Certifiable Pre-Play Communication: Full Disclosure*

Jeanne Hagenbach[†] Frédéric Koessler[‡] Eduardo Perez-Richet[§]

November 7, 2013

Abstract

This article asks when communication with certifiable information leads to complete information revelation. We consider Bayesian games augmented by a pre-play communication phase in which announcements are made publicly. We first characterize the augmented games in which there exists a fully revealing sequential equilibrium with extremal beliefs (*i.e.*, any deviation is attributed to a single type of the deviator). Next, we define a class of games for which existence of a fully revealing equilibrium is equivalent to a richness property of the evidence structure. 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 games with strategic complementarities, voting with deliberation, and persuasion games with multidimensional types.

*We thank Alfred Galichon, Johannes Horner, Navin Kartik, Yukio Koriyama, Rida Laraki, Ludovic Renou, Francisco Ruiz-Aliseda, Larry Samuelson, Karl Schlag, Joel Sobel, Bruno Strulovici and three anonymous referees for useful comments and suggestions. We also thank seminar participants at École Polytechnique, University of Mannheim, PSE, EUI Florence, University of Warwick, Yale University, NYU, UPF Barcelona, IESE Business School, University of Bern, University of Grenoble, University of Lille, Columbia University, Chicago Booth, University of Bonn, Sciences Po Paris, Humboldt University of Berlin, the International Conference of Game Theory at Stony Brook, and the conference on Mathematical Aspects of Game Theory and Applications in Toulouse.

[†]École Polytechnique – CNRS, e-mail: `jeanne.hagenbach@polytechnique.edu`

[‡]Paris School of Economics – CNRS, e-mail: `koessler@pse.ens.fr`

[§]École Polytechnique, e-mail: `eduardo.perez@polytechnique.edu`

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

JEL classification: C72; D82.

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 can make certifiable statements (*i.e.*, the availability of messages depends on types).[1] We consider general Bayesian games augmented by a communication stage at which players can publicly disclose information about their type before choosing their actions in a second stage.

In order to enforce full disclosure, players must be able to coordinate on second stage actions that deter any unilateral attempt to conceal or misrepresent 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 reveals her type, the game played at the action stage is a complete information game that depends on the type profile. Hence in a fully revealing equilibrium, each player expects to get the payoff associated with the equilibrium[2] of the complete information game that unfolds. If a player could convince all the others that her 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. The masquerade relation is best represented as a directed graph on the type set of each player, such that an arrow points from one type to another whenever the former wants

[1]The assumption of certifiable information has been introduced in sender-receiver games by Grossman (1981) and Milgrom (1981). 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, Deneckere and Severinov, 2008, Sher and Vohra, 2011, Ben-Porath and Lipman, 2012, Kartik and Tercieux, 2012).

[2]Uniqueness is assumed only in the introduction in order to simplify the exposition.

to masquerade as the latter.

This summary of players' incentives suggests a natural way to deter obfuscation at the communication stage. To support a fully revealing equilibrium, we must ensure that, for any player and any possible message from this player, other players can attribute this message to a *worst case type*, *i.e.*, a type that none of the other types who could have possibly sent this message wants to masquerade as. This idea of assigning a worst case type to any message captures the idea of Milgrom (1981) that in order to enforce full information revelation the players should exercise skepticism. Our first main result characterizes the necessary and sufficient conditions for the existence of a fully revealing sequential equilibrium (Kreps and Wilson, 1982) when we restrict players to hold *extremal beliefs* off the equilibrium path, *i.e.*, beliefs that put probability one on a single type of a deviating player.[3] We say that the communication game admits an *evidence base* if every type of a player has access to a distinct message that certifies a set of types for which it is a worst case type.[4] We show that there exists a fully revealing sequential equilibrium with extremal beliefs if and only if the communication game admits an evidence base and every certifiable subset of types admits a worst case type.

Most of our results rely on this general characterization and on the following simple observation: the existence of a worst case type for every subset of types of a player is equivalent to the acyclicity of her masquerade relation which, in turn, is equivalent to the existence of a function that weakly represents the masquerade relation of the player.[5] For the class of games satisfying this property, there exists a fully revealing equilibrium (with any beliefs off path) if and only if there exists an evidence base for each player. While apparently quite theoretical, this characterization is extremely useful to pin down sufficient conditions for the existence of a fully revealing equilibrium in large classes of games and economic applications. The first of these conditions is monotonicity. If the masquerading payoff of a player is increasing in the type

[3]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.

[4]This includes any situation in which players can certify their true type.

[5]The function w_i weakly represents the masquerade relation of player i iff, whenever type t_i of player i wants to masquerade as type s_i, we have $w_i(s_i) > w_i(t_i)$.

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). A notable exception is Giovannoni and Seidmann (2007) in which full revelation relies on a combination of two conditions:[6] single-peakedness of the masquerading payoff in the target type, and (in our terminology) a *no reciprocal masquerade* 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.

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 acyclic masquerade property holds whenever the masquerading payoff has *single crossing differences*,[7] *i.e.*, if the return from masquerading as a higher type is positive for a given true type, then it is also positive for higher true types.

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. An advantage of using increasing and single crossing differences is that there are well-known conditions under which they can be aggregated. In particular, the increasing differences condition can be aggregated when types are independent. But independence is often too restrictive in applications. Instead, we show that it is possible to work directly on the ex post masquerading relation to construct fully revealing *weak* sequential equilibria (in the sense of, e.g., Myerson (1991), it

[6]The conditions in Seidmann and Winter (1997) also imply these two conditions.

[7]Or, therefore, *increasing differences*. The terminology adopted is that of Milgrom (2004).

corresponds to what is usually called a perfect Bayesian equilibrium in the hard information literature). This applies, for example, to models with multiple senders and a single receiver: if the optimal action of the receiver is non-decreasing in types, the preferences of the players have complementarities in own type and action, then the ex post masquerading payoffs satisfy increasing differences. Then, there exists a fully revealing strong sequential equilibrium under type independence, and there exists a fully revealing weak sequential equilibrium regardless of the information structure.

To illustrate our method, we provide new applied results that contribute to different literatures. Our first application considers supermodular Bayesian games with complementarities between own actions and all types (as in Van Zandt and Vives, 2007). We show that if the preferences of the players also exhibit complementarities in own type and the actions of other players, then the ex post masquerading payoffs have increasing differences and there exists a fully revealing equilibrium.[8] Our second application contributes to the literature on deliberation before voting[9] by considering a general voting game that includes the jury model. This model can be applied to voting in a parliament for example, and has both voters and experts that testify in front of the voters. The experts could have evidence about the virtues of a proposal, and the members of the parliament may have evidence about how it would affect their constituency for example. The voters choose between two alternatives such that for each player the difference in payoff between the alternatives is non-decreasing in the types of all players. We show that the ex post masquerading payoffs satisfy increasing differences for each player under every non-unanimous rule, so that there is a fully revealing equilibrium of the voting game preceded by a debate. The case of unanimity is even simpler since the monotonicity condition is then satisfied for voters.

The sufficient conditions used above are especially suited for incomplete information games in which each player's type set is unidimensional. But the acyclic property and the weak representation of the masquerade relation can also be used to analyze information revelation

[8]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 fully revealing equilibrium.

[9]See, for example, Austen-Smith and Feddersen (2006), Gerardi and Yariv (2007), Jackson and Tan (2013), Lizzeri and Yariv (2011), Mathis (2011).

in games with multidimensional types. In particular, we prove existence of a fully revealing equilibrium in lobbying or conformity games with multidimensional objectives in which the masquerade relation of a player can be written as the sum of two terms: a first one maximized when the sender masquerades as her true type; a second one proportional to a function of the type that she masquerades as. We also study sender-receiver games where the sender has a multidimensional and type-dependent bias. In such games, for every type of the sender, the bias vector points to the direction towards which this expert wants to masquerade as. We provide sufficient conditions on the bias function to induce acyclic masquerades. These include cases in which the bias function is centrifugal (the sender wants to pretend she is further away from a central type than she really is) or mildly centripetal (the sender wants to pretend she is closer from a central type than she really is).

2 The Model

The Base Game. There is a set $N = \{1, \cdots, n\}$ of players who are to interact in a base game with action set[10] $A = A_1 \times \cdots \times A_n$. Each player i is privately informed about her type $t_i \in T_i$, where T_i is a finite set or a subset of $\mathbb{R}^K$, and $T = T_1 \times \cdots \times T_n$ is the set of type profiles endowed with its natural topology. Let $p(\cdot) \in \Delta(T)$ be a 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.[11] The preferences of the players are given by vNM utility functions $u_i : A \times T \to \mathbb{R}$.

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

Assumption 1. *For every type profile $t \in T$, the best reply correspondence of the game $\tilde{\Gamma}(t)$ is*

[10] This formulation does not exclude mixed strategy equilibria since each A_i can be replaced by the set of mixed actions $\Delta(A_i)$ and the utility functions could be extended to mixed actions in the usual way.

[11] We assume a common prior, 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 $t_i \in T_i$.

well defined, and the set of Nash equilibria of $\tilde{\Gamma}(t)$, denoted by $NE(t) \subseteq A$, is nonempty.

The Communication Phase. Before choosing their actions in A, but after learning their types, the players have the opportunity to publicly and simultaneously disclose hard evidence about their type at no cost. To formalize this, suppose that player i is restricted to send messages in a 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 $M = M_1 \times \cdots \times M_n$ 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 $M_i^{-1}(m_i) = S_i$. When T_i is not finite, we assume that all certifiable sets are compact subsets of T_i. The message structure satisfies *own type certifiability* if for every player i, every singleton $\{t_i\}$ is certifiable.

Fully Revealing Equilibria Our primary 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. 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 of the corresponding belief system obtained by Bayesian updating.[12] In the rest of the paper, the term *strong equilibrium* refers to this definition.

We will also use the notion of weak sequential equilibria in the sense of Myerson (1991). They are defined as equilibria that satisfy sequential rationality and weak belief consistency. Weak consistency here means Bayes consistency on the equilibrium path and off-path beliefs that are consistent with evidence.[13] It is implied by strong consistency. In the rest of the paper we refer to such equilibria as *weak equilibria*.

We are interested in equilibria of the augmented game in which all players perfectly reveal

[12] The notion of strong belief consistency in Kreps and Wilson (1982) is only defined for extensive form games with finite information sets; in general, it is hard to appropriately define sequential equilibria in infinite games (see Myerson and Reny, 2013); hence, when the type sets are not finite, we simply adopt the same restrictions on beliefs as those imposed by strong consistency in the finite case (see Lemma 1).

[13] They correspond to what most papers call perfect Bayesian equilibria. Because this term is used in many different ways in the literature, we find it clearer to use the terminology of Myerson (1991).

their type in the communication phase – henceforth, (weak or strong) *fully revealing equilibria*. In a fully revealing equilibrium, the second stage game on the equilibrium path corresponds to the complete information game $\tilde{\Gamma}(t)$, and therefore the action profile played on 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 fully revealing equilibrium of the augmented game such that $a^*(t)$ is played on the equilibrium path.

The Masquerade. 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 masquerade as another type s_i. For this purpose, let

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

denote the expected utility of player i on the equilibrium path of a fully revealing equilibrium if she is of type t_i, and

$$v_i(s_i|t_i) = E_{t_{-i}}\big(u_i\big(BR_i(a^*_{-i}(s_i, t_{-i}), t), a^*_{-i}(s_i, t_{-i}), t\big) \mid t_i\big),$$

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:

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

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*. We will assume the following continuity property[14] of $v_i(s_i|t_i)$. This assumption is automatically satisfied when T_i is finite. This assumption is not innocuous, but often satisfied in commonly

[14]We recall that $v_i(s_i|t_i)$ is lower semi-continuous in s_i if for every $\alpha \in \mathbb{R}$, the set $\{s_i \in T_i \mid v_i(s_i|t_i) > \alpha\}$ is open.

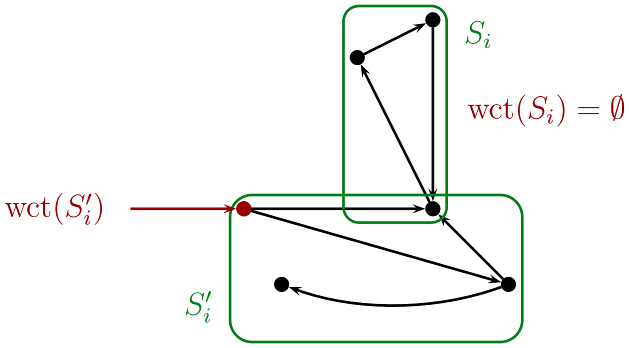


Figure 1: Masquerade relation and worst case types.

studied situations. Together with the assumptions of compactness of the certifiable subsets made above, it allows us to extend the results that hold for finite type sets to infinite type sets.

