player 2 has type v_2 and will play action e. This, in turn, implies that action a is optimal for type v_1 and that μ is an SCE.

Suppose now that μ' is a communication equilibrium under which player 1 obtains his highest payoff in every state. Then $\mu'\left((a,e)\,|\,(v_1,v_2)\right)=0$. If not, type v_2 would have an incentive to deviate and play d. Moreover,

$$\mu'\left((a,d)\,|\,(v_1,w_2)\right)+\mu'\left((a,e)\,|\,(v_1,w_2)\right)=1,$$

otherwise type v_2 would have an incentive to lie to the mediator and report message w_2. Notice that action e is strictly dominated for type w_2. Therefore, in any SCE w_2 must receive the recommendation to play d. This shows that μ is the unique SCE that maximizes the expected payoff of player 1.

We now demonstrate that μ is not an SSCE of $\tilde{G}$. By contradiction, suppose that there exists a game with communication $(G,M_1,M_2,S_1,S_2,\gamma)$ and an SE (σ,ϕ) that induces μ. Let $S_i(m_i)$ denote the set of messages that player i could receive when he announces m_i and his opponent announces an arbitrary message:

$$S_i(m_i)=\{s_i\in S_i:\gamma\left(s_i,s_{-i}|m_i,m_{-i}\right)>0\text{ for some }\left(m_{-i},s_{-i}\right)\in M_{-i}\times S_{-i}\}.$$

Furthermore, for every $i=1,2$ and every type $\tau_i\in T_i$, let $\hat{M}_i(\tau_i)$ be the set of messages used in equilibrium by τ_i :

$$\hat{M}_i(\tau_i)=\left\{m_i\in M_i:\sigma_i^M(m_i|\tau_i)>0\right\}.$$

Notice that the sets $\hat{M}_2(t_2)$, $\hat{M}_2(v_2)$ and $\hat{M}_2(w_2)$ must be pairwise disjoint otherwise the SE (σ,ϕ) cannot induce μ. Moreover, the SE (σ,ϕ) must satisfy the following condition. Consider $m_1\in\hat{M}_1(v_1)$, $m_2\in\hat{M}_2(w_2)$, and (s_1,s_2) such that $\gamma(s_1,s_2|m_1,m_2)>0$. Then

$$\sigma_1^A(a|v_1,m_1,s_1)=1.$$

If the above equality fails, type v_2 of player 2 has an incentive to deviate and choose a message from the set $\hat{M}_2(w_2)$ (and then play action d).

Consider now a pair $(\hat{m}_1,\hat{s}_1)$ such that $\hat{m}_1\in\hat{M}_1(v_1)$, $\hat{s}_1\in S_1(\hat{m}_1)$ and $\sigma_1^A(a|v_1,\hat{m}_1,\hat{s}_1)>0$. Obviously, it is sequentially rational for type v_1 to play action a only if he certain that his opponent is of type v_2 and will play action e. Formally, we have $\phi_1(v_2, e|v_1, \hat{m}_1, \hat{s}_1) = 1$. This, in turn, implies that there must be a pair $(\hat{m}_2, \hat{s}_2)$ such that: (i) $\gamma(\hat{s}_1, \hat{s}_2|\hat{m}_1, \hat{m}_2) > 0$; and (ii) $\sigma_2^A(e|v_2, \hat{m}_2, \hat{s}_2) > 0$. Suppose now that type v_2 of player 2 sends message $\hat{m}_2$ and receives $\hat{s}_2$. We claim that he must assign positive probability to the pair (v_1, a) : $\phi_2(v_1, a|v_2, \hat{m}_2, \hat{s}_2) > 0$. This follows from the fact that $\hat{m}_1 \in \hat{M}_1(v_1)$, $\gamma(\hat{s}_1, \hat{s}_2|\hat{m}_1, \hat{m}_2) > 0$ and $\sigma_1^A(a|v_1, \hat{m}_1, \hat{s}_1) > 0$. But we have now reached a contradiction. In fact, $\sigma_2^A(e|v_2, \hat{m}_2, \hat{s}_2)$ must be equal to zero when $\phi_2(v_1, a|v_2, \hat{m}_2, \hat{s}_2) > 0$ because action e is optimal for type v_2 if and only if player 1 puts probability zero on action a.

The example above shows that when there are two players it makes a difference whether or not we assume that the mediator can make mistakes. Now that we know that the concepts of SSCE and SCE are not equivalent, we face the problem of characterizing the set of SSCE for $n = 2$. This problem is complicated by the fact that the class of games with communication to consider is potentially extremely large. In fact, we have already demonstrated in Example 2 that not all SSCE can be implemented with a canonical game. In some cases the set of messages that a player can send to the mediator must be larger than his set of types. It would therefore be useful to put some restrictions on the games that it is necessary to analyze. In what follows, we develop a simple procedure that allows us to find all SSCE of a Bayesian game with two players.

Consider a two-person Bayesian game $G = (T_1, T_2, A_1, A_2, u_1, u_2, p)$. Without loss of generality, we assume that each player $i = 1, 2$ has at least one action available and define $n_i = |T_i|(|A_i| - 1)$. In the rest of the section, we use i to denote an arbitrary player and j to denote his opponent.

Our goal is to construct a class of games with communication in which all SSCE can be implemented. We let $\bar{M}_i$ denote the set of messages available to player i. We assume that $\bar{M}_i$ is equal to:

$$\bar{M}_i = T_i \times \{0, 1, \ldots, n_j\}, \tag{7}$$

with an arbitrary element denoted by $m_i = (t_i, z_i)$.

The product set $\bar{M}_1 \times \bar{M}_2$ will not play any role in our analysis. Instead, we shall often consider the set of pairs of messages in which at least one player announces the number zero. We use $\bar{M}$ to denote this set. Thus, we have:

$$\bar{M} = \left\{ (t_i, z_i)_{i=1,2} \in \bar{M}_1 \times \bar{M}_2 : z_i = 0 \text{ for some } i = 1, 2 \right\}. \tag{8}$$

Throughout the section we let ρ denote an arbitrary function from $\bar{M}$ into the set $\Delta(A)$. Given a function ρ, we construct the set $P_i(\rho)$, $i = 1, 2$, as follows:

$$P_i(\rho) = \{(t_i, a_i) \in T_i \times A_i : \sum_{t_j \in T_j} \sum_{a_j \in A_j} p(t_i, t_j) \rho(a_i, a_j | (t_i, 0), (t_j, 0)) = 0 \text{ and}$$
$$\sum_{m_j \in \bar{M}_j} \sum_{a_j \in A_j} \rho(a_i, a_j | (t_i, 0), m_j) > 0\}.$$

Intuitively, suppose that the players send their messages to the mediator who then selects a pair of recommendations (actions) randomly, according to ρ. Suppose also that each player is expected to announce his type truthfully (i.e., to be sincere) and to report the number zero. For brevity, we refer to this behavior as "correct" behavior. A pair (t_i, a_i) belongs to the set $P_i(\rho)$ if a_i is an "unexpected" recommendation for type t_i when he behaves correctly. That is, a_i is a recommendation that t_i can receive only if his opponent behaves incorrectly.

