

# Certifiable Pre-Play Communication: Full Disclosure\*

Jeanne HAGENBACH<sup>†</sup>    Frédéric KOESSLER<sup>‡</sup>    Eduardo PEREZ-RICHET<sup>§</sup>

October 9, 2012

## Abstract

This article asks when communication with certifiable information leads to complete information sharing. We consider Bayesian games augmented by a pre-play communication phase in which announcements are made publicly. We characterize the augmented games in which there exists a full disclosure sequential equilibrium with extremal beliefs (*i.e.*, any deviation is attributed to a single type of the deviator). This characterization enables us to provide different sets of sufficient conditions for full information disclosure that encompass and extend all known results in the literature, and are easily applicable. We use these conditions to obtain new insights in senders-receiver games, games with strategic complementarities, and voting with deliberation.

*Keywords:* Strategic Communication; Hard Information; Information Disclosure; Masquerade Relation; Belief Consistency; Single Crossing Differences; Supermodular Games.

*JEL classification:* C72; D82.

---

\*Thanks to be added.

<sup>†</sup>École Polytechnique – CNRS, e-mail: [jeanne.hagenbach@polytechnique.edu](mailto:jeanne.hagenbach@polytechnique.edu)

<sup>‡</sup>Paris School of Economics – CNRS, e-mail: [koessler@pse.ens.fr](mailto:koessler@pse.ens.fr)

<sup>§</sup>École Polytechnique, e-mail: [eduardo.perez@polytechnique.edu](mailto:eduardo.perez@polytechnique.edu)

# 1 Introduction

Before most individual or collective decisions, the concerned parties can communicate with each other and exchange information. The availability of communication may influence outcomes in important ways. This simple observation has given rise to a rich literature in game theory that aims at characterizing achievable equilibrium outcomes in strategic decision problems extended with communication (see, e.g., Myerson, 1994). In this paper, we adopt a different approach and try to understand when pre-play communication leads to *full disclosure* of privately held information, under the assumption that the players make certifiable statements.

More specifically, we consider a general Bayesian game augmented by a communication stage<sup>1</sup> at which players can publicly disclose information about their type before choosing their actions in a second stage. The messages exchanged in the communication phase deliver certifiable information, meaning that the players cannot lie, but may reveal their information only partially or not at all.<sup>2</sup>

In order to enforce full disclosure, players must be able to coordinate on second stage actions that deter any unilateral attempt to conceal information at the communication stage. To understand when this is possible, we define the *masquerade relation* which is a simple description of the incentives of a player with given private information (or type) to pretend that her information is different (*i.e.*, to masquerade as another type). This relation is easy to build. If in the communication phase each player fully certifies her type,<sup>3</sup> the game played at the action stage is a complete information game that depends on the type profile. Hence in a full disclosure equilibrium, each player expects to get the payoff associated with the equilibrium<sup>4</sup> of the complete information game that unfolds. If a player could convince all the others that her

---

<sup>1</sup>For most of the paper, we assume simultaneous communication, but in [Section 5](#) we show how to extend some of our results to sequential communication.

<sup>2</sup>The assumption of certifiable information has been introduced in sender-receiver games by Grossman (1981) and Milgrom (1981). See also Milgrom and Roberts (1986). It is also used in a branch of the mechanism design and implementation literature (see, e.g., Green and Laffont, 1986, Bull and Watson, 2004, 2007, Alger and Renault, 2006, Deneckere and Severinov, 2008, Ben-Porath and Lipman, 2012, Kartik and Tercieux, 2012).

<sup>3</sup>In most of the paper, we assume *own type certifiability*: each player can fully certify her true type. This assumption is relaxed in [Section 5](#).

<sup>4</sup>Uniqueness is assumed only in the introduction in order to simplify the exposition.

type is different from the truth, she might benefit by following up on her lie and best-responding to the misguided equilibrium that the other players coordinate on. If she actually benefits by masquerading as a certain target type, we say that her true type wants to masquerade as the targeted type. Note that with certifiable communication, a player cannot manage to convince other players that she is of a certain untrue type by directly lying about her type. So the masquerade relation is constructed on a somewhat hypothetical situation. However, when a player conceals some information about her type, the other players must assume that she does that in order to be perceived as different than she really is. The masquerade relation is best represented as a directed graph on the type set of each player,<sup>5</sup> such that an arrow points from one type to the other whenever the former wants to masquerade as the latter.

This summary of the players' incentives suggests a natural way to deter obfuscation at the communication stage. If a single player sends a message that does not fully certify her type, the other players should attribute this message to a *worst case type* among its possible senders, *i.e.*, a type that none of the other types who could have possibly sent this message wants to masquerade as. The idea of a worst case type was first introduced by Seidmann and Winter (1997) in a simple sender-receiver model, and captures the intuition of Milgrom (1981) that in order to enforce full disclosure the players should exercise skepticism. However, worst case types may fail to exist if there are cycles in the masquerade relation. In fact the acyclicity of the masquerade relation is equivalent to the existence of a worst case type for every subset of types, and is a sufficient condition for the existence of a full disclosure equilibrium. If the players can certify any true statement about their type, and if we restrict them to hold *extremal beliefs* off the equilibrium path, *i.e.*, beliefs that put probability one on a single type of a deviating player,<sup>6</sup> the acyclicity condition is necessary and sufficient.

---

<sup>5</sup>For most of the paper, we assume a finite type space for each player. We show in [Section 5](#) that our results extend to infinite type spaces.

<sup>6</sup>More precisely, when a player unilaterally deviates from full disclosure during the communication phase, we restrict our attention to beliefs such that every non-deviating player attributes the deviant message to a single type among its possible senders. We show that this restriction, combined with full support and strong belief consistency (Kreps and Wilson, 1982), imposes that the beliefs about the deviator are common to all non-deviators and do not depend on their types.

While apparently quite theoretical, this characterization is very useful to pin down sufficient conditions for full disclosure. The first of these conditions is monotonicity. If the masquerading payoff of a player is increasing in the type she masquerades as, the acyclicity condition is clearly satisfied. This is the case in the seller-buyer models of Milgrom (1981) and Grossman (1981) where a seller always prefers to appear as having a higher quality product. Most of the literature has followed in these steps by relying on a monotonicity condition in more complicated games (see Okuno-Fujiwara et al., 1990 and Van Zandt and Vives, 2007). Notable exceptions are Seidmann and Winter (1997), Giovannoni and Seidmann (2007) and Koessler and Renault (2012) who use different conditions in sender-receiver games. Koessler and Renault (2012) provide conditions for full disclosure of product information in seller-buyer relationships in which consumers' tastes for the product types are heterogeneous. Giovannoni and Seidmann (2007) rely on a combination of two conditions:<sup>7</sup> single-peakedness of the masquerading payoff in the target type, and a *no reciprocal masquerade* (in our terminology) condition ensuring that no two types want to masquerade as each other. We provide a simple and more general approach by showing that these two conditions prevent the existence of cycles in the masquerade relation, and we show how to identify worst case types.<sup>8</sup>

In many interesting games and economic problems, the single-peakedness or the monotonicity conditions are not satisfied. For instance, they are not satisfied in coordination games in which each player wants to be close to her ideal action and to the actions of other players, in games of influence in which each player wants to convince all others to choose her own ideal action, or in voting games such as the jury model with a non-unanimous voting rule. One of our main contributions is to show that the acyclicity condition is satisfied whenever the masquerading payoff has *single crossing differences*,<sup>9</sup> *i.e.*, if the return from masquerading as a higher target type is positive for a given true type, then it is also positive for higher true

---

<sup>7</sup>Up to minor details, the conditions in Giovannoni and Seidmann (2007) imply those of Seidmann and Winter (1997). In particular, the focus on the sign of the difference between the ideal actions of the sender and the receiver in Seidmann and Winter (1997) is unnecessary.

<sup>8</sup>This literature also gives results on uniqueness of a full disclosure equilibrium, under more restrictive conditions. The tools we develop here are tailored to provide existence results, so we will leave uniqueness aside.

<sup>9</sup>Or, therefore, *increasing differences*. The terminology adopted is that of Milgrom (2004).

types.<sup>10</sup>

The sufficient conditions mentioned so far bear on the expected masquerading payoffs at the *interim* stage, when the players only know their own type. It is often easier to verify conditions on the ex post masquerading payoffs. Ex post monotonicity implies interim monotonicity. The single-peakedness condition, on the other hand, is often difficult to aggregate. Another advantage of increasing and single crossing differences is that there are well known conditions under which they can be aggregated. We show in particular that if the ex post masquerading payoffs have increasing differences and the types are affiliated, then the interim masquerading payoffs have single crossing differences. This result is a simple corollary of Quah and Strulovici (2012), but the observation had never been made in the literature, and it could prove useful for comparative statics in general. We also mention how to directly apply the results of Quah and Strulovici (2012) to aggregate the single crossing property in our environment.

To illustrate our method, we provide a string of new applied results that contribute to different literatures. Our first result applies the single crossing differences condition to models with multiple senders and a single receiver and shows that if the optimal action of the receiver is nondecreasing in types, the preferences of the players have complementarities in own type and action, and the types are affiliated, then full disclosure is an equilibrium. The second result considers supermodular Bayesian games with complementarities between own actions and all types (as in Van Zandt and Vives, 2007). We show that if the types are affiliated and the preferences of the players also exhibit complementarities in own type and the actions of other players, then the interim masquerading payoffs satisfy the single crossing differences condition and there exists a full disclosure equilibrium.<sup>11</sup> This result implies that full disclosure is an equilibrium in the coordination and influence games already mentioned. Finally we contribute to the literature on deliberation before voting<sup>12</sup> by considering a general voting game that

---

<sup>10</sup>Clearly, this condition is written with respect to a certain linear order on the type set of the player, and the existence of this order is therefore part of the condition.

<sup>11</sup>This result is different from the result of Van Zandt and Vives (2007) which says that if the actions of others have positive or negative externalities, then there exists a full disclosure equilibrium.

<sup>12</sup>See, for example, Austen-Smith and Feddersen (2006), Gerardi and Yariv (2007), Jackson and Tan (2012), Lizzeri and Yariv (2011), Mathis (2011).

includes the jury model. The players vote between two alternatives such that for each player the difference in payoff between the alternatives is nondecreasing in the types of all players. We show that the ex post masquerading payoffs satisfy increasing differences for every non-unanimous rule, so that if types are affiliated, full disclosure is an equilibrium of the voting game preceded by a debate with hard information. The case of unanimity is even simpler since the monotonicity condition is then satisfied.

## 2 The Model

**The Base Game.** There is a set  $N = \{1, \dots, n\}$  of players who are to interact in a base game with action set  $A = A_1 \times \dots \times A_n$ . Each player  $i$  is privately informed about her type  $t_i \in T_i$ , where  $T_i$  is a finite set<sup>13</sup>, and  $T = T_1 \times \dots \times T_n$  is the type set. Let  $p(\cdot) \in \Delta(T)$  be a strictly positive (full support) common prior probability distribution over type profiles, and  $p(\cdot|t_i) \in \Delta(T_{-i})$  the interim belief of player  $i$  when she is of type  $t_i$ .<sup>14</sup> The preferences of the players are given by vNM utility functions  $u_i : A \times T \rightarrow \mathbb{R}$ ,  $i = 1, \dots, n$ .

Let  $\Gamma = \langle N, T, A, p, (u_i)_{i \in N} \rangle$  denote this Bayesian game. To every type profile  $t \in T$ , we can associate the complete information normal form game  $\tilde{\Gamma}(t) = \langle N, A, (u_i(\cdot, t))_{i \in N} \rangle$ . To avoid introducing additional conditions on  $\tilde{\Gamma}(t)$  we make the following assumption in all the paper:

**Assumption 1.** *For every type profile  $t \in T$ , the best reply correspondence of the game  $\tilde{\Gamma}(t)$  is well defined, and the set of Nash equilibria of  $\tilde{\Gamma}(t)$ , denoted by  $NE(t) \subseteq A$ , is nonempty.*<sup>15</sup>

**The Communication Phase.** Before choosing their actions in  $A$ , but after learning their types, the players have the opportunity to publicly and simultaneously<sup>16</sup> disclose hard evidence about their type at no cost. To formalize this, suppose that player  $i$  is restricted to send

<sup>13</sup>See [Section 5](#) for an extension to infinite type spaces.

<sup>14</sup>We assume a common prior to apply the notion of sequential equilibrium to a standard extensive form game. But the solution concept and our results can be readily extended to games with heterogeneous prior beliefs  $p_i(\cdot) \in \Delta(T)$  as long as  $p_i(\cdot|t_i) \in \Delta(T_{-i})$  has full support for every  $i \in N$  and  $t_i \in T_i$ .

<sup>15</sup>The set of actions  $A_i$  of each player  $i$  can be extended to the set  $\Delta(A_i)$  of mixed actions.

<sup>16</sup>We consider sequential communication protocols in [Section 5](#).

messages in a nonempty and finite set  $M_i(t_i)$  if her type is  $t_i$ . Let  $M_i = \bigcup_{t_i \in T_i} M_i(t_i)$  be the set of possible messages of player  $i$ , and  $M = M_1 \times \dots \times M_n$  the message space. A message  $m_i \in M_i$  provides hard evidence to other players that  $i$ 's type is in  $M_i^{-1}(m_i) := \{t_i \in T_i : m_i \in M_i(t_i)\}$ .

A subset  $S_i$  of  $T_i$  is certifiable if there exists a message  $m_i \in M_i$  such that  $M_i^{-1}(m_i) = S_i$ . For simplicity, we assume that the players can always certify their type by sending a message that no other type could send. This assumption is maintained throughout the paper, but we explain exactly when and how it can be relaxed in [Section 5](#).