Assumption 2 (Semicontinuity)**.** *For every player i, the function $v_i(s_i|t_i)$ is lower semi-continuous in s_i.*

We can define a binary relation $\xrightarrow{\mathcal{M}}$ on T_i, the masquerade relation, that summarizes the incentives of different types to masquerade as one another.

Definition 1 (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 relation 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 in S_i would like to masquerade as:

$$\mathrm{wct}(S_i) := \left\{ s_i \in S_i \mid \nexists\, t_i \in S_i, t_i \xrightarrow{\mathcal{M}} s_i \right\}.$$

This set may be empty, or have more than one element. It is useful to represent the masquerade relation by a directed graph on T_i, as illustrated in Figure 1. A worst-case type for S_i is an element $s_i \in S_i$ with no incoming arrow from other elements of S_i.

Evidence Base. An evidence base is a set of messages that a player can use to certify each of her possible types.

Definition 2 (Evidence Base)**.** *An evidence base for player i is a set of messages $\mathcal{E}_i \subseteq M_i$ such that there exists a one-to-one function $e_i : T_i \to \mathcal{E}_i$ that satisfies $e_i(t_i) \in M_i(t_i)$ and $t_i \in \operatorname{wct}\big(M_i^{-1}\big(e_i(t_i)\big)\big)$ for every t_i.*

Equivalently, an evidence base provides each type t_i of player i a message $e_i(t_i)$ that certifies a set in which no type of player i would like to masquerade as t_i, i.e., $M_i^{-1}(e_i(t_i)) \subseteq \{s_i \in T_i : s_i \overset{\mathcal{M}}{\not\to} t_i\}$ for every $t_i \in T_i$. Note that 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 the masquerade relation. The set of evidence bases, however, depends on the masquerade relation. For example, if T_i can be linearly ordered such that the masquerade is monotonic (i.e., $t_i \overset{\mathcal{M}}{\longrightarrow} t_i'$ for every t_i' higher than t_i) as in Milgrom (1981) or Okuno-Fujiwara et al. (1990), then there is an evidence base if and only if each type can certify that it is in a subset for which her true type is minimum: for all $t_i \in T_i$, there exists $m_i \in M_i(t_i)$ such that $t_i = \min M_i^{-1}(m_i)$. In common interest games, i.e., in games in which no type would like to masquerade as any other type, there is an evidence base if and only if each type can simply send a different message.

As another illustration, consider a player i with three possible types, $T_i = \{t^1, t^2, t^3\}$, whose masquerade relation is given by $t^1 \overset{\mathcal{M}}{\longrightarrow} t^2 \overset{\mathcal{M}}{\longrightarrow} t^3$. The message correspondence $M_i(t^1) = \{m, m^{13}, m^{12}\}$, $M_i(t^2) = \{m, m^{23}, m^{12}\}$ and $M_i(t^3) = \{m, m^{23}, m^{13}\}$ admits two evidence bases: $\{m, m^{23}, m^{13}\}$ and $\{m^{12}, m^{23}, m^{13}\}$. In contrast, the message correspondence $M_i(t^1) = \{m, m^{12}\}$, $M_i(t^2) = \{m, m^2, m^{23}, m^{12}\}$ and $M_i(t^3) = \{m, m^{23}\}$ does not admit any evidence base because type t^3 has no message certifying an event for which it is a worst case type.[15] In Section 6, we provide more intuitive examples of evidence bases related to our applications.

The existence of an evidence base is important since it is necessary for a fully revealing equilibrium to exist.[16]

Remark 1 (Evidence Base: Necessity)**.** If there exists a fully revealing (weak or strong) equilibrium, then there must exist an evidence base $\mathcal{E}_i$ for every player i.

[15]These examples also show that the existence of an evidence base is not related to the "nested range condition" (Green and Laffont, 1986) or the "minimal closure" / "normality" condition (Forges and Koessler, 2005; Bull and Watson, 2007) used to get a revelation principle with hard evidence.

[16]Mathis (2008) makes the same observation for a class of sender-receiver games.

Indeed, consider a fully revealing equilibrium communication strategy profile σ that implements some Nash equilibrium $a^*(\cdot)$ of the contingent complete information games. Then the sets of messages sent with positive probability by each type t_i under σ_i must be disjoint across types t_i. Let $\hat{\sigma}_i(t_i)$ be a selection of one message in the support of $\sigma_i(t_i)$ for each t_i, and suppose that $t_i \notin \text{wct}(M_i^{-1}(\hat{\sigma}_i(t_i)))$. 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 $\sigma_i(t_i')$, that would contradict the fact that σ is an equilibrium. Therefore, the selection $\hat{\sigma}_i(\cdot)$ must form an evidence base for $M_i(\cdot)$.

3 Characterization of Fully Revealing Strong Equilibria with Extremal Beliefs

In this section, we provide necessary and sufficient conditions for the existence of a particular type of fully revealing equilibrium in which every deviation is attributed to a single type of the deviator. The first part of the section defines these *extremal beliefs* equilibria, and discusses the consequences of the restrictions they place on equilibrium beliefs.

Extremal Beliefs. In order to support a fully revealing equilibrium, players must be able to punish any player i who sends an off the equilibrium path message m_i. 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 off the equilibrium path beliefs after unilateral deviations to be extremal in the sense that they belong to the extreme points of the simplex $\Delta(T_i)$.

Definition 3 (Extremal Beliefs). *A fully revealing equilibrium with extremal beliefs is a fully revealing 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 by sending an off the equilibrium path message m_i, while every player $j \neq i$ sends an equilibrium message that reveals 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 beliefs assumption does not require all players other than i to attribute the same type t_i' to player i, but we will show in the next paragraph that this is required by strong consistency, and we will impose it when we only require weak 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 4. *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 for 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 to extremal beliefs only make it harder to find existence results in the sense that there may be games for which fully revealing equilibria exist but can only be constructed by violating our restrictions. We use them because, under these restrictions, the existence of fully revealing equilibria can be simply characterized by properties of the masquerade relation.

Strong Consistency and Extremal Beliefs. Strong consistency has important implications for the beliefs that can be held off the equilibrium path in fully revealing equilibria with extremal beliefs. We show that after any detectable unilateral deviation by player j sending message m_j, the belief formed by other players about the type of player j depends only on m_j.

In particular, all non-deviators form the same belief, independently of their type and of the messages sent by other non-deviators.

Lemma 1 (Consistent Extremal Beliefs)**.** *In a fully revealing strong equilibrium with extremal beliefs, after any unilateral deviation of some player j in the communication stage, the off-path beliefs of all players $i \neq j$ assign probability one to the same type $t_j \in T_j$ of player j independently of the message m_{-j} and their own type t_i.*

While this result is interesting and new (to the best of our knowledge), its proof is technical and we relegate it to Appendix B. The intuition is that if the belief μ formed after a unilateral deviation is extremal, it puts probability one on a single type t'_j. For μ to be consistent, there must be a sequence of Bayes-consistent beliefs μ^k that converges to μ and is generated by a sequence of completely mixed strategies of player j that put infinitely more weight on m_j when she is of type t'_j than when she is of any other type t''_j. But if this is the case, the information contained in the strategy of player j crowds out any information about j contained in the prior, and in particular any information that the non-deviators could derive from the correlation between t_j and their own type, or what they deduce on the types of other non-deviators from their messages.[17]

The Characterization. The existence of evidence bases for each player is necessary for the existence of a fully revealing equilibrium and it can be interpreted as a richness condition on the language saying that all private information can be credibly conveyed. Worst case types are important because they allow to discourage unilateral deviations to messages off the equilibrium path. In fact, Lemma 1 implies that with extremal beliefs the deviating message m_j must be attributed to a single type $t_j \in M_j^{-1}(m_j)$ that depends only on which message was sent. If this type is not a worst case type of $M_j^{-1}(m_j)$, then there must exist another type in $M_j^{-1}(m_j)$ that gets a higher payoff by sending m_j than in equilibrium.

[17]Note however that the full support assumption on type profile is fundamental for this property to hold. 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 Fudenberg and Tirole's (1991) definition of perfect Bayesian equilibrium when types are independently distributed.

These two conditions are also sufficient as we show in the following theorem. It is not difficult to construct a fully revealing equilibrium when they are satisfied: the messages from the evidence base should be used on the equilibrium path for the players to reveal their type, and detectable deviations should be attributed to worst case types. The difficulty of the proof is to show that the equilibrium we just constructed satisfies the strong belief consistency requirement.

Theorem 1 (Characterization)**.** *There exists a fully revealing strong equilibrium with extremal beliefs that implements $a^*(\cdot)$ on and off the equilibrium path if and only if the following conditions are satisfied for every i:*

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

(ii) The correspondence $M_i(\cdot)$ admits an evidence base.

Proof. See Appendix B. □

To conclude this section, we provide two examples. The first one illustrates the Theorem 1 and shows how adding messages can destroy the existence of a fully revealing equilibrium with extremal beliefs.

Example 1. Consider a sender-receiver game in which the sender's type set is given by $T = \{t^1, t^2, t^3, t^4\}$, the masquerade relation and the certifiable subsets are given in Figure 2. By Theorem 1, there exists a fully revealing equilibrium with extremal beliefs. If we add a new message m^5 that certifies $\{t^2, t^3, t^4\}$, this is no longer true. ◇

Our next example shows that the existence of worst case types is not necessary if interior beliefs are allowed off the equilibrium path.

Example 2 (Hidden Bias)**.** Consider a sender-receiver problem[18] in which the receiver can decide between two policies A and B, or keep the status quo ϕ. The sender can have information favorable to either of the policies A and B, and also has a bias which is unknown to the receiver. We denote the types of the sender by $T = \{aA, aB, bA, bB\}$, where type aB is biased towards

[18]Note that in sender-receiver models, there is no distinction between weak and strong equilibria.

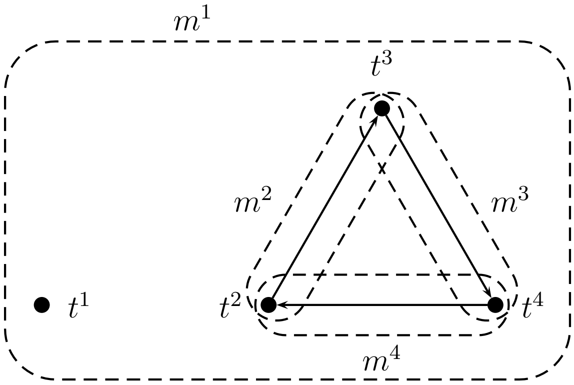


Figure 2: An example

A, and has receiver information favorable to B. We assume that all types are equally probable. The payoff matrix is given in the following table where the payoff of the sender appears first and the payoff of the receiver second.

	A	B	ϕ
aA	$1, 1$	$-1, -1$	s_ϕ, r_ϕ
aB	$1, -1$	$-1, 1$	s_ϕ, r_ϕ
bA	$-1, 1$	$1, -1$	s_ϕ, r_ϕ
bB	$-1, -1$	$1, 1$	s_ϕ, r_ϕ

Table 1: Hidden Bias – with $s_\phi, r_\phi < 1$.

The corresponding masquerade relation is represented in Figure 3. The sender can disclose her information A or B, or not disclose anything, so the certifiable sets are as represented on the figure. We assume that cheap talk is possible which means that there are several messages that certify the same subset (at least as many as there are types in the corresponding certifiable subset). We denote a generic message that certifies the complete set as m_0, and a generic message that certifies information favorable to policy X as m_X.

There exists an evidence base so full revelation is possible. Indeed, since cheap talk is allowed, there exists two messages m_A and m'_A that certify A, and that can be used respectively by aA and bA since they are both worst case types of the set $\{aA, bA\}$, and the same is true for bB and aB with two messages m_B and m'_B that certify B.