We say that a function $\rho : \bar{M} \to \Delta(A)$ is *admissible* if the following condition holds. For each pair $(t_i, a_i) \in P_i(\rho)$, $i = 1, 2$, there exists a triple $(t_j, z_j, a_j) \in \bar{M}_j \times A_j$ such that[8]

$$p(t_i, t_j) \rho(a_i, a_j | (t_i, 0), (t_j, z_j)) > 0.$$

Notice that there are two different forms of incorrect behavior. A player can either lie about his type, or the player can reveal his type truthfully and choose a number different from zero. The fact that ρ is admissible has the following implication. If type t_i behaves correctly and receives the unexpected recommendation a_i, then he is not forced to believe that the opponent lied about his type. It is conceivable that the opponent revealed his true type and announced a strictly positive integer.

Clearly, to describe what a player should believe after receiving an unexpected recommendation we need to derive a consistent system of beliefs. This, in turn, requires considering the players' trembles. Thus, we now introduce a pair of functions $f = (f_1, f_2)$ with:

$$f_i : T_i \times \{1, \ldots, n_j\} \to \{1, \ldots, n_j |T_i|\}.$$

It is useful to think at the function f_i in the following way. Consider two pairs (t_i, z_i) and (t'_i, z'_i), where both z_i and z'_i are different from zero. If $f_i(t_i, z_i) < f_i(t'_i, z'_i)$ then it is much more likely that type t_i sends (t_i, z_i) than type t'_i sends (t_i, z_i). If $f_i(t_i, z_i) = f_i(t'_i, z'_i)$ then the probabilities of the two mistakes converge to zero at the same speed.

Given an admissible function ρ and a pair of functions $f = (f_1, f_2)$ we construct the systems of beliefs as follows. Fix a pair (t_i, a_i) in the set $P_i(\rho)$. We let $\beta_i(\cdot | t_i, a_i; \rho, f) \in$

[8] Clearly, z_j must be different from zero.

$\Delta(T_j \times A_j)$ denote the following probability distributions over the types and actions of player j. Consider a pair $(t_j, a_j) \in T_j \times A_j$. We need to distinguish among different cases. First, suppose that there exists $z_j^* = 1, \ldots, n_i$ such that: (i)

$$p(t_i, t_j)\, \rho\left(a_i, a_j | (t_i, 0), \left(t_j, z_j^*\right)\right) > 0,$$

and (ii) for every triple $\left(t_j', z_j', a_j'\right)$ with

$$p\left(t_i, t_j'\right) \rho\left(a_i, a_j' | (t_i, 0), \left(t_j', z_j'\right)\right) > 0$$

(notice that $z_j' > 0$) we have

$$f_j\left(t_j, z_j^*\right) \leqslant f_j\left(t_j', z_j'\right).$$

In this case, $\beta_i(t_j, a_j | t_i, a_i; \rho, f)$ is set equal to:

$$\beta_i(t_j, a_j | t_i, a_i; \rho, f) = \frac{p(t_j, t_i) \sum\limits_{\substack{z_j \in \{1,\ldots,n_i\}: \\ f_j(t_j, z_j) = f_j\left(t_j, z_j^*\right)}} \rho\left(a_i, a_j | (t_i, 0), (t_j, z_j)\right)}{\sum\limits_{\substack{\left(t_j', z_j\right) \in T_j \times \{1,\ldots,n_i\}: \\ f_j\left(t_j', z_j\right) = f_j\left(t_j, z_j^*\right)}} p\left(t_j', t_i\right) \sum\limits_{a_j' \in A_j} \rho\left(a_i, a_j' | (t_i, 0), \left(t_j', z_j\right)\right)}. \tag{9}$$

In all other cases, we set $\beta_i(t_j, a_j | t_i, a_i; \rho, f)$ equal to zero.

Implicit in the construction of our beliefs is the idea that the deviations in which a player lies about his type are much less likely than the deviations in which the player is sincere and announces a number different from zero. Recall that ρ is admissible and, thus, there is nothing that reveals to a player who behaves correctly that his opponent was not sincere. We use the functions f_1 and f_2 to determine the most likely deviations and to compute the players' beliefs.

We are now ready to state our final result.

Theorem 4 *Consider a two-person Bayesian game G. A mapping $\mu : \hat{T} \to \Delta(A)$ is an SSCE of G if and only if there exist an admissible function ρ and a pair of functions $f = (f_1, f_2)$ such that:*

(i) $\rho((t_1, 0), (t_2, 0)) = \mu(t_1, t_2)$ for every $(t_1, t_2) \in \hat{T}$;

(ii) For every $i = 1, 2$, $t_i \in T_i$, $m_i \in \bar{M}_i$, and $\psi_i : A_i \to A_i$,

$$\sum_{t_j \in T_j} \sum_{a \in A} p(t_i, t_j)\, \rho(a | (t_i, 0), (t_j, 0))\, u_i(t_i, t_j, a) \geqslant$$

$$\sum_{t_j \in T_j} \sum_{a \in A} p(t_i, t_j)\, \rho(a | m_i, (t_j, 0))\, u_i(t_i, t_j, \psi_i(a_i), a_j);$$

(iii) For every $i = 1, 2,\ t_i \in T_i,\ z_i = 1, \ldots, n_j,$ and $\psi_i : A_i \to A_i$,

$$\sum_{t_j \in T_j} \sum_{a \in A} p(t_i, t_j)\, \rho(a|\, (t_i, z_i), (t_j, 0))\, u_i(t_i, t_j, a) \geqslant$$
$$\sum_{t_j \in T_j} \sum_{a \in A} p(t_i, t_j)\, \rho(a|\, (t_i, z_i), (t_j, 0))\, u_i(t_i, t_j, \psi_i(a_i), a_j);$$

(iv) For every $i = 1, 2,\ (t_i, a_i) \in P_i(\rho)$, and $a_i' \in A_i$,

$$\sum_{t_j \in T_j} \sum_{a_j \in A_j} \beta_i(t_j, a_j | t_i, a_i; \rho, f)\, u_i(t_i, t_j, a_i, a_j) \geqslant$$
$$\sum_{t_j \in T_j} \sum_{a_j \in A_j} \beta_i(t_j, a_j | t_i, a_i; \rho, f)\, u_i(t_i, t_j, a_i', a_j),$$

where $\beta_i(t_j, a_j | t_i, a_i; \rho, f)$ is as in equation (9).

Proof. In the appendix. ■

Consider the game with communication in which player i sends a message in $\bar{M}_i$ and receives a recommendation in A_i. The mediator chooses the recommendations according to ρ. Notice that ρ is not defined for a pair of messages (m_1, m_2) that does not belong to $\bar{M}$ (that is, when both players report a number different from zero). As we shall see, in this case the value of ρ is irrelevant and can be chosen arbitrarily.

Suppose that conditions (ii)-(iv) are satisfied. It is easy to verify that the game just described admits an SE in which each player reveals his type truthfully and reports the number zero. Moreover, after sending message $(t_i, 0)$, type t_i obeys all recommendations, both those that are expected (i.e., those that have positive probability when the opponent is sincere and announces the number zero) and those that are unexpected. Type t_i also obeys any expected recommendation if he mistakenly sends message (t_i, z_i) with $z_i > 0$. Another feature of the SE is that a correct player never assigns positive probability to the event that the opponent lied about his type.[9]

We conclude that the existence of a pair (ρ, f) satisfying conditions (i)-(iv) is sufficient for an outcome μ to be an SSCE.

On the other hand, Theorem 4 states that conditions (i)-(iv) are also necessary. In the proof, we show that any SSCE of a two person game can be implemented with the game and the SE that we have illustrated above.

[9]Recall that a player is correct if he reveals his type truthfully and announces the number zero.