**Assumption 2** (Own Type Certifiability). *For every player  $i \in N$  and every type  $t_i \in T_i$  of player  $i$ , there exists a message  $m_i \in M_i(t_i)$  such that  $M_i^{-1}(m_i) = \{t_i\}$ .*

It will be useful to consider the following property of the message correspondence which ensures that all players can certify any true statement about their type.

**Definition 1** (Full Certifiability). *The message correspondence satisfies full certifiability if for every player  $i \in N$ , every nonempty subset of her types  $S_i \subseteq T_i$  is certifiable.*

**Full Disclosure Equilibria with Extremal Beliefs.** Our equilibrium concept is the notion of sequential equilibrium of Kreps and Wilson (1982). It is defined as a profile of strategies and a belief system satisfying strong belief consistency and sequential rationality at every information set.<sup>17</sup> In the rest of the paper, the term equilibrium refers to this definition.

We are interested in equilibria of the augmented game in which all players perfectly reveal their type in the communication phase (henceforth, *full disclosure equilibria*). In a full disclosure equilibrium, the second stage game on the equilibrium path is a strict subgame corresponding to the complete information game  $\tilde{\Gamma}(t)$ , and therefore the action profile played on

---

<sup>17</sup>A pair of a strategy profile and a belief system is strongly consistent if it can be obtained as the limit of a completely mixed strategy profile and the corresponding belief system is obtained by Bayesian updating. Note that even though our extended game has two periods, the equivalence between perfect Bayesian equilibria and sequential equilibria (Fudenberg and Tirole, 1991) does not hold in general because we do not assume type independence. The fact that action sets  $(A_i)_{i \in N}$  may not be finite does not alter the definition of a sequential equilibrium because beliefs are only defined on finite information sets (over type and message profiles, but not over action profiles because the game ends after the action stage).

the equilibrium path must be in  $NE(t)$ . We choose a selection  $a^*(t)$  from  $NE(t)$  and reformulate our objective as finding conditions under which there exists a full disclosure equilibrium of the augmented game such that  $a^*(t)$  is played on the equilibrium path. Under the assumption of own type certifiability ([Assumption 2](#)), any full disclosure equilibrium is outcome equivalent to an equilibrium in which every player simply certifies her type along the equilibrium path. Without loss of generality, the paper therefore focuses on the existence of full disclosure equilibria on the equilibrium path in which every player  $i$  of type  $t_i$ , sends a message  $m_i$  such that  $M_i^{-1}(m_i) = \{t_i\}$ .<sup>18</sup>

In order to support such an equilibrium, the players must be able to punish any player  $i$  who sends a message  $m_i$  that does not fully certify her type. The other players have two levers to punish a deviator: (i) by forming appropriate beliefs about the type of the deviator within the restriction imposed by the hard evidence contained in  $m_i$ ; (ii) by coordinating on appropriate sequentially rational actions in the second stage. In order to make things tractable, we make two restrictions off the equilibrium path: one on beliefs and one on actions.

First, we restrict beliefs off the equilibrium path to be extremal in the sense of the following definition. It states that, after a unilateral deviation from full disclosure by a player  $i$ , every other player's beliefs about  $i$  are restricted to the extreme points of the convex set formed by the simplex  $\Delta(T_i)$ .

**Definition 2** (Extremal Beliefs). *A full disclosure equilibrium with extremal beliefs is a full disclosure equilibrium such that after any unilateral deviation, each player's beliefs assign probability one to a single type of the deviator.*

The second restriction concerns the second-stage equilibrium actions that can be played off the equilibrium path. To understand this restriction, suppose that player  $i$  unilaterally deviates from full disclosure and sends a message  $m_i$ , while every player  $j \neq i$  sends a message that certifies her true type  $t_j$ . Then, under extremal beliefs, all players must attribute a single type  $t'_i \in M_i^{-1}(m_i)$  to player  $i$ . The extremal belief assumption does not require all players

---

<sup>18</sup>Own type certifiability is relaxed in [Section 5](#).



other than  $i$  to attribute the same type  $t'_i$  to player  $i$ , but we will show in [Section 3](#) that this is required by strong consistency. Consequently, all non-deviators put probability one on the type profile  $(t'_i, t_{-i})$ . Then, sequential rationality requires that non-deviators play according to some action profile in  $NE(t'_i, t_{-i})$  but not necessarily  $a^*(t'_i, t_{-i})$ . We will consider only equilibria in which they do play according to  $a^*(t'_i, t_{-i})$ .

**Definition 3.** *We say that an equilibrium implements  $a^*(\cdot)$  on and off the equilibrium path if, whenever the second stage beliefs of all non-deviating players put probability one on a particular type profile  $t$ , all the non-deviating players play according to  $a^*(t)$ .*

Clearly, this restriction is without loss of generality when the complete information game  $\tilde{\Gamma}(t)$  has a unique equilibrium at every type profile  $t$ . It is also a natural assumption when there is a unique “reasonable” equilibrium of each  $\tilde{\Gamma}(t)$ . For example, if we consider a voting game with two alternatives, the unique reasonable equilibrium is one in which all voters vote for their preferred alternative. It is important to keep in mind that this restriction and the restriction on extremal beliefs only make it harder to find existence results. Therefore these restrictions should firstly be viewed as a methodological device that allows us to construct full disclosure equilibria.

**Some Definitions.** For clarity, we provide the precise definitions of several known concepts that play a role throughout the paper. To formulate these definitions, consider two partially ordered sets  $(X, \succeq)$  and  $(Y, \succeq)$ .<sup>19</sup>

**Definition 4** (Single-Peakedness). *Suppose that  $X$  is linearly ordered. A function  $f : X \rightarrow \mathbb{R}$  is single-peaked if whenever  $x \neq x'$  and  $f(x') > f(x)$ , then we have  $f(x') > f(x'')$  for every  $x''$  strictly between  $x$  and  $x'$ .<sup>20</sup>*

For the next three definitions, we adopt the terminology of [Milgrom \(2004\)](#).

---

<sup>19</sup>When there is no risk of confusion we use the same notation  $\succeq$  for orderings defined on different sets.

<sup>20</sup>It is easy to show that this definition is equivalent to the more usual definition according to which  $f(x)$  must be strictly increasing, then possibly constant at its peak, and then strictly decreasing.

**Definition 5** (Single Crossing). A function  $f : X \rightarrow \mathbb{R}$  is single crossing if for every  $x \preceq x'$ ,

$$f(x) \geq (>) 0 \Rightarrow f(x') \geq (>) 0$$

**Definition 6** (Increasing Differences). A function  $g : X \times Y \rightarrow \mathbb{R}$  has increasing differences if for every  $x \preceq x'$  and  $y \preceq y'$  we have

$$g(x', y) - g(x, y) \leq g(x', y') - g(x, y'),$$

that is if for every  $x \preceq x'$ , the difference function  $\Delta(y) = g(x', y) - g(x, y)$  is nondecreasing.

**Definition 7** (Single Crossing Differences). A function  $g : X \times Y \rightarrow \mathbb{R}$  has single crossing differences in  $(x, y)$  if for every  $x \preceq x'$ , the difference function  $\Delta(y) = g(x', y) - g(x, y)$  is single crossing.

Note that while the definition of increasing differences is symmetric, this is not the case for the definition of single crossing differences. The last definition below is due to Quah and Strulovici (2012). They show that it is a necessary and sufficient condition for aggregating the single crossing property.

**Definition 8** (Signed-Ratio Monotonicity). A pair of functions  $f : X \rightarrow \mathbb{R}$  and  $g : X \rightarrow \mathbb{R}$  satisfies signed-ratio monotonicity if they satisfy the following conditions:

(i) For every  $x$  such that  $g(x) < 0$  and  $f(x) > 0$ , and for every  $x' \succ x$ :

$$-\frac{g(x)}{f(x)} \geq -\frac{g(x')}{f(x')};$$

(ii) For every  $x$  such that  $f(x) < 0$  and  $g(x) > 0$ , and for every  $x' \succ x$ :

$$-\frac{f(x)}{g(x)} \geq -\frac{f(x')}{g(x')}.$$

### 3 A Characterization

In this section, we provide a set of necessary and sufficient conditions for the existence of a full disclosure equilibrium with extremal beliefs that implements  $a^*(\cdot)$  on and off the equilibrium path. The first part of the section discusses the consequences of our restrictions on equilibrium beliefs. In summary, it shows that in a full disclosure equilibrium with extremal beliefs, after any unilateral deviation from full disclosure by player  $i$  sending message  $m_i$ , all the non-deviating players must hold the same belief about player  $i$ , and that this belief is independent of the true type profile  $t_{-i}$  of the non-deviators. Knowing that, a reader may skip the discussion that follows and jump to the presentation of the masquerade relation.

**Consistent Beliefs.** First, we examine an implication on beliefs of the requirement of strong consistency in a full disclosure equilibrium. In such an equilibrium, any deviation is perfectly detectable. When a player  $i$  unilaterally deviates in the communication stage, all the players know the true type profile  $t_{-i}$  of the non-deviators. Then strong consistency implies that the equilibrium belief of a non-deviator about the type of player  $i$  must be the limit of Bayes-consistent beliefs generated by trembles of player  $i$  that put positive probability on all her available messages ( $M_i(t_i)$  for a player  $i$  of type  $t_i$ ). Then, since all non-deviators have a common prior and the same information (the profile  $t_{-i}$ , and the message  $m_i$ ), the beliefs generated by the trembles must be the same for all non-deviators. To summarize, we have the following observation.

**Observation 1.** In a full disclosure equilibrium, after any unilateral deviation of some player  $i$  in the communication stage, all players different from  $i$  share the same belief over type profiles.

The following lemma shows that if we add the restriction of extremal beliefs, it must also be true that the belief  $\mu$  about  $i$  commonly held by the non-deviators following a unilateral deviation by  $i$  is independent of their true type profile  $t_{-i}$ . The intuition is that if this belief  $\mu$  is extremal so that it puts probability one on a type  $t'_i$ , then the sequence of Bayes-consistent

beliefs  $\mu^k$  that converges to  $\mu$  must have been generated by completely mixed strategies of player  $i$  that put infinitely more weight on  $m_i$  when she is of type  $t'_i$  than when she is of any other type  $t''_i$ . But if this is the case, this crowds out any information about  $i$  contained in the prior, and in particular any information that the non-deviators could derive from the correlation between  $t_{-i}$  and  $t_i$ . Note however that the full support assumption on type profile is fundamental for this property to hold. The lemma also shows that any pair of a strategy profile and a belief system that satisfies full type certification and extremal beliefs with the restrictions above also satisfies strong consistency. This is important as it implies that we do not have to worry about checking that strong consistency is satisfied for the rest of the paper.

**Lemma 1** (Consistent Extremal Beliefs). *In a full disclosure equilibrium with extremal beliefs in which players fully certify their type, after any unilateral deviation of some player  $j$  in the communication stage, the beliefs of every player  $i \neq j$  assign probability one to the same type  $t_j \in T_j$  of player  $j$  independently of player  $i$  and the actual type profile  $t_{-j}$ .<sup>21</sup> Furthermore, if a pair of a strategy profile and a belief system satisfies full type certification, and is such that after a unilateral deviations from full disclosure the players have extremal beliefs that are common to all the non-deviators and do not depend on their type profile, then it can be extended at other information sets so as to satisfy strong consistency.*

*Proof.* See the Appendix. □

**The Masquerade Relation.** As Seidmann and Winter (1997) already noticed in the sender-receiver case, the key to discouraging obfuscation is to attribute any message  $m_i$  to a type  $s_i$  in the set  $M_i^{-1}(m_i)$  of its possible senders such that none of the other types in  $M_i^{-1}(m_i)$  would like to masquerade as  $s_i$ . This naturally leads us to investigate when a type  $t_i$  would like to

---

<sup>21</sup>The restriction imposed by the sequential equilibrium in the lemma also follows from the “strategic independence principle” (Battigalli, 1996), and it is explicitly required under the “no-signaling-what-you-don’t-know” condition in the definition of a perfect Bayesian equilibrium when types are independently distributed (see Fudenberg and Tirole, 1991).

masquerade as another type  $s_i$ . For this purpose, let

$$v_i(t_i|t_i) = E_{t_{-i}}(u_i(a^*(t), t) | t_i),$$

denote the expected utility of player  $i$  on the equilibrium path of a full disclosure equilibrium if she is of type  $t_i$ , and

$$v_i(s_i|t_i) = E_{t_{-i}}(u_i(BR_i(a_{-i}^*(s_i, t_{-i}), t), a_{-i}^*(s_i, t_{-i}), t) | t_i),$$

the utility that she would obtain by masquerading as  $s_i$ . In the remainder of the paper, the following notation for the utility in the expectation will be useful:

$$v_i(s_i|t_i; t_{-i}) = u_i(BR_i(a_{-i}^*(s_i, t_{-i}), t), a_{-i}^*(s_i, t_{-i}), t).$$

We call  $v_i(s_i|t_i)$  and  $v_i(s_i|t_i; t_{-i})$  the *interim* and *ex post masquerading payoff functions*.

**Definition 9** (Masquerade). *We say that  $t_i$  wants to masquerade as  $s_i$ , denoted by  $t_i \xrightarrow{\mathcal{M}} s_i$ , whenever  $v_i(s_i|t_i) > v_i(t_i|t_i)$ .*

This defines the masquerade relation  $\xrightarrow{\mathcal{M}}$  on  $T_i$ . Note that it is by definition irreflexive ( $t_i \xrightarrow{\mathcal{M}} s_i \Rightarrow t_i \neq s_i$ ), but generally not transitive. We can use this relation to define a *worst case type* for  $S_i \subseteq T_i$  as a type in  $S_i$  that no other type from  $S_i$  would like to masquerade as:

$$wct_i(S_i) := \{s_i \in S_i \mid \nexists t_i \in S_i, t_i \xrightarrow{\mathcal{M}} s_i\}.$$

Of course, this set may be empty, or have more than one element. Also note that if  $s_i$  is a worst-case type for  $S_i$ , then it is also a worst-case type for every subset of  $S_i$  containing  $s_i$ .