However the type set, which is certifiable by m_0, admits no worst case type, hence there is

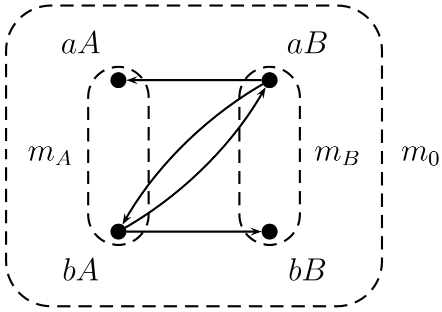


Figure 3: Hidden Bias.

no fully revealing equilibrium with extremal beliefs. We will show that a fully revealing equilibrium may nevertheless exist, depending on the values of s_ϕ and r_ϕ. In case no fully revealing equilibrium exists, we will characterize the receiver's optimal partially revealing equilibrium.

Suppose that there exists a fully revealing equilibrium, and consider a message m_0 that certifies T. This message must be off the equilibrium path since it has no worst case type. Also, it cannot be the case that the receiver responds to this message by mixing between A and B, for that would give a payoff higher than the full revelation payoff to either bA or aB. So it must be the case that the receiver chooses the status quo ϕ to respond to this message. There exists a belief that justifies the choice of the status quo by the receiver if and only if $r_\phi \geq 0$. It must also be the case that the choice of the status quo induces the sender to choose to reveal the truth rather than sending a message that reveals nothing. This is true if and only if $s_\phi < -1$. In summary, a necessary condition for the existence of a fully revealing equilibrium is that the receiver prefers the status quo to choosing randomly between the two policies, and that the sender always prefers her least favored policy to the status quo. It is easy to show that this condition is also sufficient by fixing the belief that follows any message m_0 to put equal weight on A and B.

Now suppose that $r_\phi < 0$ or $s_\phi > -1$. Then either the receiver never chooses the status quo regardless of her information, or the status quo is not an effective punishment and does not induce revelation from aB and bA. Then the best achievable situation from the point of view of the receiver is to be able to identify the types aA and bB. This is done by following each message m_0 by a belief that puts the same probability on aB and on bA, and choosing the status quo if $r_\phi \geq 0$ and any mixing between A and B otherwise. $\diamond$

4 Acyclic Masquerade

In this section, we define a class of games for which a worst case type exists for ever subset of types. Therefore the existence of fully revealing strong equilibria for this class of game is characterized only by a property of the message structure: the existence of an evidence base for each player.

4.1 Definition and Characterization

We say that a game Γ with a selection $a^*(\cdot)$ has the *interim acyclic masquerade property* if for every player i, the masquerade relation on T_i is acyclic. The following proposition characterizes acyclic masquerade relations.[19]

Proposition 1. *The following statements are equivalent*

(i) The masquerade relation of player i is acyclic.

(ii) Every finite subset $S_i \subseteq T_i$ admits a worst case type.

(iii) There exists a lower semi-continuous function $w_i : T_i \to \mathbb{R}$ such that

$$t_i \xrightarrow{\mathcal{M}} s_i \;\Rightarrow\; w_i(s_i) > w_i(t_i). \tag{WR}$$

(iv) There exists a complete, transitive and lower semi-continuous order $\succeq$ on T_i, such that

$$t_i \xrightarrow{\mathcal{M}} s_i \;\Rightarrow\; s_i \succ t_i. \tag{DM}$$

Proof. Suppose that $\xrightarrow{\mathcal{M}}$ has a cycle $t_i^1 \xrightarrow{\mathcal{M}} \cdots \xrightarrow{\mathcal{M}} t_i^k \xrightarrow{\mathcal{M}} t_i^1$ on T_i. Then $S_i = \left\{t_i^1, \cdots t_i^k\right\}$ does not have a worst case type. Now suppose that there exists $S_i \subseteq T_i$ such that $\text{wct}(S_i) = \emptyset$. Let

[19] An order $\succeq$ on T_i is lower semi-continuous if for every t_i, the set $\{s_i \in T_i \mid s_i \succ t_i\}$ is open. It is complete if for every s_i and t_i in T_i, either $s_i \succeq t_i$ or $t_i \succeq s_i$, and transitive if $t_i \succeq t_i'$ and $t_i' \succeq t_i''$ implies that $t_i \succeq t_i''$.

$s_i^1 \in S_i$. Because $\text{wct}(S_i) = \emptyset$ 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. This shows the equivalence of between the equivalence of (i) and (ii).

The equivalence between (i) and (iii) derives from Alcantud and Rodríguez-Palmero (1999), and the lower semi-continuity assumption on $v_i(s_i|t_i)$. The function w_i induces a complete, transitive and lower semi continuous order on T_i defined by $s_i \succeq t_i \Leftrightarrow w_i(s_i) \geq w_i(t_i)$, and by (iii) it must be true that $t_i \xrightarrow{\mathcal{M}} s_i \;\Rightarrow\; s_i \succ t_i$. Hence (iii) implies (iv). It is easy to see that (iv) implies that the masquerade relation is acyclic. □

In the proposition, (DM) stands for *Directional Masquerade*. It says that there is an order such that all types only want to masquerade as types that are their successors in this order. (WR) stands for *Weak Representation* since, in the language of the literature on the representation of binary relations, the function $w_i(\cdot)$ is called a weak representation of the masquerade relation. Condition (ii) means that we can find a worst case type on every subset of T_i in the finite case. In the infinite case, we would like to have a similar property. Since the only subsets on which we need worst case types are the certifiable ones, and we restricted certifiable subsets to be compact subsets, it is sufficient to show that we can find worst case types on every compact subset of T_i.

Lemma 2. *The interim acyclic masquerade property implies that for every i, every compact subset S_i of T_i admits a worst case type.*

Proof. We can use the weak representation $w_i(\cdot)$ of Proposition 1. Since it is lower semi-continuous, it reaches a minimum on every compact subset $S_i \subseteq T_i$, and the minimizer is a worst case type of S_i. □

From Lemma 2 and Theorem 1 we can immediately deduce that, in the class of games with the acyclic masquerade property, the existence of an evidence base for each player is a sufficient condition for the existence of a fully revealing strong (and therefore weak) equilibrium. From Remark 1 we know that it is also necessary.

Corollary 1. *Suppose that the interim acyclic masquerade property is satisfied. Then there exists a fully revealing (weak or strong) equilibrium that implements $a^*(\cdot)$ if and only if there exists an evidence base for every player i.*

We can directly apply Proposition 1 to the case in which the masquerading payoffs of informed players are independent of their type. For instance, this is true in the seller-buyer example of Milgrom (1981) and in the multidimensional cheap talk model of Chakraborty and Harbaugh (2010). In this case, we can represent the masquerade relation by the function $w_i(s_i) = v_i(s_i|t_i)$, which leads to the following remark.

Remark 2. Suppose that the interim masquerading payoff of each player is independent of her type. Then the interim acyclic masquerade property is satisfied.

We showed in Example 1 that adding messages could destroy full revelation. Another interesting remark is that this is no longer true for games satisfying the interim acyclic masquerade property. This is just because evidence bases are preserved under the addition of new messages, and the new certifiable subsets that are created must admit worst case types by the interim acyclic masquerade property.

4.2 Sufficient Conditions on Masquerading Payoffs

The following theorem provides a list of sufficient conditions for the masquerade relation to be acyclic. (MON) stands for *Monotonicity*, (ID) and (SCD) stand for *Increasing Differences* and *Single Crossing Differences*. (SP-NRM) is a set of two conditions, *Single Peakedness* and *No Reciprocal Masquerade*. For a reminder of standard definitions used in the statement of this theorem, see Appendix A.

Theorem 2 (Sufficient Conditions)**.** *The interim acyclic masquerade property is satisfied whenever for every i there exists a linear order $\succeq$ on T_i such that any of the following conditions is satisfied:*

(MON) *$v_i(s_i|t_i)$ is non-decreasing in s_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. See Appendix B. □

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 revelation. 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.

The sender-receiver game of Crawford and Sobel (1982) does not satisfy the (MON) property, but it satisfies (DM) because 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, the sender only wants to masquerade as a higher type. This however does not mean that she wants to masquerade as any higher type. In this case it is easy to see that (DM) is satisfied for the natural order on types. In general however it may be difficult to find an order under which (DM) holds.[20]

To our knowledge, this paper is the first to show that (SCD) and (ID) are sufficient conditions for the existence of fully revealing 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 $t'_i \succeq t_i$. The condition (SP-NRM) is

[20]In particular, the order on types for which (SCD) or (ID) holds may differ from the order induced by (DM).

used in Giovannoni and Seidmann (2007) to show the existence of a fully revealing equilibrium in a sender-receiver model.[21] The next subsection illustrates how our conditions can be directly used to get an existence result when there are multiple and asymmetrically informed senders.

To prove Theorem 2, we showed that each condition implies that the masquerade relation is acyclic. It leaves open the question of identifying the worst-case types, that is, of how to be skeptical when being certified a subset S_i of T_i. It is easy when (MON) or (DM) hold since, in any subset of types S_i, the lowest type is a worst-case type. In Appendix C, Proposition 7 also shows how to find worst case types under (SCD) and (SP-NRM).

Ex post Masquerade and Aggregation. In applications it is often easier to work with the ex post masquerading payoffs. Then we can use aggregation results to show that the interim acyclic masquerade property is satisfied. In the following lemma we recall some simple aggregation results that are useful for the applications.[22]

Lemma 3 (Ex Post Sufficient Conditions)**.** *The interim acyclic masquerade property is satisfied 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 non-decreasing 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 increasing differences in* (s_i, t_i)*, and types are independent.*

Proof. See Appendix B. ◻

In the next section, we develop a different way to work with ex post masquerade payoffs which allows us to construct fully revealing equilibria regardless of the information structure. The cost of this approach is that we need to weaken the equilibrium concept.

[21]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 existence result is also a direct corollary of Theorem 2 based on the (SP-NRM) condition.

[22]For more advanced aggregation results, we refer the reader to Quah and Strulovici (2012).

5 Weak Equilibria and Ex Post Masquerade

This section develops the idea that in order to enforce full revelation, the players can be skeptical by attributing messages that deviate from full revelation to a worst case type of the ex post masquerade relation. Indeed, if all players but i have revealed their type, the other players can condition their belief on t_{-i}. The caveat is that such out of the equilibrium path beliefs violate one of the implications of strong sequential equilibria with extremal beliefs that we derived in Lemma 1. So a cost of this approach is that it sacrifices strong consistency of beliefs. The benefit is that these equilibria are easier to construct in applications, and that the punishments used off the equilibrium path do not depend on the specifics of the information structure. For all this section we will assume that for every player i the function $v_i(s_i|t_i;t_{-i})$ is lower semicontinuous in s_i.

The Equilibrium Notion. In this section we work with fully revealing weak equilibria (see Section 2). We only seek to provide a sufficient condition for the existence of such equilibria. To do that we construct equilibria with extremal beliefs such that off the equilibrium path beliefs following a unilateral deviation are homogeneous across non-deviators.[23] The equilibria we construct also satisfy Definition 4. To summarize, the equilibria that we construct in this section are fully revealing weak equilibria with homogeneous extremal beliefs that implement the selection $a^*(\cdot)$ on and off the equilibrium path.

Ex post Masquerade Relation and Full Revelation. We start by adapting our definitions to the ex post treatment. For every t_{-i}, the ex post masquerade of player i given t_{-i} is the relation defined by $t_i \xrightarrow{\mathcal{M}(t_{-i})} s_i$ if and only if $v_i(s_i|t_i;t_{-i}) > v_i(t_i|t_i;t_{-i})$. The set of ex post worst case types of the set $S_i \subseteq T_i$ given t_{-i} is defined by $\mathrm{wct}(S_i \mid t_{-i}) := \{s_i \in S_i \mid \nexists\, t_i \in S_i, t_i \xrightarrow{\mathcal{M}(t_{-i})} s_i\}$. We say that a game satisfies the *ex post acyclic masquerade property* if for every player i, and every t_{-i}, the ex post masquerade relation of i given t_{-i} is acyclic. The characterization of acyclic masquerade relations in Proposition 1 holds for the ex post

[23] Note that with weak equilibria, Lemma 1 no longer applies, so homogeneity is not imposed by the equilibrium concept.

masquerade if we condition each statement on t_{-i} and replace masquerade relation by ex post masquerade relation, and worst case type by ex post worst case type. The sufficient conditions in Theorem 2 hold provided that the interim masquerade payoffs are replaced by the ex post masquerade payoffs. Then we have the following result

Theorem 3 (Weak Sequential Equilibria)**.** *There exists a fully revealing weak equilibrium with extremal and homogeneous beliefs that implements $a^*(\cdot)$ whenever the following conditions are satisfied for every i*

(i) For every t_{-i}, the set $M_i^{-1}(m_i)$ admits an ex post worst case type.

