Strategic Information Exchange*

Dinah Rosenberg†, Eilon Solan‡ and Nicolas Vieille§

November 20, 2012

Abstract

We analyze a toy class of two-player repeated games with two-sided incomplete information. Two players are facing independent decision problems and each of them holds information that is potentially valuable to the other player. We study to what extent, and how, information can be exchanged at equilibrium. We show that, provided one's initial information is valuable to the other player, equilibria exist at which an arbitrary amount of information is exchanged at an arbitrary high rate. The construction relies on an indefinite reciprocated exchange of information.

Introduction

Discounted repeated games with incomplete information are not quite well-understood yet. In the zero-sum framework of Aumann and Maschler (1995), Mayberry (1967) exhibits an example in which the value depends in a complex way on the discount factor. Cripps and Thomas (2003), and Peski (2008) look at games with *one-sided* information, in which each of the two players knows his own payoff function, and one

*The research of Solan was supported in part by ISF grant #212/09 and by the Google Inter-university center for Electronic Markets and Auctions.

†Economics and Decision Sciences Department, HEC Paris, and GREG-HEC. e-mail: rosenberg@hec.fr.

‡The School of Mathematical Sciences, Tel Aviv University, Tel Aviv 69978, Israel. e-mail: eilons@post.tau.ac.il.

§Economics and Decision Sciences Department, HEC Paris, and GREG-HEC. e-mail: vieille@hec.fr.

of the two is unsure of the payoff function of the other player. Cripps and Thomas (2003) prove that a Folk Theorem type of result holds in the limit where the prior belief converges to the case of complete information. Peski (2008) essentially shows that all equilibria are payoff-equivalent to equilibria that involve finitely many rounds of information revelation. Wiseman (2005) looks at situations of *common* uncertainty. Players share the same information on the underlying state of nature, and refine this information by observing actual choices and payoffs. Hörner and Lovo (2009), Hörner, Lovo and Tomala (2010), and Tomala (2012) provide a characterization of the set of belief-free equilibrium payoffs.

Our main goal in this paper is to analyze a toy class of simple-looking, yet thought-provoking, games with two-sided incomplete information. Consider two agents facing independent repeated decision problems. The two agents are unrelated, except that each has private information that is valuable for the other. Players can communicate only through their actions, which is costly. Can information be exchanged at equilibrium?

As an illustration, consider the following two-player discounted repeated game. Two biased coins C_1 and C_2 are tossed independently, once, at the outset of the game. The parameter of each coin is equal, say, to $\frac{2}{3}$. Each player i has to repeatedly guess the outcome of the coin C_i. A correct guess yields a payoff of 1, while an incorrect guess yields 0. We assume that only past actions are observed along the play, so that there is no room for statistical learning/experimentation.[1] In addition, we assume away cheap talk.

To make the game non-trivial, assume that once coins are tossed, each player gets to learn *only* the outcome of the *other player's* coin. This private information has no 'direct' value, since C_1 and C_2 are independent. In particular, it is an equilibrium for both players to ignore their private information and to repeat throughout the action that matches the most likely outcome of the coins. No information is ever exchanged, and each player's expected payoff is equal to $\frac{2}{3}$. It is readily checked that this is the unique belief-free equilibrium payoff.

Since cheap talk is assumed away, private information can be disclosed to the other player only through one's own actions. That is, disclosing information requires that a player conditions his play on the other coin's outcome, and thus, play both of his actions with some positive probability. But one of these actions will typically be myopically

[1]This assumption is discussed at length in, e.g., Mertens (1986).

suboptimal, in that it will yield a lower expected stage payoff than the other. Therefore, revelation of information is costly, and the cost depends on the player's belief on the outcome of his own coin.[2]

To illustrate this, let us ask whether a player, say player i, might be willing to 'tell' the outcome of C_j to player j at stage 1? The answer is plainly negative: if player j expects to be told the outcome of C_j in stage 1, his unique best response is to play his myopically optimal action at stage 1, and from stage 2 on to play according to the information received from player i in stage 1. But such a strategy does not provide player i with any information on the outcome of C_i. Since revealing information involves playing the myopically suboptimal action with positive probability, player i will refuse to reveal any information in stage 1.