If we think of the masquerade relation as an oriented graph on  $T_i$ , a worst-case type for  $S_i$  is an element  $s_i \in S_i$  with no incoming arrow from another element of  $S_i$ . Therefore, if the masquerade relation forms a cycle on the elements of  $S_i$ , there exists no worst-case type for  $S_i$ .

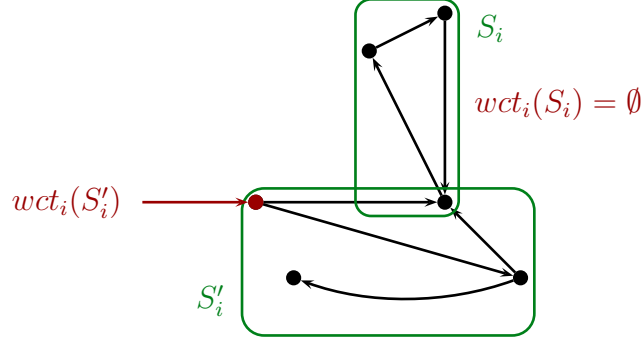


Figure 1: Masquerade relation and worst case types.

And if the masquerade relation has no cycles at all on  $T_i$ , it is easy to find a worst-case type for every subset of  $T_i$ . These intuitions are illustrated in Figure 1. They lead to the following characterization under full certifiability.<sup>22</sup>

**Theorem 1** (Characterization). *Suppose that the message correspondence satisfies full certifiability. Then the following statements are equivalent:*

- (i) *There exists a full disclosure equilibrium with extremal beliefs that implements  $a^*(\cdot)$  on and off the equilibrium path.*
- (ii) *For every  $i$  and every nonempty subset  $S_i \subseteq T_i$ , there exists a worst-case type for  $S_i$ :  $wct_i(S_i) \neq \emptyset$ .*
- (iii) *For every  $i$  the masquerade relation  $\xrightarrow{M}$  on  $T_i$  is acyclic.*

*If full certifiability is not assumed, (ii) and (iii) are still equivalent and they imply (i), which is equivalent to the existence of a worst case type on every certifiable subset of types.*

*Proof.* To show that (ii) implies (i), just notice that non-deviators can sanction any message  $m_i$  that does not fully disclose the type of player  $i$  by the belief that it must have come from a type in  $wct_i(M_i^{-1}(m_i))$ , therefore making unattractive any attempt of  $i$  to obfuscate her type.

<sup>22</sup>Hagenbach and Koessler (2012) gave the first definition of this relation and explored some of its implication. More precisely, they showed that (iii)  $\Rightarrow$  (ii)  $\Rightarrow$  (i). Our theorem extends these results to propose a characterization of full disclosure equilibria with extremal beliefs.

To show that (i) implies (ii), suppose that (ii) does not hold. Then there exists a set  $S_i \subseteq T_i$  such that  $wct_i(S_i) = \emptyset$ . Because of the full certifiability assumption, there exists a message  $m_i$  available to any type in  $S_i$  that certifies  $S_i$ . When receiving message  $m_i$  from  $i$ , the other players must assign it to some type in  $S_i$ , say  $s_i$ . But since  $wct_i(S_i) = \emptyset$ , there exists a type  $t_i \in S_i$  such that  $t_i \xrightarrow{\mathcal{M}} s_i$ . Then  $t_i$  would always deviate from the equilibrium path and send the message  $m_i$  in order to masquerade as  $s_i$ . Next, we show that (ii)  $\Leftrightarrow$  (iii). Suppose that  $\xrightarrow{\mathcal{M}}$  has a cycle  $t_i^1 \xrightarrow{\mathcal{M}} \dots \xrightarrow{\mathcal{M}} t_i^k \xrightarrow{\mathcal{M}} t_i^1$  on  $T_i$ . Then clearly  $S_i = \{t_i^1, \dots, t_i^k\}$  is such that  $wct_i(S_i) = \emptyset$ . Now suppose that there exists  $S_i \subseteq T_i$  such that  $wct_i(S_i) = \emptyset$ . Let  $s_i^1 \in S_i$ . Because  $wct_i(S_i) = \emptyset$  there exists  $s_i^2 \in S_i$  such that  $s_i^2 \xrightarrow{\mathcal{M}} s_i^1$ , but there also exists  $s_i^3 \in S_i$  such that  $s_i^3 \xrightarrow{\mathcal{M}} s_i^2$ . If  $s_i^3 = s_i^1$ , we have a cycle and we can conclude. Otherwise we can keep doing this until we obtain a cycle. This must happen eventually since  $S_i$  is finite.  $\square$

The equivalence between (i) and (ii) is essentially the same argument as in Seidmann and Winter (1997). The acyclicity condition proves to be quite useful.<sup>23</sup> Indeed, the following section shows that there are natural properties of the functions  $v_i(\cdot)$  that are often satisfied in usual games and imply acyclicity.

The existence of worst case types is not necessary if non-extremal beliefs are allowed out of the equilibrium path. To see that consider the following example.

**Example 1.** Consider an informed sender whose type is in  $T = \{t^1, t^2\}$  and an uninformed receiver who takes an action  $a \in \{a^1, a^2, a^3\}$ . The following table gives, for every pair  $(a, t)$  the payoffs of the sender and the receiver  $(u_S(a, t), u_R(a, t))$ . The receiver's ideal actions are  $a^1$

	$a^1$	$a^2$	$a^3$
$t^1$	(1,3)	(2,0)	(0,2)
$t^2$	(2,0)	(1,3)	(0,2)

under  $t^1$  and  $a^2$  under  $t^2$ , implying the following cycle in the masquerade relation  $t^1 \xrightarrow{\mathcal{M}} t^2 \xrightarrow{\mathcal{M}} t^1$ .

<sup>23</sup>It is also equivalent to the existence for every  $S_i \subseteq T_i$  of a “narcissistic type” that does not want to masquerade as any other type in  $S_i$ . Interestingly, it means that the existence of a worst case type for every subset of types is equivalent to the existence of a narcissistic type for every subset of types.

Therefore the set  $T$  has no worst case type, and there is no full disclosure equilibrium with extremal beliefs. However if any message certifying  $T$  gives rise to the belief that puts equal probability on  $t^1$  and  $t^2$ , the receiver takes action  $a^3$ , dissuading the sender from sending such a message.

## 4 Sufficient Conditions

### 4.1 Sufficient Conditions on Interim Masquerading Payoffs

In the following theorem we provide a list of sufficient conditions for the existence of a full disclosure equilibrium with extremal beliefs under full certifiability. Clearly, if a full disclosure equilibrium exists under this assumption, then it also exists when some subsets of types cannot be certified. Indeed an equilibrium under full certifiability provides ways to punish any message that is not a singleton, and these punishments are still valid when the message space is not as rich as long as own type certifiability is satisfied. (MON) stands for *Monotonicity* and (DM) for *Directional Masquerade*. (ID) and (SCD) stand for *Increasing Differences* and *Single Crossing Differences*. Finally (SP-NRM) is a set of two conditions, *Single Peakedness* and *No Reciprocal Masquerade*.

**Theorem 2** (Sufficient Conditions). *There exists a full disclosure equilibrium with extremal beliefs that implements  $a^*(\cdot)$  on and off the equilibrium path whenever for every  $i$  there exists a linear order  $\succeq$  on  $T_i$  such that either of the following conditions is satisfied:*

**(MON)**  $v_i(s_i|t_i)$  is nondecreasing in  $s_i$ .

**(DM)**  $v_i(s_i|t_i) > v_i(t_i|t_i) \Rightarrow s_i \succ t_i$ .

**(ID)**  $v_i(s_i|t_i)$  has increasing differences in  $(s_i, t_i)$ .

**(SCD)**  $v_i(s_i|t_i)$  has single crossing differences in  $(s_i, t_i)$ .



**(SP-NRM)**  $v_i(s_i|t_i)$  is single-peaked in  $s_i$  and satisfies the following no reciprocal masquerade condition:

$$v_i(s_i|t_i) > v_i(t_i|t_i) \Rightarrow v_i(s_i|s_i) \geq v_i(t_i|s_i).$$

*Proof.* First note that **(MON)** implies **(DM)** (up to a reversal of the order on  $T_i$  in case  $v_i(\cdot|t_i)$  is nonincreasing). Let  $S_i$  be any subset of  $T_i$ , and  $s_i^0 = \min S_i$ . Then for every  $s_i \in S_i$ ,  $v_i(s_i^0|s_i) \leq v_i(s_i|s_i)$ , implying that  $s_i^0 \in wct_i(S_i)$ .

For the next conditions, we start by noting that **(ID)** implies **(SCD)**. Then we first show that **(SCD)** implies that  $\xrightarrow{\mathcal{M}}$  has no 2-cycle. Suppose by contradiction that there exists a 2-cycle  $t_i^1 \xrightarrow{\mathcal{M}} t_i^2 \xrightarrow{\mathcal{M}} t_i^1$ . To fix ideas, suppose that  $t_i^1 \preceq t_i^2$  (we can do this because  $T_i$  is linearly ordered). Then we have a contradiction with **(SCD)**:

$$v(t_i^2|t_i^1) - v(t_i^1|t_i^1) > 0 > v(t_i^2|t_i^2) - v(t_i^1|t_i^2),$$

where the two inequalities come from the masquerade relation.

Now suppose that there exists a longer cycle  $t_i^1 \xrightarrow{\mathcal{M}} \dots \xrightarrow{\mathcal{M}} t_i^k \xrightarrow{\mathcal{M}} t_i^1$ . Because  $T_i$  is linearly ordered, the set  $\{t_i^1, \dots, t_i^k\}$  admits a minimal element with respect to  $\succeq$ . To fix ideas, let  $t_i^1$  be that minimal element. Then we have

$$v(t_i^2|t_i^1) - v(t_i^1|t_i^1) > 0 \quad \text{and} \quad v(t_i^1|t_i^k) - v(t_i^k|t_i^k) > 0,$$

from the fact that  $t_i^1 \xrightarrow{\mathcal{M}} t_i^2$  and  $t_i^k \xrightarrow{\mathcal{M}} t_i^1$ . And by **(SCD)** this implies  $v(t_i^2|t_i^k) - v(t_i^1|t_i^k) > 0$ , so

$$v(t_i^2|t_i^k) - v(t_i^k|t_i^k) = v(t_i^2|t_i^k) - v(t_i^1|t_i^k) + v(t_i^1|t_i^k) - v(t_i^k|t_i^k) > 0.$$

Therefore  $t_i^2 \xrightarrow{\mathcal{M}} \dots \xrightarrow{\mathcal{M}} t_i^k \xrightarrow{\mathcal{M}} t_i^2$  forms a cycle of length  $k - 1$ . By doing this over and over we end up with a 2-cycle which we already ruled out. To conclude, we have shown that  $\xrightarrow{\mathcal{M}}$  is acyclic.

For **(SP-NRM)**, note that the no reciprocal masquerade condition means that  $\xrightarrow{\mathcal{M}}$  has no 2-cycle. Let  $t_i^1 \xrightarrow{\mathcal{M}} \dots \xrightarrow{\mathcal{M}} t_i^k \xrightarrow{\mathcal{M}} t_i^1$  denote a longer cycle,  $k \geq 3$ . Let  $C = \{t_i^1, \dots, t_i^k\} \subseteq T_i$  be the set of elements of that cycle. We adopt the notation that  $t_i^{k+1} = t_i^1$ . There must be at least one  $j \in \{1, \dots, k\}$  such that  $t_i^{j+1}$  is not the immediate successor of  $t_i^j$  in  $(C, \succeq)$ . That is, there exists  $\ell \notin \{j, j+1\}$  such that  $t_i^j \prec t_i^\ell \prec t_i^{j+1}$  or  $t_i^{j+1} \prec t_i^\ell \prec t_i^j$ . Then the single-peakedness condition enables to deduce from  $v_i(t_i^{j+1}|t_i^j) > v_i(t_i^j|t_i^j)$  that  $v_i(t_i^\ell|t_i^j) > v_i(t_i^j|t_i^j)$ , that is  $t_i^j \xrightarrow{\mathcal{M}} t_i^\ell$ . Hence there exists a cycle without  $t_i^{j+1}$

$$t_i^j \xrightarrow{\mathcal{M}} t_i^\ell \xrightarrow{\mathcal{M}} \dots \xrightarrow{\mathcal{M}} t_i^j,$$

of length  $k' < k$ . But then, by repeating this operation, we eventually obtain a 2-cycle, thus contradicting the no reciprocal masquerade condition.  $\square$

Note that **(MON)** implies **(DM)** and **(NRM)**, strict monotonicity implies **(SCD)** and **(SP)**, and **(ID)** implies **(SCD)**. Except for these implications, we can construct examples showing that the different conditions are mutually exclusive, even up to a reordering of  $T_i$ .<sup>24</sup>

Most of the literature on disclosure of hard information is based on **(MON)**. When it is satisfied, every type would like to masquerade as the highest possible type. This is the case in the seller-buyer models of Grossman (1981) and Milgrom (1981). The seller's payoff is increasing in the perceived quality of her product. Then the buyer can interpret every announcement of the seller skeptically as coming from the lowest quality seller consistent with the announcement. This skeptical behavior leads to full disclosure. Another typical example mentioned in Okuno-Fujiwara et al. (1990) is a linear Cournot game with homogeneous goods and privately known marginal costs, in which the equilibrium payoff of a firm decreases when its competitors form higher beliefs about its cost.