(ii) The correspondence $M_i(\cdot)$ admits an evidence base.

Proof. See Appendix B. □

Because the ex post acyclic masquerade property implies the existence of ex post worst case types on every certifiable subset, we have the following corollary.

Corollary 2. *Suppose that the ex post acyclic masquerade property is satisfied. Then there exists a weak fully revealing equilibrium that implements $a^*(\cdot)$ if and only if there exists an evidence base for every player i.*

The following example illustrates why it is useful to work with the ex post masquerade relation. In this multiple senders example, we obtain the existence of a fully revealing weak equilibrium under mild assumptions on the preferences of the players, and no assumptions on the type distribution. To prove the existence of a fully revealing equilibrium that satisfies strong belief consistency by an aggregation result, we would have to either assume that types are independent and use Lemma 3 (*iii*), or make some unnatural assumptions on the utilities and use a more sophisticated aggregation result.

Example 3 (Multiple Senders - Single Receiver Games)**.** One player with no private information, the receiver, wants to implement her ideal action $a^*(t) \in \mathbb{R}$. The partially and asymmetrically informed players, the senders, are indexed by i. T_i is a (possibly finite) compact subset of $\mathbb{R}$ endowed with its natural order. The lower semi-continuity assumption is ensured if for every i, $u_i\big(a^*(s_i,t_{-i}),t_i,t_{-i}\big)$ is lower semi-continuous in s_i. Assume that:

(i) $a^*(\cdot)$ is non-decreasing.

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

Under these assumptions, $v_i(s_i|t_i;t_{-i}) = u_i\big(a^*(s_i,t_{-i}),t_i,t_{-i}\big)$ has increasing differences in (s_i,t_i), and therefore the ex post acyclic masquerade property is satisfied. To see that, take $s'_i \succ s_i$ and $t'_i \succ t_i$ and note that

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

where the inequality comes from the fact that $a^*(s'_i,t_{-i}) \geq a^*(s_i,t_{-i})$ by (i), and from (ii). Therefore there exists a fully revealing weak equilibrium as long as we have an evidence base for every player. If types are independent, then there exists a fully revealing strong equilibrium. $\diamond$

6 Applications

6.1 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 5. *We say that the (Bayesian) base game $\Gamma = \big\langle N,T,A,p,(u_i)_{i\in N}\big\rangle$ is supermodular if each $u_i(a,t)$ is supermodular in a_i, has increasing differences in (a_i,a_{-i}) (strategic comple-*

mentarities), and has increasing differences in (a_i, t) (complementarities between own actions and type profiles).

It is well known[24] that in this case $NE(t)$ is a complete lattice, and that its extremal elements are non-decreasing 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 2 (Supermodular Games 1, Van Zandt and Vives, 2007)**.** *Suppose that Γ is supermodular and let $a^*(\cdot)$ be either the highest equilibrium selection or the lowest equilibrium selection. Then the interim acyclic masquerade property is satisfied whenever for every i either of the following assumptions is satisfied:*

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

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

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

$$\begin{aligned} u_i\big(BR_i\big(a^*_{-i}(s_i', t_{-i}), t\big), a^*_{-i}(s_i', t_{-i}), t\big) &\geq u_i\big(BR_i\big(a^*_{-i}(s_i, t_{-i}), t\big), a^*_{-i}(s_i', t_{-i}), t\big) \\ &\geq u_i\big(BR_i\big(a^*_{-i}(s_i, t_{-i}), t\big), a^*_{-i}(s_i, t_{-i}), t\big), \end{aligned}$$

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 Lemma 3. The proof is similar with negative externalities. □

Hence, with the positive or negative externality assumption, the acyclic masquerade property holds 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

[24] See Milgrom and Roberts (1990) and Vives (1990).

on supermodular games. In order to do so, however, we need to make additional regularity assumptions.

In the remainder of this subsection, 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 Θ_i. We assume that the utility functions $u_i(\cdot)$ are defined on $A \times \Theta$ where $\Theta = \Theta_1 \times \cdots \times \Theta_n$, and that they are continuously differentiable. Finally, we assume that every equilibrium action $a_i^*(t)$, and every best-response $BR_i\big(a_{-i}^*(s_i, t_{-i})|t_i; t_{-i}\big)$ is interior. Altogether, these assumptions ensure that the best-responses $BR_i\big(a_{-i}^*(s_i, t_{-i})|t_i; t_{-i}\big)$ always satisfy a first-order condition, so that the derivatives of the ex post masquerading payoff $v_i(s_i|t_i; t_{-i})$ can be obtained by the envelope theorem. Then the only additional assumption needed to ensure that the the ex post masquerading payoff has increasing differences is that the utilities of the players have increasing differences in their own type and the actions of the others.

Proposition 3 (Supermodular Games 2)**.** *Assume that the base game Γ is supermodular, that the utility functions are continuously differentiable on $A \times \Theta$ and that every best-response $BR_i\big(a_{-i}^*(s_i, t_{-i})|t_i; t_{-i}\big)$ is interior. Let $a^*(\cdot)$ be either the highest equilibrium selection or the lowest equilibrium selection. Then, the ex post acyclic masquerade property is satisfied whenever $u_i(a_i, a_{-i}, t)$ has increasing differences in (a_{-i}, t_i). If in addition types are independent, then the interim acyclic masquerade property is satisfied.*

Proof. See Appendix B. □

An immediate corollary of this result is obtained if we replacing the condition in Proposition 3 by a separability condition between own type and others' actions.

Corollary 3. *The ex post acyclic masquerade property is satisfied whenever every best-response is interior and 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). If in addition types are independent, then the interim acyclic masquerade property is satisfied.

The two following examples are based on recent papers extending Crawford and Sobel (1982) to multi-player cheap-talk (Hagenbach and Koessler, 2010, Galeotti et al., 2013). Using Corollary 3, we show that these games satisfy the ex post acyclic masquerade condition under fairly general conditions.

Example 4 (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 non-decreasing. Players also want to coordinate their own actions with those of other players. Their utilities are given by

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

where the α_{ij} are non-negative coefficients, normalized so that $\sum_j \alpha_{ij} = 1$, and such that $\alpha_{ii} > 0$.[25] It is easy to check that Corollary 3 applies, so by Corollary 2 this game has a fully revealing weak equilibrium as long as every player has an evidence base. It has a fully revealing strong equilibrium if in addition types are independent. ◇

Example 5 (An Influence Game)**.** Galeotti et al. (2013) 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. Again, if $\theta_i(\cdot)$ is non-decreasing, Corollary 3 applies. ◇

In a simple version of the two previous examples where players' biases are constant ($\theta_i(t) = \theta(t) + b_i$), players can be divided into two groups depending on whether their biases are relatively

[25] Particular forms of this class of games have been extensively studied in the economic theory of organizations as, for example, in Alonso et al., 2008 and Rantakari, 2008.

high or low compared to others' biases.[26] Intuitively, a player with a relatively low (high) bias would like to appear only as a lower (higher) type than she truly is. Therefore, when other players skeptically interpret any vague statement of a player as the highest (lowest) type, she has no interest to deviate from full revelation. In this case, the ex-post directional masquerade condition of Lemma 3 is satisfied, so there is a fully revealing strong equilibrium whatever the prior distribution of types. There is an evidence base whenever each player whose bias is relatively low is able to certify a set of types in which her actual type is maximum, and each player whose bias is relatively high is able to certify a set of types in which her actual type is minimum. This intuitive evidence base is easy to exhibit here because the relative bias of a player is independent of her private information. In more general cases, such as non constant biases, the ex-post directional masquerade may not be satisfied but our results still guarantee the ex post acyclic masquerade property when $\theta_i(\cdot)$ is non-decreasing.

6.2 Deliberation with Hard Information

In this subsection, 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 into 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 player derives from the status quo to 0, and we denote by $\mathfrak{u}_i$ the uncertain payoff she derives from the proposal. Each player i has a private signal t_i about the proposal. We assume that the function $U_i(t) = E\big(\mathfrak{u}_i | t_1, \cdots, t_n\big)$ is non-decreasing in t. This is the case for example if every player believes the vector $(\mathfrak{u}_i, t_1, \cdots, t_n)$ to be affiliated.

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 fully revealing equilibrium

[26] For example, in the coordination game, when players' coordination motives are symmetric ($\alpha_{ij} = \alpha$, $i \neq j$), a player with a relatively low (high, respectively) bias is simply a player whose bias is smaller (higher, respectively) than the average bias in the population. Otherwise, see Appendix D for the precise meaning of a "relatively high/low" bias in the two previous examples.

that implements the sincere voting equilibrium. We interpret the pre-play communication game as deliberation with hard evidence. In the complete information voting game, the sincere best response of $i \in \mathcal{C}$ is to vote in favor of the proposal whenever $U_i(t) > 0$. The acceptance set of a player is the set of type profiles such that she favors the proposal, $\mathcal{A}_i = \{t \in T \mid U_i(t) > 0\}$.

Example 6 (The Jury Model)**.** The question of voting with private information and deliberation is 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 ω is given by a probability π 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. Then for this model, we have

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

which is indeed increasing by affiliation. Then the region $\mathcal{A}(i)$ of the type set T over which voter i favors conviction (the proposal) is characterized by

$$t \in \mathcal{A}_i \;\Leftrightarrow\; \prod_{i=1}^n \frac{q(t_i|G)}{q(t_i|I)} \geq \frac{(1-\pi)\gamma_i^C}{\pi \gamma_i^A},$$

where the expression on the left-hand side is non-decreasing in t by affiliation. Therefore we can order the players according to $\frac{(1-\pi)\gamma_i^C}{\pi\gamma_i^A}$, and the sets $\mathcal{A}_i$ are non-decreasing in i in the set containment order $\mathcal{A}_1 \subseteq \mathcal{A}_2 \subseteq \cdots \subseteq \mathcal{A}_n$. Hence the acceptance sets of the players are naturally nested. ◇

Example 7 (Altruistic Voters)**.** Suppose that the individual expected payoff of player i from the alternative is given by a non-decreasing function $\psi_i(t_i)$ 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

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

where $\varepsilon_i \in [0, 1]$. This example also satisfies our assumptions but in contrast to the jury model the players' acceptance sets are typically not nested. $\diamond$

Consider now our general model of deliberation before voting, and any rule such that $q \leq C$. For committee members, ex post masquerading payoffs are given by:

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

where $S_i(s_i, t_{-i}) = \sum_{j \in \mathcal{C} \setminus \{i\}} \mathbb{1}_{U_j(s_i,t_{-i})>0}$ is the tally of votes in favor of the alternative among all voters except i. Under the unanimous rule such that $q = C$, these payoffs take the simpler form of $v_i(s_i|t_i; t_{-i}) = U_i(t)\mathbb{1}_{U_i(t)>0}\mathbb{1}_{S(s_i,t_{-i})\geq C-1}$ which is non-decreasing in s_i. The monotonicity property is easy to understand under unanimity as every voter is in one of two situations ex post. If, on the one hand, she wants to prevent the proposal from being adopted, then she can do so by voting against it, which makes deviation from full revelation pointless. If, on the other hand, she prefers the proposal, then she only wants to masquerade as a higher type so as to increase the number of votes in favor of the proposal. Every vague message coming from a voter can then be skeptically interpreted as coming from the type most favoring the status quo.

For general rules however, voters' ex post masquerading payoffs are not monotonic. The next lemma shows that these payoffs have increasing differences in (s_i, t_i). This is because masquerading as a higher type induces more agents to vote for the alternative which is more rewarding for a high true type than for a low one as $U_i(t)$ is non-decreasing in t. The same holds for agents who do not belong to the committee, and whose masquerading payoff are:

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

Lemma 4. *For every $i \in N$, $v_i(s_i|t_i; t_{-i})$ has increasing differences in (s_i, t_i). Under unanimity,*

$v_i(s_i|t_i, t_{-i})$ *is nondecreasing in* s_i *for every* $i \in \mathcal{C}$.