6 Conclusions

In this paper, we analyze games with communication under the assumption that players behave rationally in all events, including those that have zero probability. We show that this assumption has crucial implications on the effects of communication when the players believe that some type profiles are impossible. We define the notion of SSCE and show that it coincides with the concept of SCE in games with at least three players.

The concept of SSCE assumes that the players can communicate with a trustworthy mediator, although the mediator is not required to tremble. Using techniques from Gerardi (2004) it is possible to show that, under some weak conditions, the mediator is completely superfluous. In particular, suppose that the game has at least five players. An SSCE μ is rational if for every action profile $a \in A$ and every type profile $t \in \hat{T}$ the probability $\mu(a|t)$ is a rational number. Then any convex combination of rational SSCE can be implemented in SE with unmediated communication. Notice that in games with rational parameters any SSCE can be expressed as a convex combination of two or more rational SSCE.[10]

We also show that in two-person games the concepts of SCE and SSCE do not coincide. We provide a characterization of the set of SSCE in games with two players. Perhaps, the complexity of the analysis when the mediator's trembles are not allowed suggests that, for most applications, it may be simpler to admit the possibility that the mediator makes mistakes and use the concept of SCE.

Appendix

Proof of Theorem 1

The proof is based on a series of Lemmata.

Lemma 1 *Consider an SCE μ. Let $Q = (Q_i(t_i))_{i \in N, t_i \in T_i}$ and $\{\mu_k\}_{k=1}^{\infty}$ denote the mediation range and the sequence of mappings from T into $\Delta(A)$, respectively, that satisfy constraints (3)-(5). Suppose that $B = (B_i(t_i))_{i \in N, t_i \in T_i}$ is a codominated mediation range. Then for every $i \in N$ and every $t_i \in T_i$:*

$$Q_i(t_i) \cap B_i(t_i) = \emptyset.$$

Proof. Suppose, by contradiction, that the claim is false. Then the set

$$X = \{(t, a) \in T \otimes Q : a_i \in B_i(t_i) \text{ for some } i \in N\}$$

[10] A Bayesian game G has rational parameter if for every $i \in N$, every $t \in T$, and every $a \in A$, the numbers $p(t)$ and $u_i(t, a)$ are rational. Notice also that if the game has irrational parameters, then any point in the interior of the set of SSCE can be expressed as a convex combination of two or more rational SSCE.

is nonempty.

Given any pair $(t,a) \in X$, we let $\beta(t,a|X)$ be defined by:

$$\beta(t,a|X) = \lim_{k\to\infty} \frac{p(t)\,\mu_k(a|t)}{\sum\limits_{(t',a')\in X} p(t')\,\mu_k(a'|t')}.$$

Also, for any pair (t_i, a_i) with $t_i \in T_i$ and $a_i \in Q_i(t_i)$, we let $\beta(t_i, a_i|X)$ be defined by:

$$\beta(t_i,a_i|X) = \lim_{k\to\infty} \frac{\sum\limits_{(t_{-i},a_{-i}):(t_{-i},t_i,a_{-i},a_i)\in X} p(t_{-i},t_i)\,\mu_k(a_{-i},a_i|t_{-i},t_i)}{\sum\limits_{(t',a')\in X} p(t')\,\mu_k(a'|t')}.$$

Notice that $\beta(t,a|X)$ and $\beta(t_i,a_i|X)$ are well defined since X is nonempty, and for every k and every $t' \in T$ the probability distribution $\mu_k(t')$ assigns strictly positive probability to any action profile in $Q(t')$.

Let α be the vector of nonnegative weights used in Definition 5 to demonstrate that B is a codominated system of actions. Recall that $\alpha_i(a_i'|t_i,a_i) = 0$ if $a_i \notin B_i(t_i)$.

Consider the following weighted sum of the obedience constraints:

$$y = \sum_{i\in N}\sum_{t_i\in T_i}\sum_{a_i\in Q_i(t_i)}\sum_{a_i'\in A_i}\{\beta(t_i,a_i|X)\,\alpha_i(a_i'|t_i,a_i)$$
$$\sum_{(t_{-i},a_{-i})\in(T\otimes Q)_{-i}(t_i)} \beta_i(t_{-i},a_{-i}|t_i,a_i)\,(u_i(t_{-i},t_i,a_{-i},a_i) - u_i(t_{-i},t_i,a_{-i},a_i'))\}$$

where $\beta_i(t_{-i},a_{-i}|t_i,a_i)$ is defined in equation (6).

The variable y is nonnegative since Q and $\{\mu_k\}_{k=1}^{\infty}$ satisfy constraints (3)-(5). Notice that y can be expressed as follows:

$$y = \sum_{(t,a)\in X}\sum_{i\in N}\left\{\beta(t_i,a_i|X)\,\beta_i(t_{-i},a_{-i}|t_i,a_i)\sum_{a_i'\in A_i}\alpha_i(a_i'|t_i,a_i)\,(u_i(t,a) - u_i(t,a_{-i},a_i'))\right\} =$$
$$\sum_{(t,a)\in X}\beta(t,a|X)\sum_{i\in N}\sum_{a_i'\in A_i}\alpha_i(a_i'|t_i,a_i)\,(u_i(t,a) - u_i(t,a_{-i},a_i')) < 0$$

where the inequality follows from the fact that X is nonempty, B is a codominated system and α is the corresponding vector of weights. We therefore reach a contradiction and the proof of the lemma is complete.

Lemma 2 *There exist two mediation ranges, B and Q, and a sequence $\{\eta_k\}_{k=1}^{\infty}$ of probability distributions over $T \otimes Q$ such that:*

(i) B is a codominated system;
(ii) For every $i \in N$ and every $t_i \in T_i$:

$$Q_i(t_i) = A_i \backslash B_i(t_i);$$

(iii) For every $k = 1, 2, \ldots,$ $\eta_k \in \Delta^0(T \otimes Q)$;
(iv) For every $i \in N$, for every $t_i \in T_i$, for every $a_i \in Q_i(t_i)$, and for every $a_i' \in A_i$:

$$\sum_{(t_{-i}, a_{-i}) \in (T \otimes Q)_{-i}(t_i)} \eta(t_{-i}, a_{-i}|t_i, a_i)(u_i(t_{-i}, t_i, a_{-i}, a_i) - u_i(t_{-i}, t_i, a_{-i}, a_i')) \geqslant 0$$

where $\eta(t_{-i}, a_{-i}|t_i, a_i)$ is given by:

$$\eta(t_{-i}, a_{-i}|t_i, a_i) = \lim_{k \to \infty} \frac{\eta_k(t_{-i}, t_i, a_{-i}, a_i)}{\sum_{\left(t'_{-i}, a'_{-i}\right) \in (T \otimes Q)_{-i}(t_i)} \eta_k\left(t'_{-i}, t_i, a'_{-i}, a_i\right)}.$$

Proof. Fix a (mixed-strategy) BNE $\nu^* = (\nu_1^*, \ldots, \nu_n^*)$ of the game G. ν_i^* denotes the equilibrium strategy of player i and $\nu_i^*(t_i) \in \Delta(t_i)$ is the probability distribution (over the set A_i) chosen by type t_i. Let $B_i^0(t_i) \subseteq A_i$ denote the (possibly empty) set of actions that do not belong to the support of $\nu_i^*(t_i)$. Let also $Q_i^0(t_i) = A_i \backslash B_i^0(t_i)$. Given these sets, we construct the mediation ranges $Q^0 = (Q_i^0(t_i))_{i \in N, t_i \in T_i}$ and $B = (B_i^0(t_i))_{i \in N, t_i \in T_i}$. Let ν denote the probability distribution over the set $T \times A$ given by:

$$\nu(t, a) = p(t) \prod_{i \in N} \nu_i^*(a_i|t_i).$$

There are two cases to consider. First, suppose that B^0 is a codominated system of actions. In this case, we set $Q = Q^0$ and

$$\eta_k(t, a) = \nu(t, a), \tag{10}$$

for every $(t, a) \in T \otimes Q^0$, and for every $k = 1, 2, \ldots$. The mediation range Q and the constant sequence $\{\eta_k\}_{k=1}^{\infty}$ defined above clearly satisfy condition (iv) of Lemma 2 and the proof is complete.