**(DM)** is a weaker condition. When it holds, a player of any given type only wants to

---

<sup>24</sup>This should not be too surprising since **(ID)** or **(SCD)** are conditions that bear on all the functions  $(v_i(\cdot|t_i))_{t_i \in T_i}$  jointly, while **(MON)** or **(DM)** bear on each of these functions separately. **(SP-NRM)** consists of both a condition of single-peakedness for each function  $v_i(\cdot|t_i)$  and a condition of no reciprocal masquerade bearing on them together.

masquerade as a higher type. The sender-receiver game of Crawford and Sobel (1982) satisfies this property. Indeed the sign of the difference between the ideal actions of the sender and the receiver is independent of the sender's type. If the sender's ideal action is, say, always higher than the receiver's, a sender will only ever want to masquerade as a higher type. This does not mean, however, that she would like to masquerade as any higher type.

Giovannoni and Seidmann (2007) use (SP-NRM) in a sender-receiver model.<sup>25</sup> They use the term single crossing property for (NRM). We use the terminology of Milgrom (2004) that seems more descriptive and avoids any confusion. In the proof, we show that (SP-NRM) implies acyclicity in two steps. First, (NRM) rules out the existence of cycles of size two. Second, (SP) implies that any type  $t_i$  who wants to masquerade as another type  $s_i$  also wants to masquerade as any type between  $t_i$  and  $s_i$ . This betweenness property implies that any cycle of size higher than three contains a smaller cycle.

To our knowledge, this paper is the first to show that (SCD) and (ID) are sufficient conditions for the existence of full disclosure equilibria. When (ID) holds, the return of masquerading as a higher type increases with one's true type. When (SCD) holds, if the return of masquerading as a higher type is positive for  $t_i$ , then it is also positive for  $t'_i \succeq t_i$ . As for (SP-NRM), the proof first uses (SCD) to rule out cycles of size two, and then to reduce cycles of size bigger than two to smaller cycles.

## 4.2 Identifying a Worst-Case Type

When (MON) or (DM) holds, it is easy to see that in any subset of types  $S_i$ , the lowest type for the linear order is a worst-case type. To prove the results of Theorem 2 about (ID), (SCD) and (SP-NRM), we showed that either of these conditions implies that the masquerade relation is acyclic, thus leaving open the question of identifying a worst-case type. The following result answers this question and sheds some light on how to be skeptical under (SCD) or (SP-NRM).

---

<sup>25</sup>See Section 6.1 for a further explanation of the link between Giovannoni and Seidmann (2007) and our results, and new results on senders-receiver models.

**Proposition 1** (Worst-Case Types Identification).

(a) Suppose that  $v_i(s_i|t_i)$  satisfies (SCD). Let  $S_i \subseteq T_i$ , and  $\underline{t}_i = \min S_i$ . Then

$$\emptyset \neq \arg \min_{s_i \in S_i} v_i(s_i|\underline{t}_i) \subseteq wct_i(S_i).$$

(b) Suppose that  $v_i(s_i|t_i)$  satisfies (SP-NRM). Let  $S_i \subseteq T_i$ ,  $\underline{s}_i^0 = \min S_i$  and  $\bar{s}_i^0 = \max S_i$ .

Consider the sequences  $\{\underline{s}_i^k\}_{k=0}^\infty$  and  $\{\bar{s}_i^k\}_{k=0}^\infty$  defined by

$$\underline{s}_i^{k+1} = \sup\left(\{\underline{s}_i^k\} \cup \{t_i \in S_i \mid v_i(\underline{s}_i^k|t_i) > v_i(t_i|t_i)\}\right)$$

and

$$\bar{s}_i^{k+1} = \inf\left(\{\bar{s}_i^k\} \cup \{t_i \in S_i \mid v_i(\bar{s}_i^k|t_i) > v_i(t_i|t_i)\}\right).$$

Then the sequences  $\{\underline{s}_i^k\}$  and  $\{\bar{s}_i^k\}$  converge to some limits  $\underline{s}_i \in S_i$  and  $\bar{s}_i \in S_i$  (respectively) such that  $\underline{s}_i, \bar{s}_i \in wct_i(S_i)$ .

*Proof.* See the Appendix. □

For (ID) or (SCD), we focus on the lowest type  $\underline{t}_i$  of  $S_i \subseteq T_i$ , and choose any type  $\underline{s}_i$  in  $S_i$  that minimizes the masquerading payoff of  $\underline{t}_i$ . Then (SCD) implies that no other type in  $S_i$  wants to masquerade as  $\underline{s}_i$ , which is therefore a worst-case type for  $S_i$ . For (SP-NRM), observe that by construction no type in  $S_i$  that is smaller than  $\bar{s}_i$  wants to masquerade as  $\bar{s}_i$ . In addition, if a type  $s_i$  in  $S_i$  that is higher than  $\bar{s}_i$  wants to masquerade as  $\bar{s}_i$ , then we would have  $\bar{s}_i^{k+1} \prec s_i \preceq \bar{s}_i^k$  for some  $k$ , so by single-peakedness we would have  $\bar{s}_i^{k+1} \xrightarrow{\mathcal{M}} s_i$  and  $s_i \xrightarrow{\mathcal{M}} \bar{s}_i^{k+1}$ , a contradiction of no reciprocal masquerade.

### 4.3 Sufficient Conditions on Ex Post Masquerading Payoffs

Theorem 2 provides sufficient conditions on the interim masquerading payoff functions  $v_i(s_i|t_i)$ .

We now aim to provide conditions on the type distribution and the ex post masquerading

payoffs  $v_i(s_i|t_i; t_{-i})$ . For that purpose we define the following notion:

**Definition 10.** *The beliefs satisfy **type affiliation** if each  $T_i$  is linearly ordered so that  $T$  forms a lattice and  $p(t_{-i}|t_i)$  is log-supermodular.*

The case of affiliated types includes the case in which types are independent conditional on a random variable that is affiliated with each of them, and the case in which types are independent.

Since  $v_i(s_i|t_i) = \sum_{t_{-i} \in T_{-i}} p(t_{-i}|t_i) v_i(s_i|t_i; t_{-i})$ , our objective is to find conditions under which monotonicity, directional masquerade, increasing differences or single crossing differences of the ex post masquerading payoff  $v_i(s_i|t_i; t_{-i})$  can be aggregated so that the interim masquerading payoff  $v_i(s_i|t_i)$  can satisfy one of the sufficient conditions in [Theorem 2](#). This is a well known problem (see, for example, Athey, 2002), and we use the results of Quah and Strulovici (2012) who provide conditions under which the single crossing property is stable under aggregation.

**Theorem 3** (Ex Post Sufficient Conditions). *There exists a full disclosure equilibrium with extremal beliefs that implements  $a^*(\cdot)$  on and off the equilibrium path whenever for every  $i$  there exists a linear order  $\succeq$  on  $T_i$  such that either of the following conditions is satisfied:*

- (i)  $v_i(s_i|t_i; t_{-i})$  is nondecreasing in  $s_i$ .
- (ii)  $v_i(s_i|t_i; t_{-i})$  satisfies **ex post directional masquerade**:  $v_i(s_i|t_i; t_{-i}) > v_i(t_i|t_i; t_{-i}) \Rightarrow s_i \succ t_i$ .
- (iii)  $v_i(s_i|t_i; t_{-i})$  has single crossing differences in  $(s_i, t_i)$ ; for every  $s_i, s'_i$ , for every  $t_{-i}$  and  $t'_{-i}$ , the pair of difference functions  $\Delta(t_i; t_{-i}) = v_i(s'_i|t_i; t_{-i}) - v_i(s_i|t_i; t_{-i})$  and  $\Delta(t_i; t'_{-i}) = v_i(s'_i|t_i; t'_{-i}) - v_i(s_i|t_i; t'_{-i})$  satisfies signed-ratio monotonicity in  $t_i$ ; the beliefs satisfy type affiliation.
- (iv)  $v_i(s_i|t_i; t_{-i})$  has increasing differences in  $(s_i, t_i)$ ; the beliefs satisfy type affiliation.

*Proof.* See the Appendix. □

Point (iii) is a direct consequence of Quah and Strulovici (2012, Theorem 1) which states that any positive linear combination of a family of single crossing functions is also a single crossing function as long as the members of the family satisfy the signed-ratio monotonicity pairwise. Type affiliation enables to extend the property of signed-ratio monotonicity satisfied by the pair of single crossing functions  $\Delta(t_i; t_{-i})$  and  $\Delta(t_i; t'_{-i})$  to the pair of single crossing functions  $p(t_{-i}|t_i)\Delta(t_i; t_{-i})$  and  $p(t'_{-i}|t_i)\Delta(t_i; t'_{-i})$ . Under type affiliation the interim masquerading payoff then satisfies (SCD) in  $(s_i, t_i)$  which leads to full disclosure.

The proof of point (iv) establishes that if the ex post masquerading payoff satisfies (ID) in  $(s_i, t_i)$ , then the pair of functions  $\Delta(t_i; t_{-i})$  and  $\Delta(t_i; t'_{-i})$  of  $t_i$  satisfies signed-ratio monotonicity for every pair  $t_{-i}, t'_{-i}$  in  $T_{-i}$ . The argument actually makes a general point that might be interesting in applications, and in particular for comparative statics: if a function  $f(x, y)$  is single crossing and nondecreasing in  $y$  and  $(X, Y)$  are affiliated random vectors, then  $E(f(X, Y)|y)$  is single crossing in  $y$ . This comes as a simple corollary of Quah and Strulovici (2012, Theorem 1) and extends a result of Milgrom and Weber (1982) who show that if  $f(\cdot)$  is nondecreasing in  $x$  and  $y$ , then  $E(f(X, Y)|y)$  is nondecreasing in  $y$ .

## 5 Extensions

**Sequential Communication.** So far, we have assumed simultaneous communication. It would clearly be unappealing if perturbations in the communication protocol destroyed the full disclosure equilibrium. Fortunately, we can show that our most useful result for applications, the result of Theorem 3, also holds for (public) sequential communication protocols.

A sequential communication protocol is a protocol in which all players are called sequentially, in a possibly random order, and possibly several at the same time, to make a public announcement about their type, such that when a player is called, she learns the identity of all the players that have been called before her and the content of their message. We now consider the Bayesian game  $\Gamma$  augmented by such a communication protocol, so that the players

simultaneously choose their actions once the communication phase has ended.

We show that under any sequential communication protocol, if any of the sufficient conditions in [Theorem 3](#) is satisfied, there exists a full disclosure equilibrium of the augmented game such that  $a^*(t)$  is played on the equilibrium path. To understand the key intuition behind this result, let  $J$  and  $K$  form a partition of the set  $N \setminus \{i\}$  of all players except  $i$ , let  $t_J$  denote a type profile over the players in  $J$ , and  $T_J$  denote the space of such type profiles. Each of the conditions in [Theorem 3](#) implies that the function

$$v_i(s_i|t_i; t_J) = \sum_{t_K \in T_K} v_i(s_i|t_i; t_{-i})p(t_K|t_i; t_J),$$

satisfies monotonicity, directional masquerade or single crossing differences for every partition  $\{J, K\}$ .

Now consider the candidate equilibrium in which all players send a message that certifies their true type. Suppose that player  $i$  is called after all the players in  $J$ , but before, or at the same time as, the players in  $K$ . Then, on the equilibrium path, she knows  $t_J$  when she is called, and her expected masquerading payoff is given by  $v_i(s_i|t_i; t_J)$ . But then any message of  $i$  that does not fully certify her type can be attributed to a worst case type in the set of its potential senders that takes into account the information acquired by  $i$  in the sequential communication protocol. As before, this is enough to prevent a deviation by player  $i$ .

**Theorem 4** (Sequential Communication). *Suppose that  $(T_i, \succeq)$  is linearly ordered for every  $i$ . For every sequential communication protocol, there exists a full disclosure equilibrium with extremal beliefs that implements  $a^*(\cdot)$  on and off the equilibrium path whenever either of the conditions of [Theorem 3](#) is satisfied.*

*Proof.* See the Appendix. □

**Relaxing Own Type Certifiability.** The assumption that every type  $t_i$  is able to certify the singleton  $\{t_i\}$  may be too strong, and we know from the seller-buyer example of Milgrom

(1981) that full disclosure may remain an equilibrium under weaker evidence structures. In this example, it is sufficient for a seller to be able to certify that her quality is higher than any level not higher than her true quality level. This idea can be generalized as follows. We say that the message correspondence  $M_i(\cdot)$  admits an *evidence base* if there exists a subset  $\mathcal{E}_i \subseteq M_i$  of cardinality  $|\mathcal{E}_i| = |T_i|$  such that for every  $t_i$ , there exists a message  $e_i \in \mathcal{E}_i \cap M_i(t_i)$  such that  $t_i \in wct_i(M_i^{-1}(e_i))$ . This means that we can define a one to one mapping  $\hat{e}_i : T_i \rightarrow \mathcal{E}_i$  such that  $t_i \in wct_i(M_i^{-1}(\hat{e}_i(t_i)))$ .<sup>26</sup>

As an illustration, consider a player  $i$  with three possible types,  $T_i = \{t^1, t^2, t^3\}$ , whose masquerade relation is given by  $t^1 \xrightarrow{\mathcal{M}} t^2 \xrightarrow{\mathcal{M}} t^3$ . The message correspondence  $M_i(t^1) = \{m^1, m^3, m^4\}$ ,  $M_i(t^2) = \{m^1, m^2, m^4\}$  and  $M_i(t^3) = \{m^1, m^2, m^3\}$  admits two evidence bases:  $\mathcal{E}_i = \{m^1, m^2, m^3\}$  and  $\mathcal{E}_i = \{m^4, m^2, m^3\}$ . On the contrary, the message correspondence  $M_i(t^1) = \{m^1, m^4\}$ ,  $M_i(t^2) = \{m^1, m^2, m^3, m^4\}$  and  $M_i(t^3) = \{m^1, m^3\}$  does not admit any evidence base because type  $t^3$  has no message certifying an event for which it is a worst case type.