Proof. For any $i \in \mathcal{C}$ and $t'_i \succeq t_i$, the difference

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

is non-decreasing in s_i since $S_i(s_i, t_{-i})$ is non-decreasing in s_i. For every $i \in N\backslash\mathcal{C}$ and every $t'_i \succeq t_i$, the difference $v_i(s_i|t'_i;t_{-i}) - v_i(s_i|t_i;t_{-i}) = \underbrace{\left(U_i(t'_i,t_{-i}) - U_i(t_i,t_{-i})\right)}_{\geq 0}\mathbb{1}_{S_i(s_i,t_{-i})\geq q}$ is nondecreasing in s_i. Finally, under unanimity, the ex post masquerading payoff of a player $i \in \mathcal{C}$ is $v_i(s_i|t_i;t_{-i}) = U_i(t)\mathbb{1}_{U_i(t)>0}\mathbb{1}_{S(s_i,t_{-i})\geq C-1}$ which is nondecreasing in s_i. $\square$

Then we immediately have the following result.

Proposition 4. *Under any voting rule, if* $a^*(\cdot)$ *is the sincere voting equilibrium, then the ex post acyclic masquerade property is satisfied. If* $\mathcal{C} = N$ *and the rule is unanimity, or if types are independent, then the interim acyclic masquerade property is satisfied.*

Therefore, there exists a fully revealing weak equilibrium that implements the sincere voting equilibrium as long as there exists an evidence base for each player. While other results in the voting literature suggest that unanimity may perform less well than other voting rules in terms of information revelation,[27] our results imply that with hard information, any voting rule can lead to full revelation. Schulte (2010) shows this result for the specific case of the jury model, and Mathis (2011) extends it to the case in which preferences lead to nested acceptance sets. We extend these results by showing that full revelation holds for all preferences that react to information in the same direction, even when acceptance sets are not nested.

When acceptance sets are nested, it is relatively easy to understand how to be skeptical, that is how to find a worst-case type. In this case, the identity of the pivotal voter in the

[27] See, e.g., Austen-Smith and Feddersen (2006). Gerardi and Yariv (2007) show that when voting is augmented with a cheap talk communication stage, all voting rules that differ from unanimous adoption or unanimous rejection have the same set of equilibria, while the sets obtained under any of the unanimous rules are subsets of the latter.

full information voting game is independent of the realization of t: the pivotal voter i^* is the one with the qth largest acceptance set among members of the committee. Clearly, i^* has no incentive to masquerade as any other type regardless of her true type. She therefore has an evidence base as long as she can send a different message for each of her type (in particular, cheap talk is enough to get full information revelation from her). For other players, either their acceptance set is contained in the acceptance set of i^*, or their acceptance set contains that of i^*. So ex post they are respectively either more opposed to the proposal or more in favor of the proposal than i^*. A voter of the first kind only ever wants to masquerade as a lower type so as to undermine the proposal, whereas a voter of the second type only ever wants to masquerade as a higher type. Hence, skeptical beliefs for any message sent by a player of the first kind consists in interpreting her message as stemming from a type with the most favorable information for the proposal. Conversely, for players of the second kind, skepticism consists in believing the information most favorable to the status quo. So, with nested preferences, the ex post directional masquerade property of Lemma 3 holds so the fully revealing equilibria we construct satisfy strong consistency. Furthermore, there is an evidence base for each player whenever the players who are more favorable to the proposal than the pivotal voter are able to provide any evidence in favor of it, and the others are able to provide any evidence against it.

6.3 Multidimensional Types

Norms, Lobbies, and Rewards for Masquerading. The Weak Representation approach can be fruitfully applied to show that communication games with multidimensional types satisfy the acyclic masquerade property. We start with an example inspired from the theory of conformity of Bernheim (1994). In Bernheim (1994), an agent has type in $\mathbb{R}$ and must perform an action in $\mathbb{R}$. She wants her action to be as close as possible to her type, but she also wants other agents to believe that her type is close to a norm. In our version, the type is no longer one-dimensional, and the agent sends hard information about her type instead of performing an action.

Example 8 (Conformity with Multiple Norms)**.** We consider a sender-receiver model where

T is the type set of the single sender. Here T can be any metric space, but for simplicity let $T \subseteq \mathbb{R}^K$. There is a single receiver who takes action, but potentially many other agents who do not take action but form beliefs about the type of the sender. We assume that the optimal action of the receiver if she knows t is $a(t) = t$. As in Bernheim (1994), the payoff of the sender has two components. On the one hand, she would like the receiver to implement the optimal action $a(t)$. On the other hand, she would like to conform to one of several prevailing stereotypes in society. To model that second part, suppose that upon convincing other agents that she is of type s, the sender derives a payoff proportional to $-d(s, C)$, where $C \subseteq \mathbb{R}^K$ is a finite set of social stereotypes, and $d(s, C) = \min_{c \in C} d(s, c)$ is the Euclidean distance to that set. Alternatively, the elements of C can be interpreted as the positions of lobbies that reward experts for producing information close to their positions. So the masquerading payoff of the sender can be written as

$$v(s|t) = -d(s,t) - \lambda(t) d(s, C),$$

where $\lambda(t) > 0$. The term $\lambda(t)$ captures the weight that the sender puts on the different components of the masquerading payoff, and it can vary across types (see Appendix E for a 2-dimensional illustration and $|C| = 3$ social stereotypes). It is easy to show that the masquerade relation generated by these payoffs satisfies (WR). Indeed, we have

$$v(s|t) > v(t|t) \Leftrightarrow \lambda(t)\big(d(t, C) - d(s, C)\big) > d(s,t) \Rightarrow -d(s, C) > -d(t, C),$$

so we can use $-d(\cdot, C)$ as a weak representation of the masquerade relation. Hence, there is an evidence base whenever each type is able to certify that he is at least as close to the set of social stereoptypes as he actually is. In that case, there exists a fully revealing equilibrium in which any message is skeptically interpreted as stemming from a sender who is at the maximum distance (consistent with the evidence contained in the message) from the set of social stereotypes. The interpretation is interesting as it implies that communication with evidence coupled with skepticism on the side of the receiver mutes the effect of social stereotypes. $\diamond$

In fact the logic of this example can be generalized to any masquerade relation which is the sum of three terms, where one term is maximized when the sender masquerades as her true type, the second term is proportional to a function of the type that she masquerades as, and the third term depends on the true type only.[28]

Proposition 5. *Suppose that $v(s|t) = f(s,t) + \lambda(t)g(s) + h(t)$, where $f(\cdot,t)$ admits a unique maximum $f(t,t)$, and $\lambda(t) > 0$. Then the masquerade relation associated to this masquerading payoff is acyclic.*

Proof. See Appendix B. $\square$

Biases. A more common model is to think of experts as being biased. For that part, we assume that $T \subseteq \mathbb{R}^K$. We assume that the masquerading function takes the form $v(s|t) = -\big(s - t - b(t)\big)'\Omega\big(s - t - b(t)\big)$, where $b : T \to \mathbb{R}^K$ is a bias function, and Ω is a symmetric positive semi-definite matrix. For example, if Ω is the identity matrix, then the masquerading payoff is $-||s - t - b(t)||^2$. A nice way to think of the bias function is to visualize it as a vector field on $\mathbb{R}^K$ such that at each point t, the vector $b(t)$ points to the direction towards which t would like to masquerade. In fact, if $t + b(t)$ is in T, it is exactly the type that t would prefer to masquerade as. We will provide several conditions on the bias function $b(\cdot)$ that ensure acyclicity of the masquerade relation. In every case, one of the conditions is that the vector field $b(t)$ (or a straightforward transformation of it) can be obtained as the gradient of a potential function $\phi(t)$.[29] This condition is never sufficient and needs to be completed by an assumption on $\phi(\cdot)$, which can always be interpreted as $\phi(\cdot)$ not being too concave.

The reason why the potential from which $b(\cdot)$ derives should not be too concave can be easily understood in a one dimensional example. For this consider Figure 4. In Figure 4 (a), $b(\cdot)$ derives from a convex potential, and as a consequence the vector field is centrifugal. Then it is easy to be skeptical about any message: a worst case type for every compact subset S of $\mathbb{R}$ is the point of S which is closest to the set of minimizers of $\phi(\cdot)$. In (b) and (c), the biases derive from a concave potential, and as a consequence the vector field is centripetal. Let t^* be

[28] Note that the result of Proposition 5 also extends the observation made in Remark 2.

[29] A vector field that satisfies this property is called a conservative vector field, and when $K = 3$ this is equivalent to having curl 0, that is $\nabla \times b(t) = 0$.

the maximizer of the potential $\phi(\cdot)$. Intuitively, centripetal biases may be problematic because the types to the right of t^* may want to pretend that they are to the left of t^* and vice versa creating cycles in the masquerade relation as in (c). If the intensity of biases tends to vanish for types close to t^*, however, these cycles will not be created. This is the case when the potential function is not too concave as in (b). The following propositions show that the same intuitions hold in the multidimensional case.

Before stating the result, we explain why centrifugal and centripetal biases may be relevant in practice. For this suppose that the type/action space is a one dimensional policy space (a real interval), the sender an expert and the receiver a politician. Suppose, on the one hand, that the politician reacts to the information of the expert, but also tries to cater to the median voter who thinks that the ideal policy is somewhere in the middle of the policy space, call it t_{med}. Then the politician will always pick a policy between the ideal policy of the expert and that of the median voter, and an expert who observes a signal to the right of t_{med} will want to pretend that she is further to the right than she really is, and symmetrically for the left. If, on the other hand, you suppose that it is the expert who has an incentive to cater to a consensual position, then she will have a centripetal bias. In a multidimensional setting, there may be issues on which the politician needs to cater to a median voter, and other issues on which the expert needs to cater to a consensual position (see the examples discussed after Proposition 6).

The result is the following. The proof shows directly that the masquerade is acyclic, and is relegated to the appendix.

Proposition 6. *Suppose that $b(t)$ is continuously differentiable and satisfies, for every t, $Db(t) + I \geq 0$, where $Db(t)$ is the Jacobian of $b(t)$, and $\geq$ is in the sense of positive semi-definite matrices. Suppose in addition that there exists a function $\phi : \mathbb{R}^K \to \mathbb{R}$ such that for every $t \in T$, $\Omega b(t) = \nabla\phi(t)$. Then the masquerade relation is acyclic.*

Proof. See Appendix B. ∎

Remark 3. The sense in which the result requires that the $\phi(\cdot)$ is not too concave is the following: the Hessian of $\phi(\cdot)$ is given by $\Omega Db(t)$, and we have $\Omega(Db(t) + I) \geq 0$ because Ω