We now turn to the second case and assume that B^0 is not a codominated system. We show that there exists a finite sequences of mediation ranges $\left\{B^1, \ldots, B^H\right\}$ satisfying the following properties. First, B^H is a codominated system. Second, $B_i^h(t_i) \subseteq B_i^{h-1}(t_i)$ for every $i \in N$, $t_i \in N$, and $h = 1, \ldots, H$. Finally, consider any $h = 1, \ldots, H$ and construct the mediation range Q^h by letting $Q_i^h(t_i) = A_i \backslash B_i^h(t_i)$. There exists a probability distribution $\pi^h \in \Delta(T \times A)$ such that:

$$\pi^h(t,a) = 0, \quad \forall (t,a) \notin T \otimes Q^h, \tag{11}$$

and

$$\begin{gathered}\sum_{(t_{-i},a_{-i}) \in \left(T \otimes Q^h\right)_{-i}(t_i)} \pi^h(t_{-i},t_i,a_{-i},a_i) > 0, \\ \forall i \in N, \quad \forall t_i \in T_i \quad \forall a_i \in B_i^{h-1}(t_i) \backslash B_i^h(t_i).\end{gathered} \tag{12}$$

The sequences $\{B^1,\ldots,B^H\}$ and $\{\pi^1,\ldots,\pi^H\}$ are constructed inductively as follows. If B^{h-1} is not a codominated system, then there exists a probability distribution $\pi^h \in \Delta(T \times A)$ such that

$$\begin{gathered}\sum_{t_{-i}:(t_{-i},t_i) \in T} \sum_{a_{-i} \in A_{-i}} \pi^h(t_{-i},t_i,a_{-i},a_i)\left(u_i(t_{-i},t_i,a_{-i},a_i) - u_i(t_{-i},t_i,a_{-i},a_i')\right) \geqslant 0, \\ \forall i \in N, \quad \forall t_i \in T_i \quad \forall a_i \in B_i^{h-1}(t_i), \quad \forall a_i' \in A_i,\end{gathered} \tag{13}$$

and

$$\sum_{t \in T} \sum_{i \in N} \sum_{a_i \in B_i^{h-1}(t_i)} \sum_{a_{-i} \in A_{-i}} \pi^h(t,a_i,a_{-i}) > 0.$$

We let the set $B_i^h(t_i)$ be equal to:

$$B_i^h(t_i) = \left\{a_i \in B_i^{h-1}(t_i) : \pi^h(t_{-i},t_i,a_{-i},a_i) = 0 \;\; \forall (t_{-i},a_{-i}) \text{ s.t. } (t_{-i},t_i,a_{-i},a_i) \in T \times A\right\}.$$

It is easy to check that B^h and π^h satisfy conditions (11) and (12).

In this construction, the nonnegative integer $\sum_{i \in N} \sum_{t_i \in T_i} |B_i^h(t_i)|$ is strictly decreasing in h. Thus, the construction must terminate at a codominated system.

Let $B = B^H$ and the mediation range Q be defined by $Q_i(t_i) = A_i \backslash B_i(t_i)$ for every i and every t_i. We now construct the sequence $\{\eta_k\}_{k=1}^{\infty}$ of probability distributions over $T \otimes Q$ (with full support). For every $k = 1, 2, \ldots,$ and every $(t,a) \in T \otimes Q$ let

$$\delta_k(t,a) = \nu(t,a) + \sum_{h=1}^{H} \left(\frac{1}{k}\right)^h \pi^h(t,a) + \left(\frac{1}{k}\right)^{H+1},$$

and

$$\eta_k(t,a) = \frac{\delta_k(t,a)}{\sum_{(t',a') \in T \otimes Q} \delta_k(t',a')}. \tag{14}$$

It remains to demonstrate that condition (iv) of Lemma 2 is satisfied. If $a_i \in Q_i^0(t_i)$, then the beliefs $\eta(\cdot|t_i, a_i)$ of type t_i of player i are derived from the probability distribution ν. In this case, the incentive compatibility constraint is satisfied since ν is the probability distribution induced by a BNE of the game G. On the other hand, if $a_i \in Q_i^h(t_i) \backslash Q_i^{h-1}(t_i)$ for some $h = 1, \ldots, H$, then the beliefs $\eta(\cdot|t_i, a_i)$ of t_i are derived from the probability distribution π^h. This is because the event (t_i, a_i) has strictly positive probability under π^h and zero probability under $\pi^{h'}$ for every $h' < h$. In this case, the incentive compatibility constraint follows from inequality (13). This concludes the proof of the lemma.

Lemma 3 *Let B and Q be the two mediation ranges defined in Lemma 2. Recall also that E is the maximal codominated system and the mediation range $\bar{Q}$ is defined by $\bar{Q}_i(t_i) = A_i \backslash E_i(t_i)$. Then $B = E$ and $Q = \bar{Q}$.*

Proof. B is a codominated system and $B_i(t_i) = A_i \backslash Q_i(t_i)$ for every $i \in N$ and every $t_i \in T_i$. It follows from Lemma 1 that the set $Q_i(t_i)$ does not contain any codominated action. Thus B must be equal to the maximal codominated system E.

Proof of the theorem

Suppose that μ is an SCE and Q is the associated mediation range. Lemma 1 implies that for every i and every t_i, $Q_i(t_i) \cap E_i(t_i) = \emptyset$. Recall that $R_i^\mu(t_i) \subseteq Q_i(t_i)$ (see condition (2)). Thus, $R_i^\mu(t_i) \cap E_i(t_i) = \emptyset$.

Conversely, suppose that μ is a communication equilibrium and $R_i^\mu(t_i) \cap E_i(t_i) = \emptyset$. The last condition implies that $R_i^\mu(t_i) \subseteq \bar{Q}_i(t_i)$ for every i and t_i. Consider now a type profile $t \in T$ and fix an action profile $a^t \in A$ such that $\mu(a^t|t) > 0$. For every $k = 1, 2, \ldots$, let

$$\hat{\mu}_k(a|t) = \begin{cases} 0 & \text{if } a \notin \bar{Q}(t) \\ \left(1 - \frac{1}{k}\right)\mu(a|t) + \frac{1}{k}\frac{\eta_k(t,a)}{p(t)} & \text{if } a \in \bar{Q}(t) \text{ and } a \neq a^t \\ 1 - \sum\limits_{a \neq a^t} \hat{\mu}_k(a|t) & \text{if } a = a^t \end{cases} \tag{15}$$

where $\eta_k(t, a)$ is defined in equation (10) or (14).[11]

It is clear that there exists an integer $\bar{k}$ such that for every $k > \bar{k}$ and for every $t \in T$, $\hat{\mu}_k(t) \in \Delta^0\left(\bar{Q}(t)\right)$. Moreover, for every t, $\lim_{k\to\infty} \hat{\mu}_k(t) = \mu(t)$. In other words, the mediation range $\bar{Q}$ and the sequence of functions $\{\hat{\mu}_k\}_{k=\bar{k}}^{\infty}$ satisfy conditions (3) and (4) of Definition 3.