Clearly, when own type certifiability holds, any collection of messages certifying the singletons  $\{t_i\}$  for  $t_i \in T_i$  forms an evidence base, regardless of which selection  $a^*(\cdot)$  is being considered. In general, however, an evidence base is linked to the masquerade relations and therefore to the selection  $a^*(t)$  that is to be implemented in a full disclosure equilibrium. We show that the existence of an evidence base is necessary for the existence of a full disclosure equilibrium.<sup>27</sup>

**Proposition 2** (Evidence Base: Necessity). *If there exists a full disclosure equilibrium, then there must exist an evidence base  $\mathcal{E}_i$  for every player  $i$ .*

*Proof.* Consider a full disclosure equilibrium  $\sigma$  that implements some Nash equilibria  $a^*(\cdot)$  of the contingent complete information games. Then the strategies  $\sigma_i(\cdot|t_i)$  must have disjoint

---

<sup>26</sup>Notice that the existence of an evidence base is distinct from the “full reports condition” (Lipman and Seppi, 1995) and the “nested range condition” used to get a revelation principle in settings with certifiable information (Green and Laffont, 1986; Forges and Koessler, 2005).

<sup>27</sup>Note that this proposition does not assume extremal beliefs or any restriction on the choice of actions off the equilibrium path.



supports. Let  $\hat{\sigma}_i(t_i)$  be a selection of messages in the support of  $\sigma_i(\cdot|t_i)$ , and suppose that  $t_i \notin \text{wct}_i(M_i^{-1}(\hat{\sigma}_i(t_i)))$ . Then there exists a type  $t'_i \neq t_i$  that wants to masquerade as  $t_i$  and can send the message  $\hat{\sigma}_i(t_i)$ . Since  $\hat{\sigma}_i(t_i)$  is not in the support of  $\sigma_i(\cdot|t'_i)$ , that would contradict the fact that  $\sigma$  is an equilibrium. Therefore, the set  $\{\hat{\sigma}_i(t_i)\}_{t_i \in T_i}$  must form an evidence base for  $M_i(\cdot)$ .  $\square$

The existence of an evidence base turns out to be sufficient as well, provided every certifiable subset of types admits a worst case type. The main intuition of the proof is that when these conditions are satisfied, the function  $\hat{e}_i(\cdot)$  forms an equilibrium strategy of player  $i$ .

**Proposition 3** (Evidence Base: Sufficiency). *There exists a full disclosure equilibrium with extremal beliefs that implements  $a^*(\cdot)$  on and off the equilibrium path whenever the following conditions are satisfied:*

(i) *For every player  $i$ , the correspondence  $M_i(\cdot)$  admits an evidence base.*

(ii) *For every  $m_i \in M_i$ , the set  $M_i^{-1}(m_i)$  admits a worst case type.*

*Proof.* See the Appendix.  $\square$

**Infinite Type Spaces.** It is possible to extend the characterization of [Theorem 1](#) and the subsequent results to infinite type spaces under the conditions that for each  $i$  the function  $v_i(s_i|t_i)$  is lower semi-continuous in  $s_i$  together with a compactness assumption.<sup>28</sup> More precisely, we assume that the certifiable sets of types (the sets  $M_i^{-1}(m_i)$  for  $m_i \in M_i$ ) are all compact subsets of  $T_i$ . As before, we assume that  $t_i$  can always certify the subset  $\{t_i\}$  (own type certifiability). We say that the message correspondence satisfies full compact certifiability if every nonempty compact subset of  $T_i$  is certifiable for every  $i$ . Then we have the following characterization

---

<sup>28</sup>When the set of type profiles is not finite, the notion of strong belief consistency in Kreps and Wilson (1982) is not well defined (it is defined for finite extensive form games); hence, we impose here the same restrictions on extremal beliefs as those we deduced from strong consistency in [Lemma 1](#) for the finite case.

**Theorem 5** (Infinite Type Sets: Characterization). *Suppose that the message correspondence satisfies full compact certifiability and that for every  $i$ , the masquerade payoff function  $v_i(s_i|t_i)$  is lower semicontinuous in  $s_i$ . Then the following statements are equivalent:*

- (i) *There exists a full disclosure equilibrium with extremal beliefs that implements  $a^*(\cdot)$  on and off the equilibrium path.*
- (ii) *For every  $i$  and every nonempty compact subset  $S_i \subseteq T_i$ , there exists a worst-case type for  $S_i$ :  $wct_i(S_i) \neq \emptyset$ .*
- (iii) *For every  $i$  the masquerade relation  $\xrightarrow{M}$  is acyclic on every finite subset of  $T_i$ .*

*If full compact certifiability is not assumed but every certifiable set is compact, (ii) and (iii) are still equivalent and they imply (i), which is equivalent to the existence of a worst case type on every certifiable subset of types.*

*Proof.* The only adaptation needed to the proof of [Theorem 1](#) is the proof that (iii) implies (i). But this is in fact a direct consequence of the theorem of Bergstrom (1975) which says that a binary relation  $P$  on a compact set  $X$  admits a maximal element whenever it is acyclic on every finite subset of  $X$  and the set  $P_X^{-1}(x) = \{y \in X | xPy\}$  is open. Note that a worst case type is a maximal element of the masquerade relation. The condition we need is that the set of types  $s_i$  such that  $t_i \xrightarrow{M} s_i$ , or equivalently  $v_i(s_i|t_i) > v_i(t_i|t_i)$ , is open. By definition, this has to be true if the function  $v_i(s_i|t_i)$  is lower semicontinuous.  $\square$

Naturally, we have the following corollary which extends our result on sufficient conditions to infinite type sets.

**Proposition 4** (Infinite Type Sets: Sufficient Conditions). *Suppose that  $T_i$  is a subset of  $\mathbb{R}$  ordered by the natural order of the real line, and that every certifiable set is compact. Then there exists a full disclosure equilibrium with extremal beliefs that implements  $a^*(\cdot)$  on and off the equilibrium path whenever for every  $i$ , the function  $v_i(s_i|t_i)$  is lower semicontinuous in  $s_i$*

and one of the conditions of *Theorem 2* and either *(MON)*, *(DM)*, *(ID)*, *(SCD)* or *(SP-NRM)* is satisfied.

## 6 Applications

In the following applications, full certifiability is not assumed. It is easier to assume own type certifiability to clarify the exposition, but the existence of an evidence base is sufficient for all the results to hold.

### 6.1 Senders-Receiver Games

Here we consider a game in which one player with no private information, the receiver, wants to implement her ideal action  $a^*(t) \in \mathbb{R}$ . The informed players, the senders, are indexed by  $i$ .  $T_i$  is a (possibly finite) compact subset of  $\mathbb{R}$  endowed with its natural order. Then the following result, which appears in *Giovannoni and Seidmann (2007)* under the assumption of a finite type space, is a direct consequence of our results based on the *(SP-NRM)* condition.<sup>29</sup>

**Proposition 5** (Sender-Receiver – *Giovannoni and Seidmann, 2007*). *Consider a single sender denoted by  $S$ . Then there exists a full disclosure equilibrium with extremal beliefs that implements  $a^*(\cdot)$  on and off the equilibrium path whenever  $a^*(\cdot)$  is strictly monotonic,  $u_S(a, t)$  is single-peaked in  $a$ , and  $v(s|t) = u_S(a^*(s), t)$  is continuous in  $s, t$  and satisfies the no reciprocal masquerade condition.*

We can work with *(ID)* to show the following result about multiple senders-single receiver games. Condition *(iv)* in the proposition below is always trivially satisfied when the type space is finite, and is needed in the case of a continuous type space as explained in [Section 5](#).

---

<sup>29</sup>In an earlier paper, *Seidmann and Winter (1997)* also considered sender-receiver games with a slightly different set of conditions. When the ideal action of the receiver is strictly increasing their conditions imply *(SP-NRM)* and are therefore more restrictive.

**Proposition 6** (Multiple Senders-Single Receiver). *For every sequential communication protocol, there exists a full disclosure equilibrium with extremal beliefs that implements  $a^*(\cdot)$  on and off the equilibrium path whenever the following conditions are satisfied:*

(i)  $a^*(\cdot)$  is nondecreasing.

(ii) For every sender  $i$ , the function  $u_i(a, t_i, t_{-i})$  has increasing differences in  $(a, t_i)$ .

(iii) The beliefs of the senders satisfy type affiliation.

(iv) For every  $i$ ,  $E_{t_{-i}}(u_i(a^*(s_i, t_{-i}), t_i, t_{-i}) \mid t_i)$  is lower semi-continuous in  $s_i$ .

*Proof.* We only need to prove that  $v_i(s_i \mid t_i; t_{-i}) = u_i(a^*(s_i, t_{-i}), t_i, t_{-i})$  has increasing differences in  $(s_i, t_i)$ . To see that, take  $s'_i \succ s_i$  and  $t'_i \succ t_i$  and note that

$$\begin{aligned} v_i(s'_i \mid t'_i; t_{-i}) - v_i(s_i \mid t'_i; t_{-i}) &= u_i(a^*(s'_i, t_{-i}), t'_i, t_{-i}) - u_i(a^*(s_i, t_{-i}), t'_i, t_{-i}) \\ &\geq u_i(a^*(s'_i, t_{-i}), t_i, t_{-i}) - u_i(a^*(s_i, t_{-i}), t_i, t_{-i}) \\ &= v_i(s'_i \mid t_i; t_{-i}) - v_i(s_i \mid t_i; t_{-i}), \end{aligned}$$

where the inequality comes from the fact that  $a^*(s'_i, t_{-i}) \geq a^*(s_i, t_{-i})$  by (i), and from (ii). Then, depending on whether  $T_i$  is finite or a compact subset of  $\mathbb{R}$ , we can conclude by [Theorem 2](#) or [Proposition 4](#).  $\square$

## 6.2 Supermodular Games

Suppose that each  $(T_i, \succeq)$  is a linearly ordered set, and each  $(A_i, \succeq)$  is a complete lattice. We say that the base Bayesian game is *supermodular* if each associated complete information game  $\tilde{\Gamma}(t)$  is a supermodular game in the sense of Milgrom and Roberts (1990) and Vives (1990), and the utilities exhibit complementarities in types and own actions. The following definition recalls these assumptions. These assumptions follow those of Van Zandt and Vives (2007) in their study of Bayesian games of strategic complementarities

**Definition 11.** We say that the (Bayesian) base game  $\Gamma = \langle N, A, p, (u_i)_{i \in N} \rangle$  is **supermodular** if each  $u_i(a, t)$  is supermodular in  $a_i$ , has increasing differences in  $(a_i, a_{-i})$  (strategic complementarities), and has increasing differences in  $(a_i, t)$  (complementarities between own actions and type profiles).

It is well known<sup>30</sup> that in this case  $NE(t)$  is a nonempty sublattice of  $A$ , and that its extremal elements are nondecreasing in  $t$ . Let  $a^*(\cdot)$  be either the highest equilibrium selection or the lowest equilibrium selection. The next proposition also appears in Van Zandt and Vives (2007, Proposition 20).

**Proposition 7** (Supermodular Games 1, Van Zandt and Vives, 2007). *Suppose that  $\Gamma$  is supermodular. Then, for every sequential communication protocol, there exists a full disclosure equilibrium with extremal beliefs that implements  $a^*(\cdot)$  on and off the equilibrium path whenever for every  $i$  either of the following assumptions is satisfied:*

(i)  $u_i(a_i, a_{-i}, t)$  is nondecreasing in  $a_{-i}$ . (*Positive Externalities*)

(ii)  $u_i(a_i, a_{-i}, t)$  is nonincreasing in  $a_{-i}$ . (*Negative Externalities*)

*Proof.* We prove that these assumptions imply (MON). We know that  $a_{-i}^*(s_i, t_{-i})$  is nondecreasing in  $s_i$  and  $t_{-i}$ . First assume positive externalities. Then, for  $s'_i \succeq s_i$ , we have

$$\begin{aligned} u_i(BR_i(a_{-i}^*(s'_i, t_{-i}), t), a_{-i}^*(s'_i, t_{-i}), t) &\geq u_i(BR_i(a_{-i}^*(s_i, t_{-i}), t), a_{-i}^*(s'_i, t_{-i}), t) \\ &\geq u_i(BR_i(a_{-i}^*(s_i, t_{-i}), t), a_{-i}^*(s_i, t_{-i}), t), \end{aligned}$$

where the first inequality comes from the optimality of the best response and the second inequality comes from positive externalities. This proves the monotonicity of ex post masquerading payoffs, and we can conclude with [Theorem 3](#). The proof is similar with negative externalities. □

---

<sup>30</sup>See Milgrom and Roberts (1990) and Vives (1990).

Hence, with the positive or negative externality assumption, full disclosure is an equilibrium regardless of the beliefs of the players. For this result we use the monotonicity condition. If instead we try to use the single crossing differences, we can obtain a new result on supermodular games under type affiliation. In order to do so, however, we need to make additional regularity assumptions. These conditions basically ensure that best-responses satisfy a first order condition, so that the derivatives of the function  $v_i(s_i|t_i; t_{-i})$  can be obtained by the envelope theorem.