More generally, no player is willing to disclose information, unless he expects to be reciprocated later with valuable information. That is, no player is willing to be the last one in disclosing information.[3] This suggests that equilibria improving upon $(\frac{2}{3}, \frac{2}{3})$ *must* involve an open-ended gradual exchange of valuable information. In Section 3, we show how to construct such equilibria and to implement a whole range of equilibrium payoffs.

The above example is a two-player repeated game with incomplete information on both sides and pure informational externalities – a player's payoff depends on the other player's information, but not otherwise on the other player's action. In this paper we characterize the limit set of sequential equilibrium payoffs for such games. Formally, a player's payoff depends only on his type and on his own action. Both types are drawn independently by nature at the outset of the game; the players receive private signals on the types' realizations, and the types remain fixed throughout the game. Along the play, players repeatedly choose actions, which are publicly disclosed.

Our model is admittedly quite specific, in that it rules out direct strategic interaction. This assumption of pure informational externalities plays a dual role. On one hand, it simplifies the analysis of the model and allows us to study the exchange of information in isolation from other strategic considerations. On the other hand, such games are games in which we would *least* expect that exchange of information might

[2]The variant in which players are allowed to exchange messages at a fixed cost leads to an analysis similar to that of the present paper, as we let the cost of messages vanish.

[3]As is shown later, the fact that a player may be indifferent between both actions does not open up new possibilities.

take place. Our main purpose is to come up with new tools and ideas, and we hope that any positive result in this highly non-generic setup may potentially pave the way for the analysis of other, more economically relevant set-ups.

Our main result is the following. We prove that, provided that the information held by each player is valuable to the other player, the limit set of sequential discounted equilibrium payoffs when players become more and more patient coincides with the set of all feasible payoffs, that are at least equal to the initial, myopic optimal payoffs. In the simple example discussed above, this limit set is thus equal to the set $[\frac{2}{3}, 1] \times [\frac{2}{3}, 1]$. That is, though the problem is a game of informational externalities and the other player's actions do not affect one's own payoffs, and though transmission of information is costly, at equilibrium information can be shared. Moreover, the rate of information exchange can be arbitrarily high relative to the discount rate. Our equilibria share the following features. Players start by reporting truthfully whatever information they received on their own state. This leads to a continuation game in which no player holds private information on his own state. As a result, each player is able to compute how costly it is for the other player to play his suboptimal action, and is therefore able to fine-tune his information disclosure policy, so as to provide the other player with appropriate incentives for disclosing information.

Players next exchange information in an open-ended manner. The analysis presents two main and mostly independent difficulties. One is to design open-ended equilibrium processes, according to which information is exchanged. In our construction, the bulk of information exchange takes place early in the game. Later information disclosure serves only as a means to compensate for previously incurred costs. The second difficulty consists in adjusting this continuation play so as to provide the incentives for truthfully reporting one's information on one's own state.

Starting with Crawford and Sobel (1982), the huge literature on strategic information transmission and on cheap-talk games addresses issues related to ours. The paper that is closest to our work is Aumann and Hart (2003). There, prior to playing a game once, two players, one of whom is informed of the true game to be played, exchange messages during countably many periods. Aumann and Hart (2003) characterize the set of equilibrium payoffs. Following an example of Forges (1990), they show that allowing for an unbounded communication length may increase the set of equilibrium payoffs. Results of Aumann and Hart (2003) were extended by Amitai (1996) to cheap talk with two-sided incomplete information. In particular, Amitai shows that the set of equilibrium payoffs depends on the size of the message space. There are however significant differences with our setup. On the one hand, this literature allows the game to exhibit informational and strategic interaction as well. On the other hand, communication is costless, unlike here.

The paper is organized as follows. Section 1 contains the model. The main result is stated in Section 2, and many related issues are discussed there. Section 3 provides the main insights into the proof, through the detailed discussion of the (generalized) introductory example. Proofs are provided in the Appendix.

1 Model

We study a class of two-player repeated games with incomplete information. At the outset of the game, a state of the world is drawn, and the players receive private information on the realized state. At each stage $n \geq 1$, the two players choose actions from action sets A and B respectively. Actions, and only actions, are then publicly disclosed.