[11] Lemma 3 guarantees that the mediation range $\bar{Q}$ and the mediation range Q defined in Lemma 2 coincide.

It remains to show that $\bar{Q}$ and $\{\hat{\mu}_k\}_{k=\bar{k}}^{\infty}$ satisfy condition (5). Consider a pair (t_i, a_i) with $t_i \in T_i$ and $a_i \in \bar{Q}_i(t_i)$. There are two cases to consider. First, suppose that

$$\sum_{t_{-i} \in \widehat{T}_{-i}} \sum_{a_{-i} \in A_{-i}} p(t_{-i}|t_i)\, \mu(a_{-i}, a_i | a_{-i}, a_i) > 0. \tag{16}$$

In this case, the beliefs $\beta_i(\cdot|t_i, a_i)$ are derived from and the mapping μ (and the probability distribution p). Condition (5) holds since μ is a communication equilibrium.

Suppose now that the left hand side of inequality (16) is equal to zero. Then the beliefs $\beta_i(\cdot|t_i, a_i)$ are derived from the sequence $\{\eta_k\}_{k=\bar{k}}^{\infty}$. Lemma 2 guarantees that condition (5) is satisfied.

Proof of Theorem 2

Suppose that μ is an SSCE of the Bayesian game G. We let (σ, ϕ) denote the SE of the game with communication $\left(G, (M_i)_{i \in N}, (S_i)_{i \in N}, \gamma\right)$ that induces μ. We also let $\left\{\sigma^{M,k}\right\}_{k=1}^{\infty}$ denote the sequence of completely mixed message strategy profiles that converges to σ^M and that is used to compute the system of beliefs ϕ.

We now construct a mediation range Q and a sequence of mappings $\{\mu_k\}_{k=1}^{\infty}$ from T into $\Delta(A)$ and show that the triple $(\mu, Q, \{\mu_k\}_{k=1}^{\infty})$ satisfies all the conditions of Definition 3. For every $i \in N$, and every $t_i \in T_i$, let $Q_i(t_i)$ be defined by:

$$Q_i(t_i) = \left\{a_i \in A_i : \gamma(s|m)\, \sigma_i^A(a_i|t_i, m_i, s_i) > 0 \text{ for some } (m, s) \in M \times S\right\}.$$

It can be easily verified that $R_i^{\mu}(t_i) \subseteq Q_i(t_i)$ for every i and every t_i.

Let $\{\varepsilon_k\}_{k=1}^{\infty}$ be a sequence of positive numbers such that for every $t = (t_1, \ldots, t_n) \in T$ and every $m = (m_1, \ldots, m_n) \in M$:

$$\lim_{k \to \infty} \frac{\varepsilon_k}{\prod_{i \in N} \sigma_i^{M,k}(m_i|t_i)} = 0.$$

The function $\mu_k : T \to \Delta(A)$, $k = 1, 2, \ldots,$ is defined as follows. For every $t \in T$ and every $a \in Q(t)$ we let

$$\mu_k(a|t) = \frac{\sum_{m \in M} \left(\prod_{i \in N} \sigma_i^{M,k}(m_i|t_i)\right) \sum_{s \in S : \gamma(s|m) > 0} \gamma(s|m) \left(\prod_{i \in N} \sigma_i^A(a_i|t_i, m_i, s_i)\right) + \varepsilon_k}{1 + |Q(t)|\, \varepsilon_k}. \tag{17}$$

It follows that $\mu_k(t) \in \Delta^0(Q(t))$ for every t and every k and $\lim_{k \to \infty} \mu_k(t) = \mu(t)$ for every t.

To complete the proof we need to check condition (5). For every $i \in N$, $t_i \in T_i$, and $a_i \in Q_i(t_i)$, let the set $D_i(t_i, a_i)$ be equal to:

$$D_i(t_i, a_i) = \big\{(m_i, s_i) \in M_i \otimes S_i : \gamma(s_{-i}, s_i | m_{-i}, m_i)\, \sigma_i^A(a_i | t_i, m_i, s_i) > 0 \\ \text{for some } (m_{-i}, s_{-i}) \in M_{-i} \times S_{-i}\big\}.$$

Fix a pair (t_i, a_i) with $a_i \in Q_i(t_i)$. Notice that for every pair $(m_i, s_i) \in D_i(t_i, a_i)$, the probability distribution $\phi_i(t_i, m_i, s_i)$ assigns probability zero to any pair (t_{-i}, a_{-i}) that does not belong to the set $(T \otimes Q)_{-i}(t_i)$.[12] By making a slight abuse of notation, we now view $\phi_i(t_i, m_i, s_i)$ as an element of $\Delta\left((T \otimes Q)_{-i}(t_i)\right)$.

Consider now the game in which the players announce their types to the mediator. If the profile of reports is t the mediator randomly chooses an action profile according to the probability distribution $\mu(t)$. Moreover, the mediator makes small mistakes and his trembles are described by the sequence $\{\mu_k\}_{k=1}^{\infty}$. Suppose that type t_i reveals his type truthfully and receives the recommendation $a_i \in Q_i(t_i)$. Let $\beta_i(t_i, a_i) \in \Delta\left((T \otimes Q)_{-i}(t_i)\right)$ denote the beliefs of player i. It follows from Equation (17) that the beliefs $\beta_i(t_i, a_i)$ can be expressed as:

$$\beta_i(t_i, a_i) = \sum_{(m_i, s_i) \in D_i(t_i, a_i)} d_i(m_i, s_i; t_i, a_i)\, \phi_i(t_i, m_i, s_i),$$

where $d_i(m_i, s_i; t_i, a_i) \geqslant 0$ for every $(m_i, s_i) \in D_i(t_i, a_i)$ and

$$\sum_{(m_i, s_i) \in D_i(t_i, a_i)} d_i(m_i, s_i; t_i, a_i) = 1.$$

In other words, the beliefs $\beta_i(t_i, a_i)$ are a weighted average of the collection of beliefs $(\phi_i(t_i, m_i, s_i))_{(m_i, s_i) \in D_i(t_i, a_i)}$. Notice that for every $(m_i, s_i) \in D_i(t_i, a_i)$, action a_i is optimal for player i given the beliefs $\phi_i(t_i, m_i, s_i)$. It follows that a_i is also optimal given $\beta_i(t_i, a_i)$. This concludes the proof of the theorem.

Proof of Theorem 3

Let μ be an SCE of G. It follows from the proof of Theorem 1 above that, without loss of generality, we can assume that the corresponding mediation range is $\bar{Q}$ and the mediator's trembles are described by the sequence of functions $\{\hat{\mu}_k\}_{k=\bar{k}}^{\infty}$ defined in equation (15).[13] Let $\hat{\beta}_i(t_i, a_i) \in \Delta\left(\left(T \otimes \bar{Q}\right)_{-i}(t_i)\right)$ denote the beliefs (derived from $\{\hat{\mu}_k\}_{k=\bar{k}}^{\infty}$) of player i when

[12] Recall that $\phi_i(t_i, m_i, s_i)$ denotes the beliefs of type t_i when he sends message m_i and receives s_i. The beliefs $\phi_i(t_i, m_i, s_i)$ are computed using the sequence $\left\{\sigma^{M,k}\right\}_{k=1}^{\infty}$ of completely mixed message strategies.