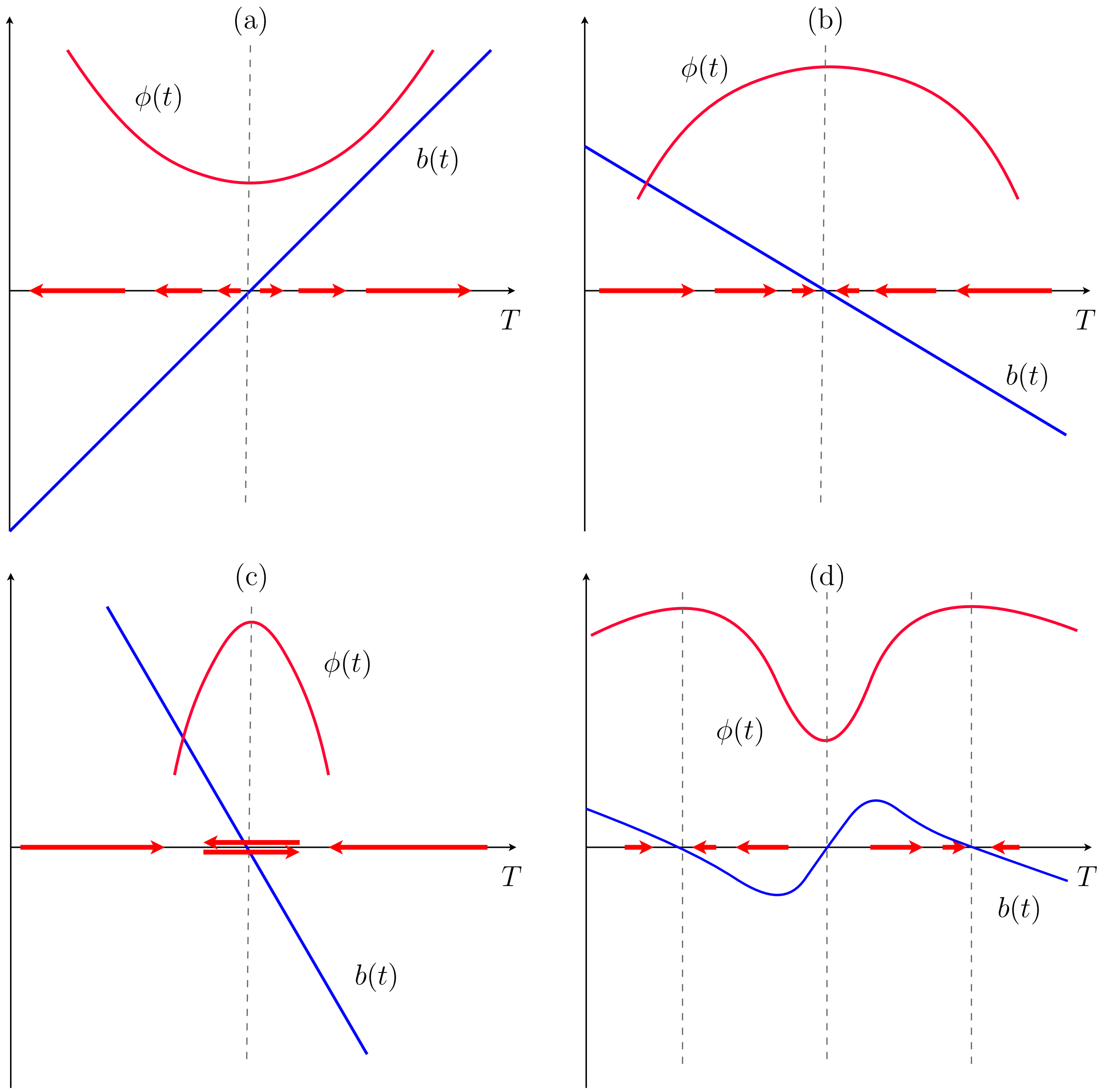


Figure 4: Illustration of Proposition 6 (biases) in the unidimensional case. The acyclic masquerade property is satisfied in figures (a), (b) and (d) because $\phi(t) + \frac{1}{2}t^2$ is convex, which is not satisfied in figure (c).

and $Db(t) + I$ are both positive semidefinite. But $\Omega(Db(t) + I)$ is the Hessian of the function $\psi(t) = \phi(t) + \frac{1}{2}t'\Omega t$ which must therefore be convex. So $\phi(\cdot)$ is not too concave in the sense that it must become convex when summed with the convex function $\frac{1}{2}t'\Omega t$.

To illustrate this proposition, consider the easy case in which $\Omega = I$ and $b(t)$ is the gradient of the concave function $-\frac{1}{2}||t||^2$. Then $Db(t) = -I$ and the conditions of the proposition hold.

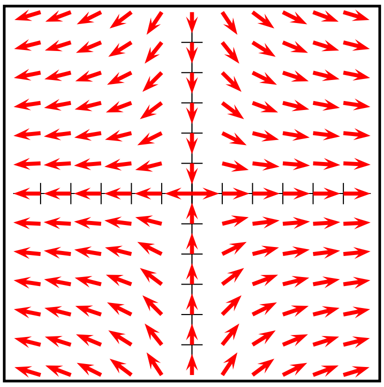


Figure 5: The bias vector field with $\phi(t) = \frac{1}{2}\left(\alpha_1 t_1^2 - \alpha_2 t_2^2\right)$.

In this case, the bias vector field $b(t)$ is centripetal, with all the biases directed towards 0. Another example is if $b(t)$ is the gradient of the function $\phi(t) = \frac{1}{2}\left(\alpha_1 t_1^2 - \alpha_2 t_2^2\right)$, where $t \in \mathbb{R}^2$, $\alpha_1, \alpha_2 \in \mathbb{R}_+$, and t_1 and t_2 are the two dimensions of the type. Hence $b(t) = \big(\alpha_1 t_1, -\alpha_2 t_2\big)$. Then, $\phi(t)$ has a saddle-point at 0 and the bias vector field $b(t)$ is centrifugal on the first dimension and centripetal on the second dimension as illustrated in Figure 5. In this case, we have $Db(t) = \left(\begin{smallmatrix} \alpha_1 & 0 \\ 0 & -\alpha_2 \end{smallmatrix}\right)$, so $Db(t) + I = \left(\begin{smallmatrix} 1+\alpha_1 & 0 \\ 0 & 1-\alpha_2 \end{smallmatrix}\right)$, and the conditions of the proposition are satisfied whenever $\alpha_2 \leq 1$.

Appendix

A 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 $(Y, \succeq)$.[30]

Definition 6 (Single-Peakedness)**.** *Suppose that X is linearly ordered. A function $f : X \to \mathbb{R}$ is single-peaked if $f(x') > f(x)$ implies $f(x'') > f(x)$ for every x'' strictly between x and x'.*

For the next three definitions, we adopt the terminology of Milgrom (2004).

[30]When there is no risk of confusion we use the same notation $\succeq$ for orderings defined on different sets.

Definition 7 (Single Crossing)**.** *A function $f : X \to \mathbb{R}$ is single crossing if for every $x \preceq x'$,*

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

Definition 8 (Increasing Differences)**.** *A function $g : X \times Y \to \mathbb{R}$ 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 non-decreasing.

Definition 9 (Single Crossing Differences)**.** *A function $g : X \times Y \to \mathbb{R}$ 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.

B Proofs

Proof of Lemma 1 (Consistent Extremal Beliefs). Let σ be a fully revealing communication strategy profile. Then each $\sigma_i : T_i \to \Delta(M_i)$ is separating in the sense that for every $t_i \neq t_i'$, the supports of $\sigma_i(t_i)$ and $\sigma_i(t_i')$ are disjoint. Let $\mu_i(t_{-i} \mid m, t_i)$ be the probability that player i puts on t_{-i} when she is of type t_i and the message profile is m. Suppose that (σ, μ) forms a fully revealing equilibrium with extremal beliefs.

Consider a sequence of mixed communication strategy profiles $\{\sigma^k\}_{k=1}^{\infty}$, where $\sigma_i^k(t_i) \in \Delta(M_i(t_i))$ is completely mixed over $M_i(t_i)$, and such that each sequence $\sigma_i^k(t_i)$ converges to $\sigma_i(t_i)$. Let $\mu_i^k(t_{-i} \mid m, t_i)$ be the beliefs computed from σ_i^k by Bayes rule:

$$\mu_i^k(t_{-i} \mid m, t_i) = \frac{\sigma_{-i}^k(m_{-i}|t_{-i})p(t_{-i}|t_i)}{\sum_{s_{-i} \in T_{-i}} \sigma_{-i}^k(m_{-i}|s_{-i})p(s_{-i}|t_i)}, \tag{1}$$

where

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

Consider an off the equilibrium path message profile m that follows a unilateral deviation by player j. Then $m_j \notin \cup_{t_j \in T_j} supp\big(\sigma_j(t_j)\big)$, whereas the messages of each player $i \neq j$ is such that $m_i \in supp\big(\sigma_i(t_i)\big)$ for some t_i.

The strong belief consistency requirement implies that for some sequence σ^k as the one defined above, the associated beliefs μ^k converge to μ. Suppose that i is not the deviator so that $i \neq j$. The extremal belief assumption implies that $\mu_i(t_{-i} \mid m, t_i) = 1$ for some t_{-i}. But then, because the prior has full support, we deduce from (1) that for every $s_{-i} \neq t_{-i}$

$$\lim_{k \to \infty} \frac{\sigma_{-i}^k(m_{-i}|s_{-i})}{\sigma_{-i}^k(m_{-i}|t_{-i})} = 0. \tag{2}$$

Now consider the type profile $s_{-i} = (s_j, t_{-ij})$, where $s_j \neq t_j$. By (2), we must have

$$\lim_{k \to \infty} \frac{\sigma_j^k(m_j|s_j)}{\sigma_j^k(m_j|t_j)} = 0.$$

Note that the expression in the limit does not depend on i or on the messages of players other than j. But then it implies that all non-deviators attribute the off the equilibrium path message m_j to the same type t_j, regardless of the messages sent by m_{-j} sent by players other than j and regardless of their own type. $\square$

Proof of Theorem 1. First we show necessity. By Remark 1, the existence of a fully revealing equilibrium implies (ii). To show that it implies (i), suppose that (i) does not hold. Then there exists a message $m_i \in M_i$ such that $\text{wct}\big(M_i^{-1}(m_i)\big) = \emptyset$. When receiving message m_i from i, the other players with extremal beliefs must assign it to some type in $M_i^{-1}(m_i)$, say s_i. But since $\text{wct}\big(M_i^{-1}(m_i)\big) = \emptyset$, there exists a type $t_i \in M_i^{-1}(m_i)$ such that $t_i \xrightarrow{\mathcal{M}} s_i$. Then player i would deviate from the equilibrium path by sending m_i when she is of type t_i since that allows her to masquerade as s_i.

Next, we show that (i) and (ii) together imply existence of a fully revealing equilibrium with extremal beliefs. By (ii), there exists an evidence base $\mathcal{E}_i$ for $M_i(\cdot)$. Let $e_i : T_i \to \mathcal{E}_i$ be the associated one-to-one mapping such that $t_i \in \text{wct}\big(M_i^{-1}(e_i(t_i))\big)$. Then we contend that, if (i) holds, there exists a fully revealing equilibrium with extremal beliefs in which the communication strategy of player i is a pure strategy given by the mapping $e_i(\cdot)$. 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 $e_i(t_i)$. 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 $s_i \in \text{wct}\big(M_i^{-1}(m_i)\big)$. 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 $m_i = e_i(s_i)$ 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 $e_i(\cdot)$, we have $s_i \in \text{wct}\big(M_i^{-1}(m_i)\big)$, which means that such a deviation cannot be beneficial for i.

To finish the proof, we show that the equilibrium we have constructed satisfies strong consistency of beliefs. The equilibrium strategy is given by the profile $e = \big(e_1, \cdots, e_n\big)$. Let $t_i^*(m_i) \in M_i^{-1}(m_i)$ be the equilibrium belief associated to any message $m_i \notin \mathcal{E}_i$. Then $t_i^*(m_i) \in \text{wct}\big(M_i^{-1}(m_i)\big)$. 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 σ^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 $e_i(t_i)$, probability $1/k$ on every message $m_i \in M_i(t_i) \setminus \mathcal{E}_i$, such that $t_i^*(m_i) = t_i$, and probability $1/k^2$ on every remaining message. Hence type t_i puts more weight on messages for which she is a worst case type $(1/k)$ than on other messages she could send $(1/k^2)$. It is then easy to see that σ^k converges to e as $k \to \infty$.

Now consider the belief μ_i^k associated to the completely mixed strategy profile σ^k for each player i. To check consistency, we need to check that the beliefs μ_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

$$\mu_i^k(t_{-i}|m, t_i) = \frac{\sigma_{-i}^k(m_{-i}|t_{-i})p(t_{-i}|t_i)}{\sum_{s_{-i} \in T_{-i}} \sigma_{-i}^k(m_{-i}|s_{-i})p(s_{-i}|t_i)}, \tag{3}$$

where

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

converges to 1 if $m_j = e_j(t_j)$ 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 $e^{-1}(m_{-i})$ which is indeed the belief that i forms about the other players on the equilibrium path.