We are motivated by situations in which one's private information is valuable to the other player, and we wish to clarify the extent to which information can be exchanged between the players out of purely strategic reasons. To this end, we make the two assumptions **A1** and **A2** below:

A1 The set of states of the world is a product set, $S \times T$, with elements denoted (s, t). The payoff of player i only depends on his own action and on the i-th component of the state. That is, the payoff function of player 1 is a function $u : S \times A \to \mathbf{R}$, while player 2's payoff is given by a function $v : T \times B \to \mathbf{R}$.

A2 Signal sets are product sets, $L = L_1 \times L_2$ and $M = M_1 \times M_2$. The random triples $(\mathbf{s}, \mathbf{l}_1, \mathbf{m}_2)$ and $(\mathbf{t}, \mathbf{l}_2, \mathbf{m}_2)$ are (stochastically) independent.[4]

Given these assumptions, the game is played as follows. At stage 0, nature chooses independently (s, l_1, m_1) and (t, l_2, m_2) according to prior distributions $p \in \Delta(S \times L_1 \times M_1)$ and $q \in \Delta(T \times L_2 \times M_2)$ respectively; player 1 is told (l_1, l_2) and player 2 is told

[4]Whenever useful to distinguish random variables from their realizations, we use bold letters for the former.

(m_1, m_2). At each stage $n \geq 1$, players choose actions $a_n \in A$ and $b_n \in B$ and the pair (a_n, b_n) is publicly announced.

Assumptions **A1** and **A2** are restrictive. Assumption **A1** ensures that the game has pure informational externalities: player 1 cares about player 2's behavior only to the extent that it conveys information about s. An alternative interpretation of the model is that each player is facing a repeated decision problem.

According to assumption **A2**, these two decision problems are independent. This independence assumption is common in games with two-sided information, see *e.g.* Zamir (1992). Besides allowing for tractability, assumption **A2** implies that behaving myopically is an equilibrium. Indeed, the (interim) belief of player 1 on his state s depends only on l_1. By assumption **A2**, repeating an action that maximizes $\mathbf{E}[u(\mathbf{s}, a) \mid l_1]$ is not informative about $(\mathbf{t}, \mathbf{l_2})$. The belief held by player 2 on t does not change along the play, and it is a best-reply to repeat an action that maximizes $\mathbf{E}[v(\mathbf{t}, b) \mid m_2]$, and vice-versa.

Strategies of the two players will be denoted σ and τ respectively throughout. A behavior strategy of player i maps his private information and the public history of past actions, into a mixed action. Accordingly, behavior strategies are maps $\sigma : L \times H \to \Delta(A)$ and $\tau : M \times H \to \Delta(B)$, where $H = \cup_{n \geq 0}(A \times B)^n$ is the set of finite sequences of action profiles. Every strategy profile (σ, τ) (together with the prior distributions p and q) induces a probability distribution over the set of infinite plays. Expectations under this distribution are denoted by $\mathbf{E}_{p,q,\sigma,\tau}$. The expected payoff of player 1, say, is thus equal to $(1-\delta)\mathbf{E}_{p,q,\sigma,\tau}\left[\sum_{n=1}^{\infty} \delta^{n-1} u(\mathbf{s}, \mathbf{a}_n)\right]$, where $\delta \in [0, 1)$ is the common discount factor.

2 Main results and comments

The main question we ask is whether and how valuable information can be exchanged at equilibrium. Our main result is a characterization of the limit set of sequential equilibrium payoffs, as players become very patient. Loosely put, it reads as follows. Provided that each player holds information that is valuable to the other, a Folk Theorem holds. Information can thus be exchanged, at an arbitrarily high rate, when players are patient enough.

2.1 Valuable information

In equilibrium, a player will play a myopically suboptimal action only if he expects to receive information in return, the marginal value of which offsets the cost incurred when playing the suboptimal action. In particular, a *necessary* condition for improving upon myopic play is that *each* player holds information that is *valuable* to the other. We here formalize this insight.