[13] The mediation range $\bar{Q}$ is defined in Section 3. It contains all the actions that are not codominated.

he reveals his true type t_i and receives recommendation $a_i \in \bar{Q}_i(t_i)$. That is, for every $t_i \in T_i$, every $a_i \in \bar{Q}_i(t_i)$, and every $(t_{-i}, a_{-i}) \in \left(T \otimes \bar{Q}\right)_{-i}(t_i)$,

$$\hat{\beta}_i(t_{-i}, a_{-i}|t_i, a_i) = \lim_{k\to\infty} \frac{p(t_{-i}, t_i)\,\hat{\mu}_k(a_{-i}, a_i|t_{-i}, t_i)}{\sum\limits_{\left(t'_{-i}, a'_{-i}\right)\in\left(T\otimes\bar{Q}\right)_{-i}(t_i)} p\left(t'_{-i}, t_i\right)\hat{\mu}_k\left(a'_{-i}, a_i|t'_{-i}, t_i\right)}. \tag{18}$$

Before constructing the game with communication and the SE that implement the SCE μ, we prove a preliminary result. Given the sequence $\{\hat{\mu}_k\}_{k=\bar{k}}^{\infty}$, we show that there exists a finite sequence $\{\tilde{\mu}_1, \ldots, \tilde{\mu}_L\}$ of functions from T into $\Delta(A)$ that satisfy the following two properties: (i) $\tilde{\mu}_1(t) = \mu(t)$ for every $t \in T$; and (ii) for every $t_i \in T_i$, every $a_i \in \bar{Q}_i(t_i)$, and every $(t_{-i}, a_{-i}) \in \left(T \otimes \bar{Q}\right)_{-i}(t_i)$,

$$\hat{\beta}_i(t_{-i}, a_{-i}|t_i, a_i) = \frac{p(t_{-i}, t_i)\,\tilde{\mu}_{\ell^*}(a_{-i}, a_i|t_{-i}, t_i)}{\sum\limits_{\left(t'_{-i}, a'_{-i}\right)\in\left(T\otimes\bar{Q}\right)_{-i}(t_i)} p\left(t'_{-i}, t_i\right)\tilde{\mu}_{\ell^*}\left(a'_{-i}, a_i|t'_{-i}, t_i\right)},$$

where $\hat{\beta}_i(t_{-i}, a_{-i}|t_i, a_i)$ is defined in equation (18) and ℓ^* is the smallest integer $\ell = 1, \ldots, L$ for which

$$\sum_{\left(t'_{-i}, a'_{-i}\right)\in(T\otimes Q)_{-i}(t_i)} \tilde{\mu}_\ell\left(a'_{-i}, a_i|t'_{-i}, t_i\right) > 0.$$

The functions $\tilde{\mu}_1, \ldots, \tilde{\mu}_L$ are constructed inductively. Let $\tilde{\mu}_1 = \mu$. We assume that $\tilde{\mu}_\ell$ is given for $\ell \geqslant 1$. Let F_ℓ denote the set:

$$F_\ell = \left\{(t, a) \in T \otimes \bar{Q} : \tilde{\mu}_\ell(a|t) = 0\right\}.$$

If the set F_ℓ is empty we stop the process and set $L = \ell$. Otherwise, for every $(t, a) \in F_\ell$ we define

$$\tilde{\mu}_{\ell+1}(a|t) = \lim_{k\to\infty} \frac{\hat{\mu}_k(a|t)}{2\left[\sum\limits_{(t', a')\in F_\ell} \hat{\mu}_k(a'|t')\right]}.$$

Moreover, for every pair $(t, a) \in \left(T \otimes \bar{Q}\right) \backslash F_\ell$ we let

$$\tilde{\mu}_{\ell+1}(a|t) = \frac{1 - \sum\limits_{a' : (t, a')\in F_\ell} \tilde{\mu}_{\ell+1}(a'|t)}{\left|\left\{a' \in A : (t, a') \in \left(T \otimes \bar{Q}\right) \backslash F_\ell\right\}\right|}.$$

Clearly, the process must stop after finitely many iterations. It is easy to check that $\{\tilde{\mu}_1, \ldots, \tilde{\mu}_L\}$ satisfy properties (i) and (ii) above. Finally, notice that for every $\ell = 1, \ldots, L$, every $t \in T$, and every $a \in A \backslash \bar{Q}(t)$, $\tilde{\mu}_\ell(a|t) = 0$.

Consider now the following game with communication $\left(G, (M_i)_{i\in N}, (S_i)_{i\in N}, \gamma\right)$. The set of messages available to player i is:

$$M_i = T_i \times \{1, \ldots, L\},$$

where L is the length of the sequence $\{\tilde{\mu}_1, \ldots, \tilde{\mu}_L\}$ defined above. The set of messages S_i is equal to A_i, the set of actions available to player i.

We now describe how the mediator selects an action profile (i.e., the function $\gamma : M \to \Delta(A)$). Let t be the profile of types announced by the players. If $t \in \widehat{T} \backslash T$ then the mediator chooses an action profile randomly according to the probability distribution $\mu(t)$. Suppose now that $t \in T$. Let $\tilde{\ell} = 1, \ldots, L$ denote the second largest integer announced by the players (or the largest integer if this is announced by two or more players). Then the mediator selects an action profile randomly according to the probability distribution $\tilde{\mu}_{\tilde{\ell}}(t)$, where $\tilde{\mu}_{\tilde{\ell}}$ is the $\tilde{\ell}$-th element of the sequence $\{\tilde{\mu}_1, \ldots, \tilde{\mu}_L\}$.

We now construct an SE (σ, ϕ) of the game $\left(G, (M_i)_{i\in N}, (S_i)_{i\in N}, \gamma\right)$ that induces μ. For every $i \in N$, $t_i \in T_i$, we let $M_i^*(t_i)$ denote the set of messages in which player i reveals his type truthfully. Formally,

$$M_i^*(t_i) = \{(t_i, z_i) : z_i = 1, \ldots, L\}.$$

In equilibrium, every type $t_i \in T_i$ of player $i \in N$ sends the message $(t_i, 1)$:

$$\sigma_i^M((t_i, 1) | t_i) = 1.$$

Moreover, type t_i obeys every recommendation after sending a message $m_i \in M_i^*(t_i)$. As we shall show below, after sending a message $m_i \in M_i^*(t_i)$, type t_i assigns probability zero to the event that his opponents have lied about their types. That is, type t_i assigns probability zero to the event that type t_j of player $j \neq i$ sent a message $m_j \notin M_j^*(t_j)$. Therefore, it is not necessary to specify the actions that type t_j chooses after he sends a message $m_j \notin M_j^*(t_j)$.[14]

It remains to describe the sequence of completely mixed message strategy profiles $\left\{\sigma^{M,k}\right\}_{k=1}^{\infty}$ that converges to σ^M and that determines the players' beliefs. Consider type $t_i \in T_i$. We let:

$$\sigma_i^{M,k}(m_i|t_i) = \begin{cases} \left(\frac{1}{k}\right)^{z_i - 1} & \text{if } m_i = (t_i, z_i), \text{ where } z_i = 2, \ldots, L, \\ \left(\frac{1}{k}\right)^{2L+1} & \text{if } m_i = (t_i', z_i), \text{ where } t_i' \neq t_i, \text{ and } z_i = 1, \ldots, L, \\ 1 - \sum\limits_{m_i' \neq (t_i, 1)} \sigma_i^{M,k}(m_i'|t_i) & \text{if } m_i = (t_i, 1). \end{cases}$$

[14]Of course, he will play the action that maximizes his expected payoff.