Next consider an information set that follows a detectable unilateral deviation. 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 (3). But now we have the following:

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

In the last case, the k^2 comes from the fact that at least one player other than i and j has used a non-detectable deviation (probability $1/k$), and j has used a message which she sends with probability lower than $1/k$. 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 $t_\ell = e_\ell^{-1}(m_\ell)$ 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. □

Proof of Theorem 2 (Interim Sufficient Conditions). For (MON), it is sufficient to note that for $t_i \neq s_i$, $t_i \xrightarrow{\mathcal{M}} s_i$ implies by monotonicity that $t_i \prec s_i$. Hence a cycle in the masquerade relation would also be a cycle for $\succ$ on T_i, which would contradict its linearity. 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):

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

where the two inequalities come from the masquerade relation. Now suppose that there exists a longer cycle $t_i^1 \xrightarrow{\mathcal{M}} \cdots \xrightarrow{\mathcal{M}} t_i^k \xrightarrow{\mathcal{M}} t_i^1$. Because T_i is linearly ordered, the set $\{t_i^1, \cdots, t_i^k\}$ admits a minimal element with respect to $\succeq$. To fix ideas, let t_i^1 be that minimal element. Then we have $v\big(t_i^2|t_i^1\big) - v\big(t_i^1|t_i^1\big) > 0$ and $v\big(t_i^1|t_i^k\big) - v\big(t_i^k|t_i^k\big) > 0$ from the fact that $t_i^1 \xrightarrow{\mathcal{M}} t_i^2$ and $t_i^k \xrightarrow{\mathcal{M}} t_i^1$. Since t_i^1 is a minimal element in $\{t_i^1, \cdots, t_i^k\}$, we have $t_i^1 \prec t_i^k$, and applying (SCD) to the first of these two inequalities yields $v\big(t_i^2|t_i^k\big) - v\big(t_i^1|t_i^k\big) > 0$. Hence, we have

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

This inequality implies that $t_i^2 \xrightarrow{\mathcal{M}} \cdots \xrightarrow{\mathcal{M}} t_i^k \xrightarrow{\mathcal{M}} t_i^2$ 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 $t_i^1 \xrightarrow{\mathcal{M}} \cdots \xrightarrow{\mathcal{M}} t_i^k \xrightarrow{\mathcal{M}} t_i^1$ denote a longer cycle, $k \geq 3$. We adopt the notation that $t_i^{k+1} = t_i^1$. It must be the case that there exists $\ell \notin \{j, j+1\}$ such that $t_i^j \prec t_i^\ell \prec t_i^{j+1}$ or $t_i^{j+1} \prec t_i^\ell \prec t_i^j$. Indeed, otherwise we would have $t_i^1 \prec t_i^2 \prec \cdots \prec t_i^k \prec t_i^1$, a contradiction since $\preceq$ is a linear order on T_i. Therefore, by single-peakedness, $v_i(t_i^{j+1}|t_i^j) > v_i(t_i^j|t_i^j)$ implies 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}

$$t_i^j \xrightarrow{\mathcal{M}} t_i^\ell \xrightarrow{\mathcal{M}} t_i^{\ell+1} \xrightarrow{\mathcal{M}} \cdots \xrightarrow{\mathcal{M}} t_i^{j-1} \xrightarrow{\mathcal{M}} t_i^j,$$

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

Proof of Lemma 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) Let $\Delta(t_i;t_{-i}) = v_i(s_i'|t_i;t_{-i}) - v_i(s_i|t_i;t_{-i})$, for $s_i' \succ s_i$. Then $\Delta(\cdot)$ is non-decreasing in t_i. But then $\Delta(t_i) = E\big(\Delta(t_i;t_{-i})|t_i\big) = E\big(\Delta(t_i;t_{-i})\big)$ by independence, and it is a non-decreasing function of t_i. Therefore $v_i(s_i|t_i)$ satisfies (ID). $\square$

Proof of Theorem 3 (Weak Sequential Equilibria). Pick an evidence base $\mathcal{E}_i$ for each player, and consider the strategy $e_i(\cdot)$ for each player in which i plays according to her evidence base mapping. By definition of an evidence base, this strategy profile is separating. Suppose that all players believe that the message $e_i(t_i)$ is sent by t_i only. Then the beliefs are consistent on the equilibrium path. Now consider a unilateral deviation $m_i \neq e_i(t_i)$ of player i of type t_i. If $m_i = e_i(s_i)$ for some $s_i \neq t_i$, this deviation cannot be beneficial as other players will believe that m_i was sent by type s_i which is a worst case type of $M_i^{-1}(m_i)$. Now suppose that $m_i \notin \mathcal{E}_i$, so m_i is an off-path message. Assume that the beliefs formed by other players after observing m_i puts probability 1 on a type $s_i^*(m_i,t_{-i}) \in \mathrm{wct}(S_i \mid t_{-i})$. This is possible because all other players have sent an equilibrium message which is correctly interpreted as their true type, so all players know t_{-i}. This belief is an extremal belief that is consistent with the evidence contained in m_i. The interim payoff of player i if she sends m_i is therefore given by

$$v_i(m_i|t_i) = E\left(v_i\big(s_i^*(m_i,t_{-i}|t_i;t_{-i})\big) \mid t_i\right) \leq E\left(v_i\big(t_i|t_i;t_{-i}\big) \mid t_i\right) = v_i(s_i|t_i),$$

where the inequality comes from the fact that $s_i^*(m_i,t_{-i})$ is an ex post worst case type. But this shows that m_i is not a profitable deviation and concludes the proof. $\square$

Proof of Proposition 3. 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. We show that this function has increasing differences in (s_i, t_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 satisfies the following first order condition

$$\frac{\partial}{\partial a_i} u_i\big(BR_i(a^*_{-i}(s_i, t_{-i}), t), a^*_{-i}(s_i, t_{-i}), t\big) = 0. \tag{FOC}$$

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

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

and a second time

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

The first term under the summation is non-negative because $u_i(a_i, a_{-i}, t)$ has increasing differences in (a_{-i}, t_i); the second term is also non-negative since the supermodularity of the base game implies that $a^*(\cdot)$ is non-decreasing. □

Proof of Proposition 5. We have for any $s \neq t$, $v(s|t) > v(t|t) \Leftrightarrow g(s) - g(t) > \frac{1}{\lambda(t)}\big(f(t,t) - f(s,t)\big) \Rightarrow g(s) > g(t)$, where the last implication follows from the fact that $\lambda(t) > 0$ and $f(t,t) > f(s,t)$. Therefore the function $g(\cdot)$ is a weak representation for the masquerade relation. □

Proof of Proposition 6. We define the function $\psi(t) = \phi(t) + \frac{1}{2} t' \Omega t$. The function $\psi(\cdot)$ must be convex since $Db(t) + I \geq 0$ implies that the Hessian of $\psi(\cdot)$ satisfies $D^2\psi = \Omega\,(Db(t) + I) \geq 0$. $\psi(\cdot)$ also inherits the continuous differentiability of $\phi(\cdot)$. Then $\nabla\psi(t)$ satisfies the cyclical monotonicity condition of Rockafellar (1972, p. 238). That is, for every finite sequence of distinct types $t(1), \cdots, t(k)$, we have

$$\sum_{\ell=1}^{k} \big(\nabla\psi\big(t(\ell)\big)\big)'\big(t(\ell+1) - t(\ell)\big) \leq 0,$$

with the convention that $t(k+1) = t(1)$. But that implies

$$\sum_{\ell=1}^{k} b\big(t(\ell)\big)'\Omega\big(t(\ell+1) - t(\ell)\big) + \underbrace{\sum_{\ell=1}^{k} t(\ell)'\Omega\big(t(\ell+1) - t(\ell)\big)}_{\mathcal{T}} \leq 0. \tag{4}$$

We can rewrite $\mathcal{T}$ as follows

$$\mathcal{T} = \sum_{\ell=1}^{k} t(\ell+1)'\Omega t(\ell) - \sum_{\ell=1}^{k} t(\ell+1)'\Omega t(\ell+1) = -\sum_{\ell=1}^{k} t(\ell+1)'\Omega\big((t(\ell+1) - t(\ell)\big).$$

Then combining the initial expression of $\mathcal{T}$ and the one we just derived, we can write that

$$\begin{aligned}\mathcal{T} &= \frac{1}{2}\sum_{\ell=1}^{k} t(\ell)'\Omega\big(t(\ell+1) - t(\ell)\big) - \frac{1}{2}\sum_{\ell=1}^{k} t(\ell+1)'\Omega\big((t(\ell+1) - t(\ell)\big) \\ &= -\frac{1}{2}\sum_{\ell=1}^{k} \big(t(\ell+1) - t(\ell)\big)'\Omega\big(t(\ell+1) - t(\ell)\big)\end{aligned}$$

Going back to (4), we now have:

$$\sum_{\ell=1}^{k} b\big(t(\ell)\big)'\Omega\big(t(\ell+1) - t(\ell)\big) - \frac{1}{2}\sum_{\ell=1}^{k} \big(t(\ell+1) - t(\ell)\big)'\Omega\big(t(\ell+1) - t(\ell)\big) \leq 0.$$

But that is exactly

$$\sum_{\ell=1}^{k} \Big(v\big(t(\ell+1)|t(\ell)\big) - v\big(t(\ell)|t(\ell)\big)\Big) \leq 0.$$

And this rules out the possibility that $t(1), \cdots, t(k)$ forms a cycle of the masquerade relation as we would then have for every $\ell = 1, \ldots, k$, $v\big(t(\ell+1)|t(\ell)\big) - v\big(t(\ell)|t(\ell)\big) > 0$. Since the cyclical monotonicity condition must hold for every finite sequence $t(1), \cdots, t(k)$, we have proved that the masquerade relation must be acyclic. $\square$

C Identifying a Worst-Case Type

The following result identifies a worst case type under any condition of Theorem 2 and under (DM).

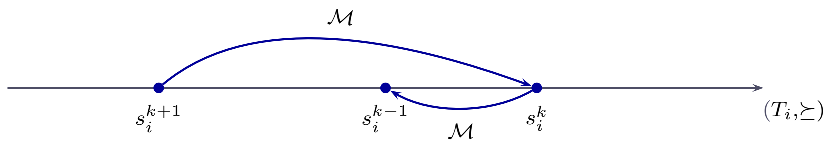
Proposition 7. *Suppose that $v_i(s_i|t_i)$ satisfies (MON), (DM), (SP-NRM) or (SCD). Let S_i be a compact subset of T_i and $s_i^0 = \min S_i$. Then the sequence*

$$s_i^{k+1} = \begin{cases} \inf\{t_i \in S_i \mid v_i(s_i^k|t_i) > v_i(t_i|t_i)\} & \textit{if } \{t_i \in S_i \mid v_i(s_i^k|t_i) > v_i(t_i|t_i)\} \neq \emptyset, \\ s_i^k & \textit{otherwise,} \end{cases}$$

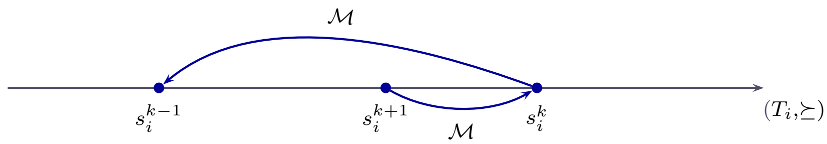
is non-decreasing and converges to some limit $s_i^\infty \in S_i$ such that $s_i^\infty \in \mathrm{wct}(S_i)$.[31]

Proof. The result is obvious under (DM) and (MON) because in that case $s_i^\infty = s_i^0$. Assume (SCD). First notice that if $s_i^{k+1} = s_i^k$ then $\{t_i \in S_i \mid t_i \xrightarrow{\mathcal{M}} s_i^k\} = \emptyset$, and hence $s_i^k \in \mathrm{wct}(S_i)$. To show that the sequence is non-decreasing we show that if $\{t_i \in S_i \mid t_i \xrightarrow{\mathcal{M}} s_i^k\} \neq \emptyset$, then $s_i^{k+1} = \inf\{t_i \in S_i \mid t_i \xrightarrow{\mathcal{M}} s_i^k\} > s_i^k$. By way of contradiction consider the smallest k such that $s_i^{k+1} < s_i^k$ ($k \geq 1$ because $s_i^1 \geq s_i^0$). Then, notice that $s_i^{k+1} = s_i^{k-1}$ is impossible because (SCD) implies (NRM). But $s_i^{k+1} < s_i^k$, $s_i^{k+1} \neq s_i^{k-1}$, is also impossible because in that case we are in one of the two following situations:

(i)



(ii)



In both situations (SCD) implies $s_i^{k+1} \xrightarrow{\mathcal{M}} s_i^{k-1}$, a contradiction with $s_i^k = \inf\{t_i \in S_i \mid t_i \xrightarrow{\mathcal{M}} s_i^{k-1}\}$. A similar proof applies for (SP-NRM). ☐

[31] The same proposition is true by replacing $s_i^0 = \min S_i$ by $s_i^0 = \max S_i$ and inf by sup.