In the remainder of this section, we assume that each  $A_i$  is a finite product of closed intervals of the real line with the natural lattice order, and each  $T_i$  is a subset of a real interval  $\Theta_i$ . We assume that the utility functions  $u_i(\cdot)$  are defined on  $A \times \Theta$  where  $\Theta = \Theta_1 \times \dots \times \Theta_n$ , and that they are continuously differentiable. Finally, we assume that every equilibrium action  $a_i^*(t)$ , and every best-response  $BR_i(a_{-i}^*(s_i, t_{-i})|t_i; t_{-i})$  is interior. Altogether, these assumptions ensure that the best-responses  $BR_i(a_{-i}^*(s_i, t_{-i})|t_i; t_{-i})$  always satisfy a first-order condition.

**Proposition 8** (Supermodular Games 2). *Assume that the base game  $\Gamma$  is supermodular, that the utility functions are continuously differentiable on  $A \times \Theta$  and that every best-response  $BR_i(a_{-i}^*(s_i, t_{-i})|t_i; t_{-i})$  is interior. Then, for every sequential communication protocol, there exists a full disclosure equilibrium with extremal beliefs that implements  $a^*(\cdot)$  on and off the equilibrium path whenever the following assumptions are satisfied:*

- (i) *The beliefs satisfy type affiliation.*
- (ii)  *$u_i(a_i, a_{-i}, t)$  has increasing differences in  $(a_{-i}, t_i)$ . (Increasing Differences in Own Type and Others Actions)*

*Proof.* In order to avoid cumbersome notations, we write the proof in the case where each action set  $A_i$  is one-dimensional. The generalization to higher dimensions is straightforward but heavy. With our assumptions, we can define the function  $v_i(s_i|t_i; t_{-i})$  on  $\Theta_i \times \Theta_i \times \Theta_{-i}$ , and it is continuously differentiable. To conclude the proof by [Theorem 2](#), it is sufficient to show

that this function has increasing differences in  $(s_i, t_i)$  on  $\Theta_i \times \Theta_i \times \Theta_{-i}$ . It is well known that this is the case if  $\partial^2 v_i(s_i|t_i, t_{-i})/\partial s_i \partial t_i \geq 0$ .

The assumptions we made ensure that every best-response satisfy the following first order condition

$$\frac{\partial}{\partial a_i} u_i(BR_i(a_{-i}^*(s_i, t_{-i}), t), a_{-i}^*(s_i, t_{-i}), t) = 0. \quad (\text{FOC})$$

Using the chain rule and (FOC) a first time, we have

$$\frac{\partial}{\partial s_i} v_i(s_i|t_i; t_{-i}) = \sum_{j \neq i} \frac{\partial}{\partial a_j} u_i(BR_i(a_{-i}^*(s_i, t_{-i}), t), a_{-i}^*(s_i, t_{-i}), t) \frac{\partial}{\partial s_i} a_j^*(s_i, t_{-i}),$$

and a second time

$$\frac{\partial^2}{\partial s_i \partial t_i} v_i(s_i|t_i; t_{-i}) = \sum_{j \neq i} \frac{\partial^2}{\partial a_j \partial t_i} u_i(BR_i(a_{-i}^*(s_i, t_{-i}), t), a_{-i}^*(s_i, t_{-i}), t) \frac{\partial}{\partial s_i} a_j^*(s_i, t_{-i}).$$

The first term under the summation is nonnegative by (ii), and the second term is also nonnegative since the supermodularity of the base game implies that  $a^*(\cdot)$  is nondecreasing.  $\square$

An immediate corollary of this result is obtained by replacing (ii) by a separability condition between own type and others' actions.

**Corollary 1.** *For every sequential communication protocol, there exists a full disclosure equilibrium with extremal beliefs that implements  $a^*(\cdot)$  on and off the equilibrium path whenever every best-response is interior and the following set of assumptions is satisfied:*

(i) *The beliefs satisfy type affiliation.*

(ii) *For every  $i$ , there exist functions  $g_{ij}(\cdot)$  and  $h_i(\cdot)$  such that*

$$u_i(a_i, a_{-i}, t) = \sum_{j \in N} g_{ij}(a_j, t) + h_i(a_i, a_{-i}, t_{-i}),$$

where  $h_i(\cdot)$  has increasing differences in  $(a_i, a_{-i})$ ,  $g_{ii}(\cdot)$  has increasing differences in  $(a_i, t)$ , and  $g_{ij}(\cdot)$ ,  $i \neq j$ , has increasing differences in  $(a_j, t_i)$ .

The following examples are applications of this result.

**Example 2** (A Coordination Game). Each player has an ideal action  $\theta_i(t) \in \mathbb{R}$ , where  $A_i = \mathbb{R}$ ,  $(T_i, \succeq)$  is a linearly ordered set and  $\theta_i(\cdot)$  is nondecreasing in each  $t_j$ . Players also want to coordinate their own actions with those of other players. Their utilities are given by

$$u_i(a, t) = -\alpha_{ii}(a_i - \theta_i(t))^2 - \sum_{j \neq i} \alpha_{ij}(a_i - a_j)^2,$$

where the  $\alpha_{ij}$  are nonnegative coefficients, normalized so that  $\sum_j \alpha_{ij} = 1$ , and such that  $\alpha_{ii} > 0$ . Particular forms of this class of games have been extensively studied in the economic theory of organizations (see, for example, Alonso et al., 2008, Rantakari, 2008, Myatt and Wallace, 2011, Calvo-Armengol et al., 2011). It is easy to check that the game is supermodular, and that the utility functions are separable in own type and the actions of other players. Therefore we can apply [Corollary 1](#) to get existence of a full disclosure equilibrium when types are affiliated.

**Example 3** (An Influence Game). Galeotti et al. (2011) and Loginova (2012) consider a game in which players try to influence others to play their favorite actions by selectively transmitting information. We consider a more general payoff and information structure with the restriction that players communicate hard information. Each player  $i$  has an unknown ideal action  $\theta_i(t) \in \mathbb{R}$ . Her final payoff is given by  $-\sum_{j=1}^N \alpha_{ij}(a_j - \theta_i(t))^2$ , with  $\alpha_{ij} \geq 0$ , hence she would like all players to play as close as possible to her own ideal action. It is easy to show that under complete information there exists a unique equilibrium in which all players play  $a_i^*(t) = \theta_i(t)$ . If  $\theta_i(t)$  is increasing in  $t$ , this game is supermodular and has increasing differences in types and others' actions. By [Corollary 1](#), there exists a full disclosure equilibrium implementing  $a^*(t)$  under type affiliation. If in addition there is a constant bias, *i.e.*,  $\theta_i(t)$  can be written as  $\theta(t) + b_i$ ,  $b_i \in \mathbb{R}$ , then it is immediate to show that the masquerade relation satisfies (DM), and therefore



by [Theorem 2](#) there exists a full disclosure equilibrium whatever the prior distribution of types (even without type affiliation).

### 6.3 Deliberation with Hard Information

In this section, the base game is a voting game in which a proposal may be adopted to replace the status quo if it is supported by at least  $q$  members of the committee. The set of players is partitioned between the committee,  $\mathcal{C} \subseteq N$ , whose members can cast a vote in the election, and other players who are inactive in the election but may disclose information in the communication phase. Let  $C$  be the size of the committee. Without loss of generality, we can normalize the utility each voter derives from the status quo to 0, and we denote by  $u_i$ , a random variable, the uncertain payoff she derives from the proposal. Each voter  $i$  has a private signal  $t_i$  about the proposal. Consider the function

$$U_i(t) = E(u_i | t_1, \dots, t_n).$$

We assume that it is nondecreasing in  $t$ . This is the case for example if every player believes the vector  $(u_i, t_1, \dots, t_n)$  to be affiliated.

As is standard in voting theory, the complete information voting game has multiple equilibria, but only one in weakly undominated strategies: the sincere voting equilibrium. We can use the tools developed in the rest of the paper to provide conditions under which there exists a full disclosure equilibrium that implements the sincere voting equilibrium. We interpret the pre-play communication game as deliberation with hard evidence. Note that the sincere best response of  $i \in \mathcal{C}$  in the complete information voting game is to vote in favor of the proposal whenever  $U_i(t) > 0$ .

**Example 4** (The Jury Model). The question of voting with private information and deliberation are often studied within the framework of the jury model. This model is a particular case of our framework in which the status quo is to acquit and the proposal is to convict. There is a

state of the world  $\omega \in \{I, G\}$  (innocent or guilty) and the signals of the players are drawn independently according to a distribution  $q(t_i|\omega)$  that satisfies affiliation. The prior on  $\omega$  is given by a probability  $\pi$  that the defendant is guilty. Each voter has a cost  $\gamma_i^C > 0$  for unjustified conviction and  $\gamma_i^A > 0$  for unjustified acquittal. This model is a particular case of our model in which

$$U_i(t) = \gamma_i^A \pi \underbrace{\frac{\prod_{i=1}^n q(t_i | G)}{\prod_{i=1}^n q(t_i | G) + \prod_{i=1}^n q(t_i | I)}}_{\Pr(G|t)} - \gamma_i^C (1 - \pi) \underbrace{\frac{\prod_{i=1}^n q(t_i | I)}{\prod_{i=1}^n q(t_i | G) + \prod_{i=1}^n q(t_i | I)}}_{\Pr(I|t)}.$$

Note that in the jury model, the regions of the type space over which each voters prefers one alternative to the other are naturally nested.

**Example 5** (Altruistic Voters). Suppose that the individual expected payoff of player  $i$  from the alternative is given by a nondecreasing function  $w_i(t_i)$  that only depends on her type, but that she is altruistic either out of generosity, or because she internalizes the danger of a revolution if others are too unhappy. She then evaluates the expected value of the alternative according to the function

$$U_i(t) = (1 - \varepsilon_i)w_i(t_i) + \varepsilon_i E \left( \sum_{j \neq i} w_j(t_j) \mid t_i \right),$$

where  $\varepsilon_i$  is a (typically small) number between 0 and 1. This example satisfies our assumptions, and in contrast with the jury model the regions of the type space over which each voter prefers the alternative to the status quo are typically not nested.

**Unanimity Rule.** With the unanimity rule, we have for members of the committee:

$$v_i(s_i|t_i; t_{-i}) = U_i(t) \mathbb{1}_{U_i(t) > 0} \mathbb{1}_{S(s_i, t_{-i}) \geq C-1},$$

and for players who are not members of the committee

$$v_i(s_i|t_i; t_{-i}) = U_i(t) \mathbb{1}_{S(s_i, t_{-i}) \geq C},$$

where

$$S_i(s_i, t_{-i}) = \sum_{j \in \mathcal{C} \setminus \{i\}} \mathbb{1}_{U_j(s_i, t_{-i}) > 0}.$$

In both cases, the monotonicity condition (MON) is satisfied because  $S_i(s_i, t_{-i})$  is nondecreasing in  $s_i$ . This leads to the following result.

**Proposition 9** (Deliberation: Unanimity). *Under the unanimity rule, for every sequential communication protocol, there exists a full disclosure equilibrium with extremal beliefs that implements the sincere voting equilibrium.*

**Non-Unanimous Rules.** Now suppose that the rule is  $q < C$ . For players who are not members of the committee

$$v_i(s_i|t_i; t_{-i}) = U_i(t) \mathbb{1}_{S_i(s_i, t_{-i}) \geq q}.$$

For players who belong to the committee we have

$$v_i(s_i|t_i; t_{-i}) = U_i(t) \mathbb{1}_{U_i(t) > 0} \mathbb{1}_{S_i(s_i, t_{-i}) \geq q-1} + U_i(t) \mathbb{1}_{U_i(t) < 0} \mathbb{1}_{S_i(s_i, t_{-i}) \geq q}.$$

These function do not satisfy the monotonicity property but we can show that they have increasing differences in  $(s_i, t_i)$ .

**Lemma 2.** *With non-unanimous rules, for every  $i$ , the function  $v_i(s_i|t_i; t_{-i})$  has increasing differences in  $(s_i, t_i)$ .*

*Proof.* For  $t'_i \succeq t_i$ , the difference

$$v_i(s_i|t'_i; t_{-i}) - v_i(s_i|t_i; t_{-i}) = \underbrace{(U_i(t'_i, t_{-i})\mathbb{1}_{U_i(t'_i, t_{-i})>0} - U_i(t_i, t_{-i})\mathbb{1}_{U_i(t_i, t_{-i})>0})}_{\geq 0} \mathbb{1}_{S_i(s_i, t_{-i}) \geq q-1} \\ + \underbrace{(U_i(t'_i, t_{-i})\mathbb{1}_{U_i(t'_i, t_{-i})<0} - U_i(t_i, t_{-i})\mathbb{1}_{U_i(t_i, t_{-i})<0})}_{\geq 0} \mathbb{1}_{S_i(s_i, t_{-i}) \geq q}$$

is nondecreasing in  $s_i$  since  $S_i(s_i, t_{-i})$  is nondecreasing in  $s_i$ .  $\square$

Hence by [Theorem 2](#) we know that if the beliefs satisfy type affiliation, there exists a full disclosure equilibrium with extremal beliefs that implements the sincere voting equilibrium. This directly leads to the following result.

**Proposition 10** (Deliberation with Non-Unanimous Rules). *If beliefs satisfy type affiliation, then for every sequential communication protocol there exists a full disclosure equilibrium with extremal beliefs that implements the sincere voting equilibrium.*

This result applies in particular to our examples. [Proposition 9](#) and [Proposition 10](#) together show that under relatively mild assumptions, deliberation with hard information leads to full disclosure regardless of the voting rule. While other results in the voting literature suggest that unanimity may perform less well than other voting rules in terms of information revelation,<sup>31</sup> our results imply that with hard information, unanimity can lead to full disclosure regardless of the information structure, whereas more structure may be needed for other voting rules to induce full disclosure. Schulte (2010) shows this result for the specific case of the jury theorem, and Mathis (2011) extends it to nested preferences. By contrast, our result shows that full disclosure holds for all preferences that react to information in the same direction, as long as affiliation is satisfied.