Clearly, $\sigma_i^{M,k}(m_i|t_i) > 0$ for every $m_i \in M_i$ when k is sufficiently large.

Given this profile of strategies, it is easy to check that player $i = 1, \ldots, n$ does not have an incentive to deviate when he announces the message m_i. Notice, in fact, that is takes at least two players to convince the mediator to ignore the function μ (recall that μ coincides with $\tilde{\mu}_1$ on T). Thus, player i does not gain by announcing a number z_i greater than one. Moreover, since μ is a communication equilibrium, it is optimal for i to reveal his type truthfully.

Consider now the action stage. Suppose that type t_i of player i sent a message $m_i \in M_i^*(t_i)$ and received a recommendation $a_i \in \bar{Q}_i(t_i)$. Given the trembles specified above, player i's beliefs $\phi_i(t_i, m_i, a_i)$ coincide with $\hat{\beta}_i(t_i, a_i)$ defined in equation (18). Clearly, this implies that it is sequentially rational for i to play a_i.

Proof of Theorem 4

Necessity

Fix an SSCE μ of a two-person Bayesian game G. Let $(G, M_1, M_2, S_1, S_2, \gamma)$ and (σ, ϕ) be the game with communication and the SE, respectively, that induce μ. We also let $\{\sigma_i^{M,k}\}_{k=1}^{\infty}$ denote the sequence of completely mixed message strategies of player $i = 1, 2$ (with $\lim_{k\to\infty} \sigma_i^{M,k} = \sigma_i^M$).

For each player i and each type $t_i \in T_i$, we let $M_i^0(t_i)$ denote the set message that are not played in equilibrium by type t_i. Formally:

$$M_i^0(t_i) = \left\{m_i \in M_i : \sigma_i^M(m_i|t_i) = 0\right\}.$$

We now use the sequence of mixed strategies $\{\sigma_i^{M,k}\}_{k=1}^{\infty}$ to partition $M_i^0(t_i)$ as follows. Two messages m_i and m_i' in $M_i^0(t_i)$ belong to the same element of the partition if and only if

$$\lim_{k\to\infty} \frac{\sigma_i^{M,k}(m_i|t_i)}{\sigma_i^{M,k}(m_i'|t_i)}$$

exists, is finite and different from zero. We let $\{M_i^0(t_i, 1), \ldots, M_i^0(t_i, r(t_i))\}$ denote the partition of $M_i^0(t_i)$.

For each player $i = 1, 2$, we let the set $\tilde{M}_i$ be equal to:

$$\tilde{M}_i = \{(t_i, z_i) : t_i \in T_i \text{ and } z_i = 0, \ldots, r(t_i)\},$$

and the set $\tilde{M}$ be equal to:

$$\tilde{M} = \left\{(t_i, z_i)_{i=1,2} \in \tilde{M}_1 \times \tilde{M}_2 : z_i = 0 \text{ for some } i = 1, 2\right\}.$$

Let the function $\tilde{\rho} : \tilde{M} \to \Delta(A)$ be given by:

$$\tilde{\rho}(a|(t_1,0),(t_2,0)) = \sum_{m\in M} \sum_{s\in S:\gamma(s|m)>0} \left(\prod_{i=1,2} \sigma_i^M(m_i|t_i)\right) \gamma(s|m) \left(\prod_{i=1,2} \sigma_i^A(a_i|t_i,m_i,s_i)\right),$$

and for every $t_j \in T_j$ and $z_j = 1,\dots,r(t_j)$,

$$\tilde{\rho}(a|(t_i,0),(t_j,z_j)) = \sum_{m_i\in M_i} \sum_{m_j\in M_j^0(t_j,z_j)} \sum_{s\in S:\gamma(s|m_i,m_j)>0} \{\sigma_i^M(m_i|t_i)$$

$$[\lim_{k\to\infty} \frac{\sigma_j^{M,k}(m_j|t_j)}{\sum_{m'_j\in M_j^0(t_j,z_j)} \sigma_j^{M,k}(m'_j|t_j)}] \gamma(s|m_i,m_j)\, \sigma_i^A(a_i|t_i,m_i,s_i)\, \sigma_j^A(a_j|t_j,m_j,s_j)\}.$$

For each player $i = 1,2$, we construct a function $\tilde{f}_i : \left\{(t_i,z_i) \in \tilde{M}_i : z_i \neq 0\right\} \to \mathbb{N}_{++}$ that satisfies the following condition. Consider any pair (t_i,z_i) and (t'_i,z'_i) (with $z_i = 1,\dots,r(t_i)$ and $z'_i = 1,\dots,r(t'_i)$). Let m_i be an element of $M_i^0(t_i,z_i)$ and m'_i be an element of $M_i^0(t'_i,z'_i)$. Suppose that the following limit exists:

$$\lim_{k\to\infty} \frac{\sigma_i^{M,k}(m_i|t_i)}{\sigma_i^{M,k}(m'_i|t'_i)}.$$

Then we have:

$$\begin{aligned}
\tilde{f}_i(t_i,z_i) &< \tilde{f}_i(t'_i,z'_i) \quad \text{if } \lim\nolimits_{k\to\infty} \tfrac{\sigma_i^{M,k}(m_i|t_i)}{\sigma_i^{M,k}(m'_i|t'_i)} = \infty \\
\tilde{f}_i(t_i,z_i) &> \tilde{f}_i(t'_i,z'_i) \quad \text{if } \lim\nolimits_{k\to\infty} \tfrac{\sigma_i^{M,k}(m_i|t_i)}{\sigma_i^{M,k}(m'_i|t'_i)} = 0 \\
\tilde{f}_i(t_i,z_i) &= \tilde{f}_i(t'_i,z'_i) \quad \text{otherwise}
\end{aligned}$$

We do not impose any restriction on the relationship between $\tilde{f}_i(t_i,z_i)$ and $\tilde{f}_i(t'_i,z'_i)$ when the limit defined above does not exist. Notice that the existence and the value of the limit does not depend on how we choose message m_i in $M_i^0(t_i,z_i)$ and message m'_i in $M_i^0(t'_i,z'_i)$.

The domain $\tilde{M}$ of the function $\tilde{\rho}$ defined above may be different from the domain $\bar{M}$ of the function ρ defined in equation (8). In particular, for some type t_i in T_i, the number $r(t_i)$ may be different from $n_j = |T_j|(|A_j|-1)$. For the same reason, the domain of $\tilde{f}_i$ may be different from the domain of f_i of Section 5. To complete the proof of the first part of the theorem, we now modify the domains of the functions $\tilde{\rho}$, $\tilde{f}_1$, and $\tilde{f}_2$ so that they match the domains of ρ, f_1, and f_2, respectively.

Consider player $i = 1, 2$, and type $t_i \in T_i$. Suppose that $r(t_i) \neq n_j$. There are two cases to consider depending on whether $r(t_i)$ is larger or smaller than n_j. First, suppose that $r(t_i) < n_j$. In this case, we simply add $n_j - r(t_i)$ new messages: $(t_i, r(t_i) + 1), \ldots, (t_i, n_j)$. Moreover, we assume that the function $\tilde{\rho}$ treats these new messages exactly as message $(t_i, 0)$. Formally, we assume:

$$\tilde{\rho}(a | (t_i, z_i), (t_j, 0)) = \tilde{\rho}(a | (t_i, 0), (t_j, 0)),$$

for every $z_i = r(t_i) + 1, \ldots, n_j$, for every $t_j \in T_j$ and every $a \in A$. Finally, we let $\tilde{f}_i(t_i, r(t_i) + 1), \ldots, \tilde{f}_i(t_i, n_j)$ be arbitrary positive integers.