Consider player 1 at the interim stage, after observing $l = (l_1, l_2)$. In the absence of additional information, his highest payoff is

$$u_\star(l_1) := \max_{a \in A} \mathbf{E}[u(\mathbf{s}, a) \mid l_1]. \tag{1}$$

If player 1 could observe m_1 as well, this highest payoff would instead be equal to $\max_{a \in A} \mathbf{E}[u(\mathbf{s}, a) \mid l_1, m_1]$. Thus, from an interim perspective, the marginal value of the information held by player 2 is $u_{\star\star}(l_1) - u_\star(l_1)$, where

$$u_{\star\star}(l_1) := \mathbf{E}[\max_{a \in A} \mathbf{E}[u(\mathbf{s}, a) \mid \mathbf{l}_1, \mathbf{m_1}] \mid l_1].$$

The payoff $u_\star(l_1)$ is a function of the posterior belief of player 1 given l_1. The inequality $u_{\star\star}(l_1) \geq u_\star(l_1)$ always holds.

Definition 1 *The information held by player 2 is valuable to player 1 at $l_1 \in L_1$ if*

$$u_{\star\star}(l_1) > u_\star(l_1). \tag{2}$$

Condition (2) is an interim requirement. It holds if and only if, after observing l_1, player 1 assigns positive probability to the event that his optimal action would change, if he were to learn m_1.

The *ex ante* value of the information held by player 2 is thus $u_{\star\star} - u_\star$, where $u_{\star\star} := \mathbf{E}[u_{\star\star}(\mathbf{l}_1)]$, and $u_\star := \mathbf{E}[u_\star(\mathbf{l}_1)]$. Plainly, one has $u_{\star\star} \geq u_\star$, and $u_{\star\star} > u_\star$ if and only if the information of player 2 is valuable to player 1 at some $l_1 \in L_1$.

The functions $v_\star(m_2)$, $v_{\star\star}(m_2)$, and their expectations $v_\star$, $v_{\star\star}$ are defined in a symmetric way.

2.2 Main results

We now state our main result.

Theorem 1 *Assume that $u_{\star\star}(l_1) > u_\star(l_1)$ and $v_{\star\star}(m_2) > v_\star(m_2)$ for every $l_1 \in L_1$ and $m_2 \in L_2$. Then, as $\delta \to 1$, the set of sequential equilibrium payoffs converges to the set $[u_\star, u_{\star\star}] \times [v_\star, v_{\star\star}]$.*

More generally, introduce the sets

$$\tilde{L}_1 := \{l_1 \in L_1, u_{\star\star}(l_1) > u_\star(l_1)\} \text{ and } \tilde{M}_2 := \{m_2 \in M_2, v_{\star\star}(m_2) > v_\star(m_2)\}.$$

The set $\tilde{L}_1$ is the set of signals at which the information held by player 2 has a positive value to player 1.

Theorem 2 *Assume that the sets $\tilde{L}_1$ and $\tilde{M}_2$ are both non-empty. Then any payoff in the interior of $[u_\star, u_\star(1 - q(\tilde{M}_2)) + v_{\star\star}q(\tilde{M}_2)] \times [v_\star, v_\star(1 - p(\tilde{L}_1)) + v_{\star\star}p(\tilde{L}_1)]$ is a sequential equilibrium payoff for each δ close enough to 1.*

A few observations are in place. Note first that $(u_\star, v_\star)$ is the minmax point of the repeated game. Indeed, player 1 can guarantee $u_\star$ by repeating an action that maximizes $\mathbf{E}[u(\mathbf{s}, a) \mid l_1]$, ignoring player 2's actions. And player 2 can bring player 1's payoff down to at most $u_\star$ – using any non-revealing strategy (that does not depend on m_1).

Note next that the limit set of feasible and individually rational payoffs is equal to $[u_\star, u_{\star\star}] \times [v_\star, v_{\star\star}]$. Thus, players can achieve payoffs arbitrarily close to $(u_{\star\star}, v_{\star\star})$ (as $\delta \to 1$) by disclosing l_2 and m_2 in the first stages, and by then playing myopically. Thus, Theorem 1 states that a Folk Theorem holds when each player always values the information held by the other player.

Theorem 2 generalizes this finding. As soon as there is a positive *ex ante* probability that a player will value the other player's information, there exist equilibria in which players engage in information exchange. Theorem 2 also shows that, to some extent, the Folk Theorem holds *only* if each player always values the other's information. The logic again is that a player will not be willing to bear the cost of disclosing information if he does not value the information held by the other player. Thus, player 2's payoff cannot exceed $v_\star$ unless $l_1 \in \tilde{L}_1$, and vice-versa. Yet, Theorem 2 does not provide a general characterization of the limit set for the following reason. It might be the case that $l_1 \notin \tilde{L}_1$, and the maximum in (1) is achieved at two different actions. In such a knife-edge case, player 1 can *costlessly* disclose information to player 2.