---

<sup>31</sup>See, e.g., Austen-Smith and Feddersen (2006).

# Appendix

## A Proofs

*Proof of Lemma 1 (Consistent Extremal Beliefs).* Consider a sequence of strictly positive communication strategy profiles  $\sigma_i^k : T_i \rightarrow \Delta(M_i)$ ,  $i = 1, \dots, n$ ,  $k = 1, 2, \dots, \infty$ , where  $\sigma_i^k(t_i)$  is completely mixed over  $M_i(t_i)$ , converging to some full disclosure communication strategy profile  $\sigma_i : T_i \rightarrow \Delta(M_i)$ ,  $i = 1, \dots, n$ , where  $M_i^{-1}(\sigma_i(t_i)) = \{t_i\}$  for every  $i \in N$  and  $t_i \in T_i$ . Denote by  $\mu^k(t_j | m_j, t_{-j})$  the belief, computed by Bayes rule from  $(\sigma_i^k)_{i \in N}$  and the priors, of players other than  $j$  following a unilateral deviation by player  $j$  from full disclosure. Let  $\mu(t_j | m_j, t_{-j}) = \lim_{k \rightarrow \infty} \mu^k(t_j | m_j, t_{-j})$ . We have to show that  $\mu(t_j | m_j, t_{-j}) = 1$  for some  $t_{-j} \in T_{-j}$  implies  $\mu(t_j | m_j, s_{-j}) = 1$  for every  $s_{-j} \in T_{-j}$ . We have:

$$\mu^k(t_j | m_j, t_{-j}) = \frac{\sigma_j^k(m_j | t_j)p(t_j, t_{-j})}{\sum_{t'_j \in T_j} \sigma_j^k(m_j | t'_j)p(t'_j, t_{-j})}. \quad (1)$$

Hence, since  $\lim_{k \rightarrow \infty} \mu^k(t_j | m_j, t_{-j}) = 1$  and the priors have full support, we get

$$\lim_{k \rightarrow \infty} \frac{\sigma_j^k(m_j | t'_j)}{\sigma_j^k(m_j | t_j)} = 0, \quad \text{for every } t'_j \neq t_j,$$

which implies  $\lim_{k \rightarrow \infty} \mu^k(t_j | m_j, s_{-j}) = \lim_{k \rightarrow \infty} \frac{\sigma_j^k(m_j | t_j)p(t_j, s_{-j})}{\sum_{t'_j \in T_j} \sigma_j^k(m_j | t'_j)p(t'_j, s_{-j})} = 1$  for every  $s_{-j} \in T_{-j}$ .

This proves the first part of the lemma. For the second part, consider a full disclosure strategy  $\sigma$  such that  $M_i^{-1}(\sigma_i(t_i)) = \{t_i\}$ . Suppose that this strategy is supported by extremal beliefs out of the equilibrium path such that  $\mu(t_j | m_j, t_{-j})$  puts probability one on a type  $t_j^*(m_j) \in M_j^{-1}(m_j)$  regardless of  $t_{-j}$ . Let  $N(t_j)$  be the number of messages  $m_j \in M_j(t_j)$  such that  $t_j = t_j^*(m_j)$ ,  $m_j \neq \sigma_j(t_j)$ .

Then consider the sequence of completely mixed communication strategy profiles  $\sigma^k$  such that  $\sigma_i^k(\cdot | t_i)$  puts probability  $1 - \frac{N(t_i)}{k} - \frac{|M_i(t_i)| - N(t_i) - 1}{k^2}$  on the message  $\sigma_i(t_i)$ , probability  $1/k$  on every message  $m_i \in M_i(t_i)$  such that  $t_i^*(m_i) = t_i$ ,  $m_i \neq \sigma_i(t_i)$ , and probability  $1/k^2$  on every

remaining message. It is easy to see that  $\sigma^k$  converges to  $\sigma$ .

Now consider the belief system  $\mu^k$  associated to the completely mixed strategy profile  $\sigma^k$ . We are only concerned about the limit of these beliefs for a player  $i \neq j$  at the information set where she knows  $t_{-j}$  and has received a message  $m_j$  from player  $j$  such that  $M_j^{-1}(m_j) \neq \{t_j\}$ , which is given by (1). Since  $\sigma_j^k(m_j|t_j^*(m_j))$  is proportional to  $1/k$ , and for any  $t_j \neq t_j^*(m_j)$ ,  $\sigma_j^k(m_j|t_j)$  is proportional to  $1/k^2$ , we have that for every  $s_{-j}$ ,  $\lim_{k \rightarrow \infty} \mu^k(t_j^*(m_j) | m_j, s_{-j}) = 1$ . This proves that  $\mu^k$  converges to  $\mu$  at every relevant information set. Since what  $\mu$  is at other information sets is irrelevant, we can fix it to be equal to the limit of  $\mu^k$  (which exists), and this concludes the proof.  $\square$

*Proof of Proposition 1 (Worst-Case Types Identification).*

(a) First, consider the case of (SCD). Let  $\underline{s}_i \in \arg \min_{s_i \in S_i} v_i(s_i|\underline{t}_i)$  and  $t_i \in S_i \setminus \{\underline{s}_i\}$ . If  $t_i \succ \underline{s}_i$ , then by (SCD)

$$v_i(t_i|\underline{t}_i) \geq v_i(\underline{s}_i|\underline{t}_i) \Rightarrow v_i(t_i|t_i) \geq v_i(\underline{s}_i|t_i),$$

hence  $t_i$  does not want to masquerade as  $\underline{s}_i$ . If  $t_i \prec \underline{s}_i$ , suppose  $t_i \xrightarrow{\mathcal{M}} \underline{s}_i$ . Then, by (SCD)

$$v_i(t_i|t_i) < v_i(\underline{s}_i|t_i) \Rightarrow v_i(t_i|\underline{t}_i) < v_i(\underline{s}_i|\underline{t}_i).$$

But this contradicts the definition of  $\underline{s}_i$ . Since this holds for any  $t_i \in S_i \setminus \{\underline{s}_i\}$ , and no type ever wants to masquerade as herself, we have  $\underline{s}_i \in wct_i(S_i)$ .

(b) Next, consider the case of (SP-NRM). By construction,  $\{\underline{s}_i^k\}$  is a nondecreasing sequence and  $\{\overline{s}_i^k\}$  is a nonincreasing sequence, and since  $S_i$  is bounded (it is finite) the two sequences converge. By (SP), since  $\underline{s}_i^k \xrightarrow{\mathcal{M}} \underline{s}_i^{k-1}$ , we have  $\underline{s}_i^k \xrightarrow{\mathcal{M}} s_i$  for every  $s_i \in S_i$  such that  $\underline{s}_i^{k-1} \preceq s_i \prec \underline{s}_i^k$ . Similarly  $\overline{s}_i^k \xrightarrow{\mathcal{M}} s_i$  for every  $s_i \in S_i$  such that  $\overline{s}_i^k \prec s_i \preceq \overline{s}_i^{k-1}$ . For the rest of the proof we call this property the betweenness property.

We prove that  $\underline{s}_i \in wct_i(S_i)$ . The proof for  $\overline{s}_i$  is symmetric and therefore omitted. By definition, if  $s_i \in S_i$  satisfies  $s_i \succ \underline{s}_i$  then  $s_i$  does not want to masquerade as  $\underline{s}_i$ . If  $\underline{s}_i = \underline{s}_i^0$ , the

proof is over. Suppose that  $\underline{s}_i \succ \underline{s}_i^0$  and that there exists some  $s_i \in S_i$  such that  $s_i \prec \underline{s}_i$  that wants to masquerade as  $\underline{s}_i$ . Then there exists  $k$  such that  $\underline{s}_i^k \preceq s_i \prec \underline{s}_i^{k+1} \preceq \underline{s}_i$ . But then the betweenness property implies that  $s_i \xrightarrow{\mathcal{M}} \underline{s}_i^{k+1} \xrightarrow{\mathcal{M}} s_i$  which contradicts (NRM).  $\square$

*Proof of Theorem 3 (Ex Post Sufficient Conditions).*

(i) For every  $s'_i \succeq s_i$ ,  $v_i(s'_i|t_i; t_{-i}) \geq v_i(s_i|t_i; t_{-i})$  and the inequality is preserved by taking expectations, hence  $v_i(s_i|t_i)$  satisfies (MON).

(ii) Suppose  $s_i \prec t_i$ . Then by ex post directional masquerade,  $v_i(s_i|t_i; t_{-i}) \leq v_i(t_i|t_i; t_{-i})$ , and taking expectations  $v_i(s_i|t_i) \leq v_i(t_i|t_i)$ . Therefore if  $v_i(s_i|t_i) > v_i(t_i|t_i)$  it must be the case that  $s_i \succ t_i$ , which means that (DM) is satisfied.

(iii) By affiliation, the function  $p(t_{-i}|t_i)$  is log-supermodular and strictly positive, therefore  $p(t_{-i}|t_i)\Delta(t_i; t_{-i})$  is single crossing in  $t_i$ , and for any two  $t_{-i}$  and  $t'_{-i}$ , the pair of functions  $p(t_{-i}|t_i)\Delta(t_i; t_{-i})$  and  $p(t'_{-i}|t_i)\Delta(t_i; t'_{-i})$  satisfies signed-ratio monotonicity. Then, we use Quah and Strulovici (2012, Theorem 1) to show that the aggregation of these single crossing functions  $\Delta(t_i) = \sum_{t_{-i} \in T_{-i}} p(t_{-i}|t_i)\Delta(t_i; t_{-i})$  is single crossing in  $t_i$ . This proves that  $v_i(s_i|t_i)$  satisfies (SCD) leading to full disclosure by Theorem 2.

(iv) Take  $s'_i \succ s_i$ . Since  $v_i(s_i|t_i; t_{-i})$  has (ID) in  $(s_i, t_i)$ ,  $\Delta(t_i; t_{-i}) = v_i(s'_i|t_i; t_{-i}) - v_i(s_i|t_i; t_{-i})$  is nondecreasing and therefore single crossing in  $t_i$ . Furthermore, for any  $t_{-i}$  and  $t'_{-i}$  in  $T_{-i}$ , the pair of functions  $\Delta(t_i; t_{-i})$  and  $\Delta(t_i; t'_{-i})$  of  $t_i$  satisfies the signed-ratio monotonicity property of Quah and Strulovici (2012). Indeed, suppose that  $\Delta(t_i; t_{-i}) < 0$  and  $\Delta(t_i; t'_{-i}) > 0$  and take  $\tilde{t}_i \succ t_i$ . First, we have  $\Delta(\tilde{t}_i; t'_{-i}) \geq \Delta(t_i; t'_{-i}) > 0$ . Next, whether  $\Delta(\tilde{t}_i; t_{-i})$  is such that  $\Delta(t_i; t_{-i}) \leq \Delta(\tilde{t}_i; t_{-i}) < 0$  or such that  $\Delta(t_i; t_{-i}) < 0 \leq \Delta(\tilde{t}_i; t_{-i})$ , we have  $-\Delta(\tilde{t}_i; t_{-i}) \leq -\Delta(t_i; t_{-i})$ . Hence

$$-\frac{\Delta(t_i; t_{-i})}{\Delta(t_i; t'_{-i})} \geq -\frac{\Delta(\tilde{t}_i; t_{-i})}{\Delta(\tilde{t}_i; t'_{-i})},$$

which proves signed-ratio monotonicity. From (iii), we conclude that  $v_i(s_i|t_i)$  satisfies (SCD).

$\square$

*Proof of Theorem 4 (Sequential Communication).* The starting point of the proof is to notice that: whenever (i) is satisfied, the functions  $v_i(s_i|t_i; t_J)$  satisfy monotonicity in  $s_i$  for every  $J \subseteq N \setminus \{i\}$ ; whenever (ii) is satisfied, they satisfy directional masquerade; and when (iii) is satisfied they have single crossing differences in  $(s_i, t_i)$ . This implies that in each case we can define a masquerade relation on  $T_i$  for every  $J$  and every realization of  $t_J$ , and that this masquerade relation is acyclic. Therefore, we can associate a worst case type knowing  $t_J$  to every subset  $S_i \subseteq T_i$ . Let  $wct_i(S_i|t_J)$  denote the set of these worst case types.

Pick any sequential protocol. Our candidate equilibrium is one in which the communication strategy of player  $i$  when she is called is to exactly certify her type by sending a message  $m_i^*(t_i)$  such that  $M_i^{-1}(m_i^*(t_i)) = \{t_i\}$ , regardless of the message history and of the period in which she is called. At the action stage, the strategy is to play according to  $a_i^*(t)$  if all players have exactly certified their type. If player  $i$  believes that the other players are of type  $t_{-i}$ , and she knows that the other players believe she is of type  $s_i$ , she best responds to  $a_{-i}^*(s_i, t_{-i})$ . We do not need to specify what she does elsewhere, only that she best responds to her beliefs and the strategies of other players.