We now turn to the case $r(t_i) > n_j$. Consider a message (t_i, z_i), with $z_i = 1, \ldots, r(t_i)$. We keep (t_i, z_i) in the set $\tilde{M}_i$ if and only if the following condition is satisfied (if the condition is violated we delete the message). There exists a pair $(t_j, a_j) \in T_j \times A_j$ such that: (i)

$$\sum_{t_i' \in T_i} \sum_{a_i \in A_i} p(t_i', t_j) \tilde{\rho}(a_i, a_j | (t_i', 0), (t_j, 0)) = 0;$$

(ii)

$$\sum_{a_i \in A_i} p(t_i, t_j) \tilde{\rho}(a_i, a_j | (t_i, z_i), (t_j, 0)) > 0;$$

and (iii) $\tilde{f}_i(t_i, z_i) \leqslant \tilde{f}_i(t_i, z_i')$ for every $z_i' = 1, \ldots, z_i - 1, z_i + 1, \ldots, r(t_i)$ such that

$$\sum_{a_i \in A_i} \tilde{\rho}(a_i, a_j | (t_i, z_i'), (t_j, 0)) > 0.$$

It follows from the definition of $\tilde{\rho}$ and $\tilde{f}_i$ that at most n_j messages can be kept in the set $\tilde{M}_i$.

Consider now the function $\tilde{\rho}$ defined over the set $\bar{M}$ and the function $\tilde{f}_i$ defined over the set $T_i \times \{1, \ldots, n_j\}$. It is easy to verify that $\tilde{\rho}$ is admissible and that $\tilde{\rho}$ and $\tilde{f} = \left(\tilde{f}_1, \tilde{f}_2\right)$ satisfy conditions (i)-(iv) of Theorem 4.

Sufficiency

Consider a two-person game G. Suppose that there exist an admissible function ρ and a function $f = (f_1, f_2)$ that satisfy conditions (i)-(iv) of Theorem 4. Let $\mu : \hat{T} \to \Delta(A)$ be such that $\rho((t_1, 0), (t_2, 0)) = \mu(t_1, t_2)$ for every $(t_1, t_2) \in \hat{T}$. We now show that μ is an SSCE of G.

Consider the following game with communication $\left(G, \bar{M}_1, \bar{M}_2, A_1, A_2, \gamma\right)$, where $\bar{M}_i$ is defined in equation (7) and $\gamma : \bar{M}_1 \times \bar{M}_2 \to \Delta(A)$ is defined as follows. Recall that the set $\bar{M}$ defined in equation (8) is a proper subset of $\bar{M}_1 \times \bar{M}_2$. For each $m = (m_1, m_2) \in \bar{M}$, we let $\gamma(m) = \rho(m)$. For $m \in \left(\bar{M}_1 \times \bar{M}_2\right) \setminus \bar{M}$, we let $\gamma(m)$ be an arbitrary element of $\Delta(A)$.

We claim that the game with communication just described admits an SE with the following features. Consider player i with type t_i. At the message stage he announces message $(t_i, 0)$. After announcing this message, type t_i obeys any recommendation he receives from the mediator. Also, suppose that type t_i sends (t_i, z_i), with $z_i \neq 0$, and receives a recommendation a_i such that

$$\sum_{t_j \in T_j} \sum_{a_j \in A_j} p(t_i, t_j) \rho(a_i, a_j | (t_i, z_i), (t_j, 0)) > 0.$$

In this case, type t_i obeys the recommendation and plays a_i. We do not specify the actions chosen by t_i in all other circumstances (see below). It is clear that the strategy profile described above induces the outcome μ.

It follows from condition (ii) of Theorem 4 that type t_i does not have an incentive to announce a message different from $(t_i, 0)$ or to disobey a recommendation that has positive probability when both players follow their equilibrium strategies at the message stage. Furthermore, condition (iii) of Theorem 4 implies that after sending (mistakenly) message (t_i, z_i), $z_i \neq 0$, type t_i does not want to disobey the recommendation to play an action a_i that has positive probability when the opponent follows his equilibrium strategy.

It remains to show that after sending message $(t_i, 0)$, type t_i has an incentive to obey the unexpected recommendations (i.e., those recommendations that have zero probability when the opponent follows his equilibrium strategy). To do this we need to construct a consistent system of beliefs. We consider the following sequence of mixed message strategies. Let $\{\varepsilon_k\}_{k=1}^{\infty}$ be a sequence of positive numbers in the unit interval converging to zero. Along the sequence, we assume that type t_i reports message (t_i, z_i), $z_i \neq 0$, with probability $(\varepsilon_k)^{f_i(t_i, z_i)}$. Moreover, type t_i announces any message (t_i', z_i'), $t_i' \neq t_i$ and $z_i = 0, \ldots, n_j$, with probability $(\varepsilon_k)^{n_j |T_i| + 1}$.

Consider type t_i and suppose that he sends message $(t_i, 0)$ and receives the unexpected recommendation a_i. The trembles described above guarantee that the beliefs of type t_i are given by $\beta_i(\cdot | t_i, a_i; \rho, f)$ (see equation (9)). It then follows from condition (iv) of Theorem 4 that it is optimal for type t_i to play action a_i.

References

[1] Aumann, R. J. (1974): "Subjectivity and Correlation in Randomized Strategies," *Journal of Mathematical Economics*, 1, 67-96.

[2] Aumann, R. J., and S. Hart (2003): "Long Cheap Talk," *Econometrica*, 71, 1619-1660.

[3] Barany, I. (1992): "Fair Distribution Protocols or How the Players Replace Fortune," *Mathematics of Operations Research*, 17, 327-340.

[4] Ben-Porath, E. (1998): "Correlation without Mediation: Expanding the Set of Equilibrium Outcomes by 'Cheap' Pre-play Procedures," *Journal of Economic Theory*, 80, 108-122.

[5] Ben-Porath, E. (2003): "Cheap Talk in Games with Incomplete Information," *Journal of Economic Theory*, 108, 45-71.

[6] Forges, F. (1986): "An Approach to Communication Equilibria," *Econometrica*, 54, 1375-1385.

[7] Forges, F. (1990): "Universal Mechanisms," *Econometrica*, 58, 1341-1364.

[8] Gerardi, D. (2004): "Unmediated Communication in Games with Complete and Incomplete Information," *Journal of Economic Theory*, 114, 104-131.

[9] Harsanyi, J. C. (1967-68): "Games with Incomplete Information Played by 'Bayesian' Players," *Management Science*, 14, 159-182, 320-334, 486-502.

[10] Kreps, D. M., and R. Wilson (1982): "Sequential Equilibria," *Econometrica*, 50, 863-894.

[11] Myerson, R. B. (1982): "Optimal Coordination Mechanisms in Generalized Principal-Agent Problems," *Journal of Mathematical Economics*, 10, 67-81.

[12] Myerson, R. B. (1986): "Multistage Games with Communication," *Econometrica*, 54, 323-358.

[13] Selten, R. (1975): "Re-examination of the Perfectness Concept for Equilibrium Point in Extensive Games," *International Journal of Game Theory*, 4, 25-55.

[14] Urbano, A., and J. Vila (2002): "Computational Complexity and Communication: Coordination in Two-Player Games," *Econometrica*, 70, 1893-1927.