The extension to an arbitrary number of players is outside of the scope of the paper. With more than two players, there exist cases where some player i holds no private information, yet receives information in equilibrium. The basic intuition is that player i may receive information from some other player j, who is compensated by a third player k.

Throughout we rely on the following leading example, already discussed in the Introduction.

Leading example: All four sets S, T, A and B are equal to $\{0,1\}$. A player's payoff, say player 1's payoff, is 1 if his action matches his state ($s = a$) and 0 otherwise ($s \neq a$). For illustration purposes, we will use the leading example with various signalling structures. Note that $u_\star(l_1) = \max\{p_{l_1}, 1 - p_{l_1}\}$, where $p_{l_1} := p(s = 1 \mid l_1)$ is the probability assigned to state $s = 1$ when observing l_1. The action $a = 1$ (resp., $a = 0$) is myopically optimal if and only if $p_{l_1} \geq \frac{1}{2}$ (resp., $p_{l_1} \leq \frac{1}{2}$).

The introduction discusses the case where each player knows the other player's state, and does not get any private information on his own state. In that specific case, one has $u_{\star\star} = v_{\star\star} = 1$, and $u_\star = \max\{p, 1-p\}$, $v_\star = \max\{q, 1-q\}$, where p is the *ex ante* probability of state $s = 1$. The condition in Theorem 1 thus holds, and the limit set of sequential equilibrium payoffs is $[u_\star, 1] \times [v_\star, 1]$.

2.3 Comments

2.3.1 Correlated states

We assume that the states s and t are independent, and our Folk Theorem is not robust to the introduction of correlation. One obvious reason is that myopic play need no longer be an equilibrium. Indeed, not only myopically optimal actions may change along the play, as players may make useful inferences from the other player's play, but players may have incentives to manipulate these inferences. Example 1 below provides a simple illustration of this.

We view our main result to be the claim that efficient exchange of information is an equilibrium outcome. We believe that this insight is likely to be robust. Indeed, by imposing assumption **A2**, we have here "stacked the deck" against the exchange of information. In a sense, introducing correlation between states may make information exchange easier to implement. This remains speculative; we have no formal statement to offer, and challenging technical difficulties emerge in the correlated case.

Leading Example 1

Consider the following version of the leading example. States and signals are functions of an auxiliary variable ω, the state of the world, which may take four possible values ω_k, $k \in \{1, 2, 3, 4\}$. The probabilities of the different values, and the relationship between ω and states and signals, are as follows:

state of the world	prob.	state of nature s of player 1	state of nature t of player 2	signal l of player 1	signal m of player 2
ω_1	$\frac{1}{6}$	0	0	$\bar{l}$	$\bar{m}$
ω_2	$\frac{1}{3}$	0	0	$\underline{l}$	$\bar{m}$
ω_3	$\frac{1}{3}$	1	0	$\bar{l}$	$\underline{m}$
ω_4	$\frac{1}{6}$	1	1	$\underline{l}$	$\underline{m}$

When behaving in a myopically optimal way, players play as follows. In stage 1, player 2 plays $b = 0$. Indeed, the probability he assigns to state $t = 0$ is either 1 or $\frac{2}{3}$. Meanwhile, player 1 plays $a = 1$ if $l = \bar{l}$, and $a = 0$ if $l = \underline{l}$. Indeed, the probability assigned to state $s = 0$ is $\frac{1}{3}$ in the former case, and $\frac{2}{3}$ in the latter case. In stage 2, player 1 repeats his stage 1 action. On the other hand, player 2 can deduce ω from player 1's stage 1 choice. If player 1 played $a = 1$, player 2 repeats his stage 1 action. If instead player 1 played $a = 0$, player 2 will play either $b = 0$ or $b = 1$, depending on player 2's information. In the former case, players repeat forever their stage 2 action. In the latter, player 1 deduces ω from player 2 stage 2's action, and obtains a payoff of 1 in all later stages. This creates an incentive for player 1 to deviate in stage 1, and to always play $a = 0$, in order to learn the value of ω in stage 3. Thus, it is not an equilibrium to behave myopically, for sufficiently high discount factors. In this example, information exchange is a *consequence* of equilibrium behavior.