D The Coordination and Influence Games with Constant Biases

The Coordination Game. In the coordination game of Example 4 with constant biases, it is easy to show that every player i's best response takes the form $BR_i(a_{-i};t) = \alpha_{ii}\big(\theta(t) + b_i\big) + \sum_{j\neq i}\alpha_{ij}a_j$, and that equilibrium actions under complete information are given by $a_i^*(t) = \theta(t) + B_i$ for every i, with $B_i \equiv \sum_{j\in N}\gamma_{ij}b_j$, $\gamma_{ij} \equiv \beta_{ij}\,\alpha_{jj} \in (0,1)$ and the β_{ij} are the coefficients of the matrix

$$\beta \equiv \begin{pmatrix} 1 & -\alpha_{12} & \cdots & -\alpha_{1n} \\ -\alpha_{21} & \ddots & \ddots & \vdots \\ \vdots & -\alpha_{ij} & \ddots & \vdots \\ -\alpha_{n1} & \cdots & \cdots & 1 \end{pmatrix}^{-1}.$$

Next, we show that for every player i such that $\sum_{j\neq i}\alpha_{ij}(B_i - B_j) \geq 0$, $v_i(s_i \mid t_i; t_{-i})$ satisfies ex post directional masquerade for the initial order on T_i $v_i(s_i \mid t_i; t_{-i}) > v_i(t_i \mid t_i; t_{-i}) \Rightarrow s_i \succ t_i$. If this is true, the interim masquerading payoff satisfies (DM) by Lemma 3 and for every message m_i, $s_i = \min M_i^{-1}(m_i)$ is a worst case type of $M_i^{-1}(m_i)$. To see that it holds, observe that

$$\begin{aligned} & v_i(s_i \mid t_i; t_{-i}) > v_i(t_i \mid t_i; t_{-i}) \\ \Leftrightarrow\quad & u_i(BR_i(a^*_{-i}(s_i, t_{-i}); t_i, t_{-i}), a^*_{-i}(s_i, t_{-i}); t_i, t_{-i}) > u_i(a_i^*(t_i, t_{-i}), a^*_{-i}(t_i, t_{-i}), t_i, t_{-i}). \end{aligned}$$

To simplify the notations, let $s = (s_i, t_{-i})$ and $t = (t_i, t_{-i})$. Noting that player i's utility when she plays a best response is given by $a_i^2 - \sum_{j\neq i}\alpha_{ij}a_j^2$, the previous inequality becomes:

$$[BR_i(a^*_{-i}(s); t)]^2 - \textstyle\sum_{j\neq i}\alpha_{ij}[a_j^*(s)]^2 > [a_i^*(t)]^2 - \sum_{j\neq i}\alpha_{ij}[a_j^*(t)]^2. \tag{5}$$

We use the form of player i's best response and of equilibrium actions to get $BR_i(a^*_{-i}(s); t) = \alpha_{ii}\big(\theta(t) + b_i\big) + \sum_{j\neq i}\alpha_{ij}\big(\theta(s) + B_j\big)$. From the fact that $B_i = \alpha_{ii}b_i + \sum_{j\neq i}\alpha_{ij}B_j$,[32] we get $BR_i(a^*_{-i}(s); t) = \alpha_{ii}\theta(t) + (1-\alpha_{ii})\theta(s) + B_i$. We insert this expression into Inequality (5) so

[32]We know that $a_i^*(t) = \theta(t) + B_i$. From the expression of $BR_i(a^*_{-i}(s); t)$ that we just calculated, we deduce that $B_i = \alpha_{ii}b_i + \sum_{j\neq i}\alpha_{ij}B_j$ since $a_i^*(t) = BR_i(a^*_{-i}(t); t)$.

that:

$$
\begin{aligned}
& v_i(s_i \mid t_i; t_{-i}) > v_i(t_i \mid t_i; t_{-i}) \\
\Leftrightarrow \quad & \big(\theta(t_i, t_{-i}) - \theta(s_i, t_{-i})\big)\Big[\alpha_{ii}(1-\alpha_{ii})\big(\theta(t_i, t_{-i}) - \theta(s_i, t_{-i})\big) + 2\textstyle\sum_{j\neq i}\alpha_{ij}(B_i - B_j)\Big] < 0.
\end{aligned}
$$

If $\sum_{j\neq i}\alpha_{ij}(B_i - B_j) \geq 0$, then this inequality implies $s_i \succ t_i$ as $\theta(\cdot)$ is non-decreasing.

The same calculation shows that, for every player i such that $\sum_{j\neq i}\alpha_{ij}(B_i - B_j) \leq 0$, $v_i(s_i \mid t_i) > v_i(t_i \mid t_i) \Rightarrow s_i \prec t_i$. For every message m_i of such players, $s_i = \max M_i^{-1}(m_i)$ is a worst case type of $M_i^{-1}(m_i)$. Players for which $\sum_{j\neq i}\alpha_{ij}(B_i - B_j)$ is negative (positive, respectively) are said to have a a relatively low (high, respectively) bias.

The Influence Game. In the influence game (Example 5) with constant biases, equilibrium actions under complete information are given by $a_i^*(t) = \theta(t) + b_i$ for every i. We have:

$$
v_i(s_i \mid t_i; t_{-i}) - v_i(t_i \mid t_i; t_{-i}) = \big(\theta(s_i, t_{-i}) - \theta(t_i, t_{-i})\big)\sum_{j\neq i}\alpha_{ij}\Big[\big(\theta(t_i, t_{-i}) - \theta(s_i, t_{-i})\big) + 2(b_i - b_j)\Big].
$$

Hence, when $b_i > \frac{\sum_{j\neq i}\alpha_{ij}b_j}{\sum_{j\neq i}\alpha_{ij}}$, $v_i(s_i \mid t_i, t_{-i}) - v_i(t_i \mid t_i, t_{-i}) > 0$ implies $s_i \succ t_i$, and when $b_i < \frac{\sum_{j\neq i}\alpha_{ij}b_j}{\sum_{j\neq i}\alpha_{ij}}$, $v_i(s_i \mid t_i, t_{-i}) - v_i(t_i \mid t_i, t_{-i}) > 0$ implies $s_i \prec t_i$. Therefore, in this example, player i is is said to have a relatively high (low, respectively) bias when $b_i - \frac{\sum_{j\neq i}\alpha_{ij}b_j}{\sum_{j\neq i}\alpha_{ij}}$ is positive (negative, respectively).

E Conformity with Multiple Norms: An Illustration

Figure 6 illustrates a masquerade relation for Example 8 with three social stereotypes, $C = \{c^1, c^2, c^3\}$ and a two-dimensional type space. The set of types is $\{c^1, c^2, c^3, t^1, \ldots, t^8\}$. Each of the three areas represents the set of types that are closer to the stereotype that belongs to it than to any other stereotype. A player whose type t is different from t^5 puts a weight $\lambda(t)$ on social stereotypes higher than one, so she always want to masquerade as the social stereotype which is the closest to her. For high values of $\lambda(t)$, a type may even want to masquerade as

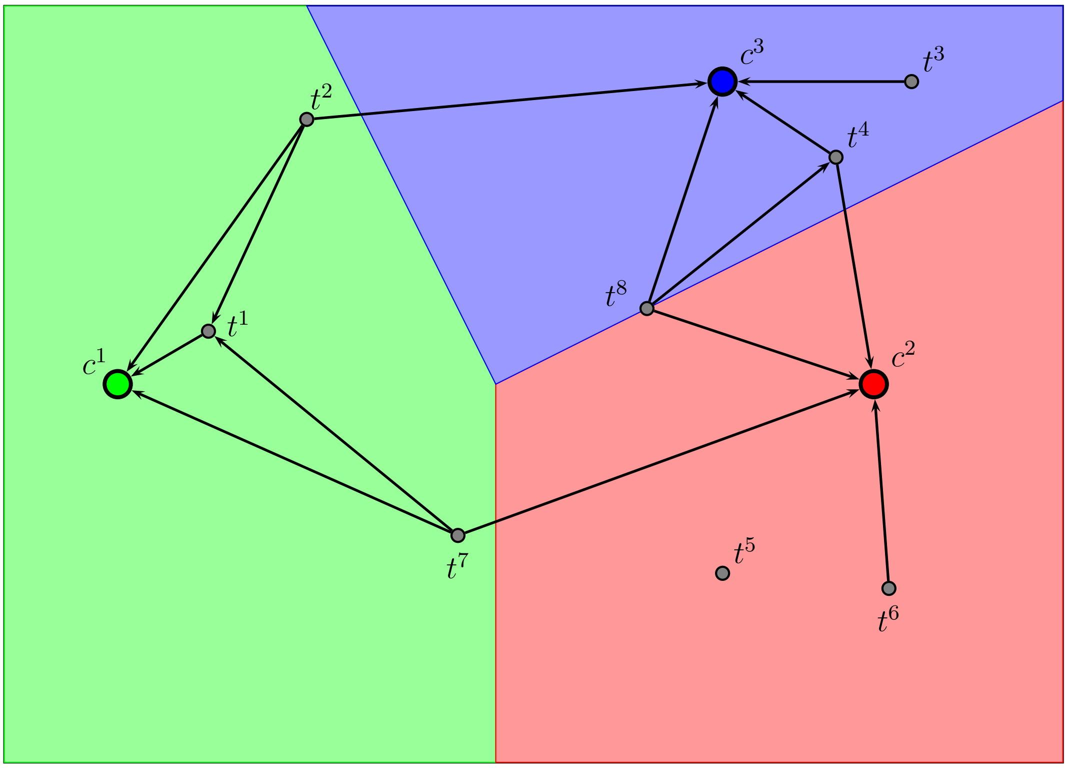


Figure 6: Illustration of Example 8 (conformity with multiple norms) with three social stereotypes, $C = \{c^1, c^2, c^3\}$.

multiple stereotypes when these stereotypes are not too far away (e.g., type t^2, t^4, t^7 and t^8). The acyclic masquerade property is satisfied; for any subset of types, a type that maximizes the minimal distance to a stereotype always constitutes a worst case type.

References

Alcantud, J. and C. Rodríguez-Palmero (1999): "Characterization of the Existence of Semicontinuous Weak Utilities," *Journal of Mathematical Economics*, 32, 503–509.

Alonso, R., W. Dessein, and N. Matouschek (2008): "When Does Coordination Require Centralization?" *American Economic Review*, 98, 145–179.

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. L. Lipman (2012): "Implementation with Partial Provability," *Journal of Economic Theory*, 147, 1689–1724.

Bernheim, B. D. (1994): "A Theory of Conformity," *Journal of Political Economy*, 102, 841–877.

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.

Chakraborty, A. and R. Harbaugh (2010): "Persuasion by Cheap Talk," *American Economic Review*, 100, 2361–82.

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 (2013): "Strategic Information Transmission Networks," *Journal of Economic Theory*, forthcoming.

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 (2010): "Strategic Communication Networks," *Review of Economic Studies*, 77, 1072–1099.

JACKSON, M. AND X. TAN (2013): "Deliberation, Disclosure of Information, and Voting," *Journal of Economic Theory*, 148, 2–30.

KARTIK, N. AND O. TERCIEUX (2012): "Implementation with Evidence," *Theoretical Economics*, 7, 323–355.

KREPS, D. M. AND R. WILSON (1982): "Sequential Equilibria," *Econometrica*, 50, 863–894.

LIZZERI, A. AND L. YARIV (2011): "Sequential Deliberation," mimeo.

MATHIS, J. (2008): "Full Revelation of Information in Sender-Receiver Games of Persuasion," *Journal of Economic Theory*, 143, 571–584.

——— (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 University Press.

MILGROM, P. AND J. ROBERTS (1990): "Rationalizability, learning, and equilibrium in games with strategic complementarities," *Econometrica*, 1255–1277.

MYERSON, R. B. (1991): *Game Theory, Analysis of Conflict*, Harvard University Press.

——— (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.

MYERSON, R. B. AND P. J. RENY (2013): "Sequential Equilibria of Multi-Stage Games with Infinite Sets of Types and Actions," *preliminary draft*.

OKUNO-FUJIWARA, A., M. POSTLEWAITE, AND K. SUZUMURA (1990): "Strategic Information Revelation," *Review of Economic Studies*, 57, 25–47.

QUAH, J. K.-H. AND B. STRULOVICI (2012): "Aggregating the Single Crossing Property," *Econometrica*, 80, 2333–2348.

RANTAKARI, H. (2008): "Governing Adaptation," *The Review of Economic Studies*, 75, 1257.

ROCKAFELLAR, R. (1972): *Convex Analysis*, Princeton University Press, Princeton NY.

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.

Sher, I. and R. Vohra (2011): "Price Discrimination Through Communication," *mimeo.*

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.