At the end of the communication phase and at the moment they are called, the players know the players who have been called before them. For each player  $i$ , let  $P(i) \subseteq N \setminus \{i\}$  denote the set of predecessors of  $i$  in the realized calling order. At the time she is called, player  $i$  forms the belief  $t_j^*(m_j)$  about each player  $j$  that preceded her and sent the message  $m_j$ , where  $t_j^*(m_j) \in wct_j(M_j^{-1}(m_j)|t_{P(j)}^*(m_{P(j)}))$  (the notation  $t_{P(j)}^*(m_{P(j)})$  stands for the vector  $(t_\ell^*(m_\ell))_{\ell \in P(j)}$ ). At the end of the communication phase, when she is called to choose her action, each player  $i$  keeps the same belief about all the players in  $P(i)$ , and forms the belief  $t_j^*(m_j)$  about every player  $j$  that communicated after her and sent a message  $m_j$ , where again  $t_j^*(m_j) \in wct_j(M_j^{-1}(m_j)|t_{P(j)}^*(m_{P(j)}))$ . The beliefs  $t_j^*(\cdot)$  are well defined (recursively), and at the end of the communication stage, all the players other than  $j$  hold the same belief about player  $j$ . Note that these beliefs are consistent on the equilibrium path, and satisfy the hard information constraint.



To show that this is an equilibrium, we have to show that it is sequentially rational for each player  $i$  to exactly certify her type when she is called regardless of the history of messages  $m_{P(i)}$ . When player  $i$  is called, she believes that all the players  $j \in P(i)$  are of type  $t_j^*(m_j)$ . And she expects all the players that will be called after her to exactly reveal their type. Suppose  $i$  is of type  $t_i$ . If she sends the equilibrium message  $m_i^*(t_i)$ , she expects to eventually receive the payoff  $v_i(t_i|t_i; t_{P(i)}^*(m_{P(i)}))$ . If on the other hand she sends another message  $m_i$ , this message will be interpreted as  $t_i^*(m_i)$  which is a worst case type and hence delivers an expected payoff

$$v_i(t_i^*(m_i)|t_i; t_{P(i)}^*(m_{P(i)})) \leq v_i(t_i|t_i; t_{P(i)}^*(m_{P(i)}))$$

and player  $i$  has therefore no incentive to deviate.

The only remaining point that must be checked to conclude the proof is that the beliefs we defined to support the equilibrium satisfy strong consistency. It is easy to understand that the same approach we used in [Lemma 1](#) and in [Proposition 3](#) will work, but we refrain from writing this burdensome proof to save on space.  $\square$

*Proof of Proposition 3 (Evidence Base: Sufficiency).* Suppose that for every  $i$ , there exists an evidence base  $\mathcal{E}_i$  for  $M_i(\cdot)$  and let  $\hat{e}_i : T_i \rightarrow \mathcal{E}_i$  be the associated one to one mapping such that  $t_i \in wct_i(M_i^{-1}(\hat{e}_i(t_i)))$ . Then we contend that if (ii) holds, there exists a full disclosure equilibrium in which the communication strategy of player  $i$  is given by the mapping  $\hat{e}_i(\cdot)$ . To show that, we now construct extremal beliefs that support this equilibrium. Consider a unilateral deviation of player  $i$  of type  $t_i$  who plays a message  $m_i$  instead of  $\hat{e}_i(t_i)$ . If  $m_i \notin \mathcal{E}_i$ , then the deviation is detected, and can be prevented by the belief that the type of player  $i$  is some  $s_i \in wct_i(M_i^{-1}(m_i))$ . Now suppose that  $m_i \in \mathcal{E}_i$ . Then the deviation cannot be detected by the other players. But then it must be the case that  $m_i = \hat{e}_i(s_i)$  for some  $s_i \neq t_i$ . And the belief associated to  $m_i$  is therefore the “on the equilibrium path” belief that  $i$  is of type  $s_i$ . Then by construction of  $\hat{e}_i(\cdot)$ , we have  $s_i \in wct_i(M_i^{-1}(m_i))$ , which means that such a deviation cannot be beneficial for  $t_i$ .

To finish the proof, we only need to show that the equilibrium we have constructed satisfies strong consistency of beliefs. The proof is similar to the proof of the second point in [Lemma 1](#) but slightly more involved because own type certifiability is not satisfied. Remember that the equilibrium strategy is given by the profile  $\hat{e}$ . Let  $t_i^*(m_i) \in M_i^{-1}(m_i)$  be the equilibrium belief associated to any message  $m_i \notin \mathcal{E}_i$  (hence  $t_i^*(m_i) \in wct_i(M_i^{-1}(m_i))$ ). Let  $N(t_i)$  be the number of messages  $m_i \in M_i(t_i) \setminus \mathcal{E}_i$  such that  $t_i = t_i^*(m_i)$ .

Let  $\sigma^k$  be a sequence of completely mixed communication strategy profiles such that  $\sigma_i^k(\cdot|t_i)$  puts probability  $1 - \frac{N(t_i)}{k} - \frac{|M_i(t_i)| - N(t_i) - 1}{k^2}$  on the message  $\hat{e}_i(t_i)$ , probability  $1/k$  on every message  $m_i \in M_i(t_i)$ ,  $m_i \notin \mathcal{E}_i$ , such that  $t_i^*(m_i) = t_i$ , and probability  $1/k^2$  on every remaining message. It is then easy to see that  $\sigma^k$  converges to  $\hat{e}$ .

Now consider the belief  $\mu_i^k$  associated to the completely mixed strategy profile  $\sigma^k$  for each player  $i$ . To check consistency, we need to check that the beliefs  $\mu_i^k$  converge to the equilibrium beliefs at two kinds of information set.

First consider an information set on the equilibrium path. That is, all the players have observed a message profile  $m$  such that  $m_i \in \mathcal{E}_i$  for every  $i$ . Then

$$\mu_i^k(t_{-i}|m, t_i) = \frac{\sigma_{-i}^k(m_{-i}|t_{-i})p(t_{-i}|t_i)}{\sum_{s_{-i} \in T_{-i}} \sigma_{-i}^k(m_{-i}|s_{-i})p(s_{-i}|t_i)}, \quad (2)$$

where

$$\sigma_{-i}^k(m_{-i}|t_{-i}) = \prod_{j \neq i} \sigma_j^k(m_j|t_j),$$

converges to 1 if  $m_j = \hat{e}_j(t_j)$  for every  $j \neq i$  and to 0 otherwise. Hence in the limit,  $\mu_i^k(t_{-i}|m, t_i)$  puts probability 1 on the vector  $\hat{e}^{-1}(m_{-i})$  which is indeed the belief that  $i$  forms about the other players on the equilibrium path.

Next consider an information set in which a unilateral deviation is detected. That is all the players but  $j$  have sent a message profile  $m_{-j} \in \mathcal{E}_{-j}$ , whereas  $j$  has sent a message  $m_j \notin \mathcal{E}_j$ . Then the belief formed by  $j$  about other players can be analyzed as we just did and satisfies strong consistency. We need to show that this is true for other players as well so consider a

player  $i \neq j$ . Her belief about other players is still given by (2). But now we have the following:

$$\sigma_{-i}^k(m_{-i}|t_{-i}) = \begin{cases} O(1/k) & \text{if } m_\ell \in \hat{e}(t_\ell) \text{ for every } \ell \notin \{i, j\} \text{ and } t_j^*(t_j) = m_j \\ O(1/k^2) & \text{if } m_\ell \in \hat{e}(t_\ell) \text{ for every } \ell \notin \{i, j\} \text{ and } t_j^*(t_j) \neq m_j \\ o(1/k^{n-1}) & \text{otherwise.} \end{cases}$$

Therefore,  $\mu_i^k(t_{-i}|m, t_i)$  must converge to a belief that puts probability 1 on the unique profile  $t_{-i}$  that satisfies  $t_\ell = \hat{e}_\ell^{-1}(m_\ell)$  for  $\ell \notin \{i, j\}$ , and  $t_j = t_j^*(m_j)$ . This is exactly the belief we used to construct our equilibrium, and this concludes the proof.  $\square$

## References

- ALGER, I. AND R. RENAULT (2006): “Screening Ethics when Honest Agents Care about Fairness,” *International Economic Review*, 47, 59–85.
- ALONSO, R., W. DESSEIN, AND N. MATOUSCHEK (2008): “When Does Coordination Require Centralization?” *American Economic Review*, 98, 145–179.
- ATHEY, S. (2002): “Monotone comparative statics under uncertainty,” *The Quarterly Journal of Economics*, 117, 187–223.
- AUSTEN-SMITH, D. AND T. FEDDERSEN (2006): “Deliberation, Preference Uncertainty and Voting Rules,” *American Political Science Review*, 100, 209–218.
- BATTIGALLI, P. (1996): “Strategic independence and perfect Bayesian equilibria,” *Journal of Economic Theory*, 70, 201–234.
- BEN-PORATH, E. AND B. LIPMAN (2012): “Implementation and Partial Provability,” *Journal of Economic Theory*, forthcoming.
- BERGSTROM, T. (1975): “Maximal Elements of Acyclic Relations on Compact Sets,” *Journal of Economic Theory*, 10, 403–404.
- BULL, J. AND J. WATSON (2004): “Evidence Disclosure and Verifiability,” *Journal of Economic Theory*, 118, 1–31.
- (2007): “Hard Evidence and Mechanism Design,” *Games and Economic Behavior*, 58, 75–93.
- CALVO-ARMENGOL, A., J. DE MARTÍ, AND A. PRAT (2011): “Communication and Influence,” *mimeo*.

- CRAWFORD, V. P. AND J. SOBEL (1982): “Strategic Information Transmission,” *Econometrica*, 50, 1431–1451.
- DENECKERE, R. AND S. SEVERINOV (2008): “Mechanism design with partial state verifiability,” *Games and Economic Behavior*, 64, 487–513.
- FORGES, F. AND F. KOESSLER (2005): “Communication Equilibria with Partially Verifiable Types,” *Journal of Mathematical Economics*, 41, 793–811.
- FUDENBERG, D. AND J. TIROLE (1991): “Perfect Bayesian Equilibrium and Sequential Equilibrium,” *Journal of Economic Theory*, 53, 236–260.
- GALEOTTI, A., C. GHIGLINO, AND F. SQUINTANI (2011): “Strategic Information Transmission in Networks,” mimeo.
- GERARDI, D. AND L. YARIV (2007): “Deliberative Voting,” *Journal of Economic Theory*, 134, 317–338.
- GIOVANNONI, F. AND D. J. SEIDMANN (2007): “Secrecy, Two-Sided Bias and the Value of Evidence,” *Games and Economic Behavior*, 59, 296–315.
- GREEN, J. R. AND J.-J. LAFFONT (1986): “Partially Verifiable Information and Mechanism Design,” *Review of Economic Studies*, 53, 447–456.
- GROSSMAN, S. J. (1981): “The Informational Role of Warranties and Private Disclosure about Product Quality,” *Journal of Law and Economics*, 24, 461–483.
- HAGENBACH, J. AND F. KOESSLER (2012): “Full Disclosure in Decentralized Organizations,” mimeo.
- JACKSON, M. AND X. TAN (2012): “Disclosure of Information, Deliberation and Voting,” *Journal of Economic Theory*, forthcoming.
- KARTIK, N. AND O. TERCIEUX (2012): “Implementation with Evidence,” *Theoretical Economics*, forthcoming.
- KOESSLER, F. AND R. RENAULT (2012): “When Does a Firm Disclose Product Information?” *Rand Journal of Economics*, forthcoming.
- KREPS, D. M. AND R. WILSON (1982): “Sequential Equilibria,” *Econometrica*, 50, 863–894.
- LIPMAN, B. L. AND D. SEPPI (1995): “Robust Inference in Communication Games with Partial Provability,” *Journal of Economic Theory*, 66, 370–405.
- LIZZERI, A. AND L. YARIV (2011): “Sequential Deliberation,” mimeo.
- LOGINOVA, U. (2012): “Strategic communication in networks: The choice between soft and hard information transmission,” mimeo.

- MATHIS, J. (2011): “Deliberation with Evidence,” *American Political Science Review*, 105, 516–529.
- MILGROM, P. (1981): “Good News and Bad News: Representation Theorems and Applications,” *Bell Journal of Economics*, 12, 380–391.
- (2004): *Putting auction theory to work*, Cambridge Univ Pr.
- MILGROM, P. AND J. ROBERTS (1986): “Relying on the Information of Interested Parties,” *Rand Journal of Economics*, 17, 18–32.
- (1990): “Rationalizability, learning, and equilibrium in games with strategic complementarities,” *Econometrica*, 1255–1277.
- MILGROM, P. AND J. WEBER (1982): “A Theory of Auctions and Competitive Bidding,” *Econometrica*, 50, 1089–1122.
- MYATT, D. AND C. WALLACE (2011): “Endogenous Information Acquisition in Coordination Games,” *Review of Economic Studies*, forthcoming.
- MYERSON, R. B. (1994): “Communication, Correlated Equilibria and Incentive Compatibility,” in *Handbook of Game Theory*, ed. by R. J. Aumann and S. Hart, Elsevier Science B. V., vol. 2, chap. 24, 827–847.
- OKUNO-FUJIWARA, A., M. POSTLEWAITE, AND K. SUZUMURA (1990): “Strategic Information Revelation,” *Review of Economic Studies*, 57, 25–47.
- QUAH, J. AND B. STRULOVICI (2012): “Aggregating the Single Crossing Property,” *Econometrica*, forthcoming.
- RANTAKARI, H. (2008): “Governing Adaptation,” *The Review of Economic Studies*, 75, 1257.
- SCHULTE, E. (2010): “Information Aggregation and Preference Heterogeneity in Committees,” *Theory and Decision*, 69, 97–118.
- SEIDMANN, D. J. AND E. WINTER (1997): “Strategic Information Transmission with Verifiable Messages,” *Econometrica*, 65, 163–169.
- VAN ZANDT, T. AND X. VIVES (2007): “Monotone Equilibria in Bayesian Games of Strategic Complementarities,” *Journal of Economic Theory*, 134, 339–360.
- VIVES, X. (1990): “Nash equilibrium with strategic complementarities,” *Journal of Mathematical Economics*, 19, 305–321.