2.3.2 Pure strategies

Our construction relies on randomizing as a means of fine-tuning incentives. There are knife-edge cases where information may be exchanged using pure strategies. Yet such cases are highly non-generic. As an illustration, consider the version of the leading example that is described in the introduction.[5] Here, myopic play is the unique pure equilibrium outcome. To see this, assume to the contrary that there is a pure equilibrium and a stage n, in which at least one of the players first discloses the true

[5]Each player knows the other state, no player has private information on his own state.

state. The unique best reply of the other player is to play myopically in every stage, including stage n, and is therefore a pooling strategy. But the unique best reply to a pooling strategy is also pooling, a contradiction.

2.3.3 Finite horizon

Our equilibrium constructions rely on an indefinite, gradual and reciprocated exchange of information, and thus require an infinite horizon. Results for the finite-horizon case do not seem to be as clear-cut.

In many cases, myopic play is the unique equilibrium outcome when the horizon is bounded. To illustrate this claim, we state Proposition 1, which applies to (any version of) the leading example. The proof is in the Appendix.

Proposition 1 *Assume that $p_{l_1} \neq \frac{1}{2}$ and $q_{m_2} \neq \frac{1}{2}$, for every $l_1 \in L_1$ and $m_2 \in M_2$. Then for every $K \in \mathbf{N}$, $(u_\star, v_\star)$ is the unique Nash equilibrium payoff of the K-stage game.*

In this statement, p_{l_1} is the interim probability assigned by player 1 to his own state being state 1. The meaning of q_{m_2} is similar. Proposition 1 may be rephrased as follows. If at the interim stage, players always have a unique optimal action, then myopic play is the unique equilibrium outcome.

Yet, as soon as the setup is enriched, other possibilities arise. As an illustration, consider the binary example, and add to each action set one action, denoted by 2, and which yields payoff $\frac{2}{3}$ irrespective of the state. The optimal action of player 1, as a function of the belief assigned to state $s = 1$, is given by Figure 1 below, and the structure of player 2's best-response is similar.

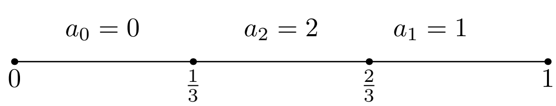


Figure 1: The optimal action of player 1

Assume that the two states are equally likely, that player 1 learns $\mathbf{t}$, and that player 2 learns $\mathbf{s}$. Suppose that in stage 1 player 1 plays $[\frac{1}{3}(a_0), \frac{2}{3}(a_1)]$ if $\mathbf{t} = 0$, and $[\frac{2}{3}(a_0), \frac{1}{3}(a_1)]$ if $\mathbf{t} = 1$, and suppose that player 2 plays in an analogous way. Player 1's belief in stage 2 is either $[\frac{1}{3}(0), \frac{2}{3}(1)]$ or $[\frac{2}{3}(0), \frac{1}{3}(1)]$, depending on player 2's action in stage 1. In the former case,[6] we let player 1 play either a_0 or a_2 depending on $\mathbf{t}$. In

[6]We let player 1 repeat a_2 forever if player 2 played b_2.

the latter case, we let player 1 play either a_1 or a_2, depending on $\mathbf{t}$. Let the behavior of player 2 in stage 2 be analogous. Provided δ is high enough, this strategy pair is an equilibrium, in which players exchange all information in two stages.

2.3.4 Observed payoffs

We assume that payoffs are not observed. This is consistent with most of the literature on repeated games with incomplete information, see e.g. Aumann and Maschler (1995), Hart (1985) or more recently, Hörner and Lovo (2009). Cripps and Thomas (2003) and Peski (2008) look at games with one-sided information, in which each of the players knows his own payoff function, and one of the two is unsure about the payoff function of the other – again, payoffs are not observed along the play.

There are exceptions. In Wiseman (2005), there is incomplete but symmetric information about the state of the world, and payoffs are publicly observed as they are received. The focus is on strategic learning. Strategic experimentation games, also known as bandit games, are another exception. Similar to our model, these are incomplete information games with pure informational externalities. A distinguishing feature is that payoffs are random and observed. Most attention has been paid to the symmetric information setting. Here again, there is no privately held information. We refer to Bergemann and Valimaki (2008) for an overview.

2.3.5 Contribution games

There is an analogy between our games and dynamic games of public good contributions (see, e.g., Admati and Perry (1991), Marx and Matthews (2000)) or dynamic resolution of the hold-up problem (see, e.g., Che and Sakovics (2004) or Pitchford and Snyder (2004)). Information disclosure is costly and has to be reciprocated, just as contributions are. Consequently, in the two setups, the pattern of disclosure/contributions has to be gradual and open-ended. This analogy should not be over-emphasized. Contribution games exhibit strategic externalities, whereas our games do not. The intricacy of our analysis instead stems from incomplete information.

According to Theorem 1, there exist equilibria in which most information is exchanged with an asymptotically negligible delay. This stands in contrast to conclusions for contribution games, see Compte and Jehiel (2004). The driving force behind this difference is the following. Here, the *cost* of disclosing information is the opportunity cost of playing a suboptimal action, while the *amount* of information thus being disclosed depends on how correlated actions are with private information. Thus, the cost and amount are in a sense "orthogonal". By contrast, there is a one-to-one link between the cost of a given contribution, and the amount being contributed.

3 The analysis of the leading example

We illustrate the main ideas of the proof with the leading example: states and actions are binary, and a player gets a payoff of 1 when his action matches his state. In addition, we assume throughout this section that each player knows the other state. The complete proofs of Theorems 1 and 2 are given in the Appendix.

Given a strategy profile, we denote by $\mathbf{p}_n$ the belief of player 1 at stage n – the posterior probability of $s = 1$, given the information of player 1. Thus, we have $u_\star(\mathbf{l}_1) = \max\{\mathbf{p}_1, 1 - \mathbf{p}_1\}$ and $u_{\star\star} = 1$. The belief of player 2 is denoted $\mathbf{q}_n$. For each stage n, we abuse notations and denote by $u_\star(p_n) := \max_{a\in A} u(p_n, a)$ the optimal payoff in stage n, where $u(p_n, a) := \mathbf{E}_{p_n}[u(\cdot, a)]$ is the mixed extension of u. As a supremum of affine functions, $u_\star(\cdot)$ is convex. An action a is (myopically) *optimal* at stage n if $u(p_n, a) = u_\star(p_n)$. It is *suboptimal* otherwise.

We start in Section 3.1 by discussing the case where players have no private information on their own state (that is, the sets L_1 and M_2 are singletons) – the self-ignorant case. Thus, the beliefs p_n and q_n are common knowledge at stage n. This enables a player to compute the opportunity costs incurred by the other player, and to adjust accordingly the amount of information he discloses. We next deal in Section 3.2 with the more general case, where players receive private information on their own state. Building on Section 3.1, we show how to devise equilibria in which players report truthfully this private information.

3.1 No private information on one's own state

Since each player knows the other player's state, and has no private information on his own state, we may identify the distribution p with the *ex ante* probability of the state $s = 1$, and identify q with the probability of state $t = 1$. For concreteness, we assume that $p > \frac{1}{2}$ and $q > \frac{1}{2}$.

3.1.1 A first equilibrium profile

We here construct *one* equilibrum profile that will later serve as a building block. Starting with player 1, players randomize in turn, so long as the randomizing player plays his currently suboptimal action. As soon as this fails to be the case, players stop randomizing and repeat their optimal action. Thus, the play path looks as follows. In a first phase of random duration, player 1 plays suboptimally in odd stages, and optimally in even ones, while player 2 plays optimally and suboptimally in odd and even stages respectively. In a second phase, players repeat their optimal actions. The equilibrium logic is that suboptimal play in any stage is reciprocated by information in the following stage.

For tractability, we design the randomizations in such as way that the evolution of beliefs (and of randomizations) follows a cyclical pattern in Phase 1. The evolution of beliefs on the play path is illustrated in Figure 2. It involves a few parameters, $\bar{x}$, $\underline{x}$, p^* and q^*, that will later be chosen so as to meet equilibrium requirements.

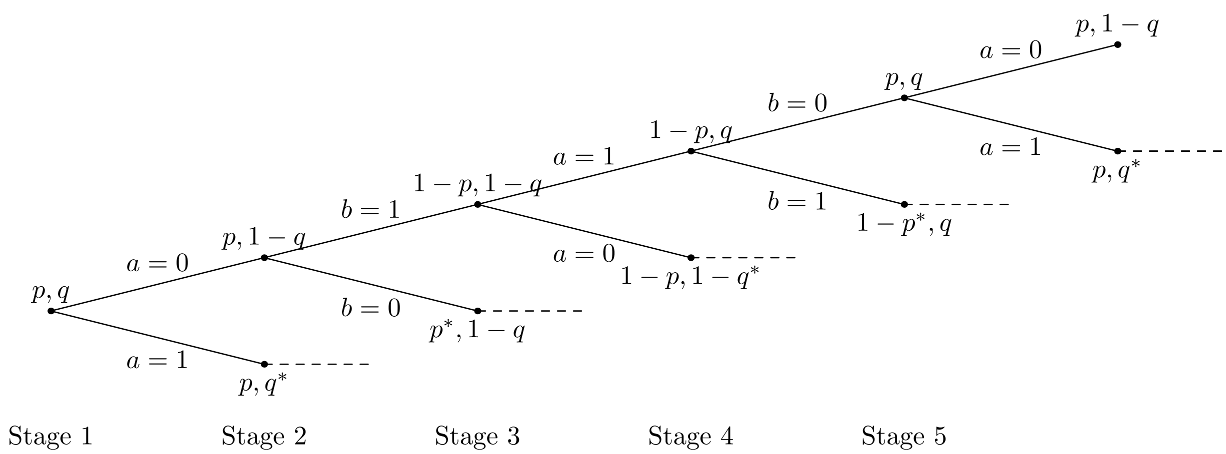


Figure 2: The play as long as both players play suboptimally.

How to read Figure 2? Start with stage 1. Player 2 plays his optimal action (which is $b = 1$ since $q > \frac{1}{2}$), and the belief of player 1 in stage 2 is equal to p, as in stage 1. Meanwhile, player 1 randomizes and plays the suboptimal action $a = 0$ with probability $\bar{x}$ if $\mathbf{t} = 1$, and $\underline{x}$ if $\mathbf{t} = 0$, and player 2 updates his belief to $\mathbf{q}_2 = 1 - q$ or to $\mathbf{q}_2 = q^*$, depending on the action choice of player 1.

Roles are reversed in stage 2. If player 1 played the optimal action $a = 1$ in stage 1, players stop randomizing and repeat their optimal actions $a = 1$ and $b = 1$. If instead player 1 played the suboptimal action $a = 0$, player 2 randomizes and assigns to the suboptimal action, which is now $b = 1$, a probability of $\bar{y}$ if $\mathbf{s} = 1$ and of $\underline{y}$ if $\mathbf{s} = 0$.

Consider next stage 3, and assume that $a = 0$ was played in stage 1. If the optimal action $b = 0$ was played in stage 2, players repeat their optimal actions, which are $a = 1$ and $b = 0$. Otherwise, player 1 reciprocates and plays as in stage 1, with the roles of states/actions being exchanged. That is, following the same logic as in stage 1, the probability assigned to the suboptimal action (which is now $a = 1$) is set to $\bar{x}$ if the true state $\mathbf{t}$ is the one that player 2 currently considers more likely (state 0), and is set to $\underline{x}$ otherwise. This ensures that the belief of player 2 in stage 4 is equal to either q or $1 - q^*$. And so on. Observable deviations are ignored.

We now discuss the parameter values. Two sets of conditions have to hold. To start with, for given $q^* > q$, the values of $\bar{x}$ and $\underline{x}$ should be set to $\bar{x} = \dfrac{1-q}{q} \times \dfrac{q^* - q}{q + q^* - 1} \in (0, 1)$ and $\underline{x} = \dfrac{q}{1-q} \times \dfrac{q^* - q}{q + q^* - 1} \in (0, 1)$, so that Bayesian updating by player 2 leads to the desired beliefs $1 - q$ and q^* respectively.[7] From the perspective of player 2, the suboptimal action is then played with probability $x := q\bar{x} + (1 - q)\underline{x}$. (One can check that x and q^* satisfy $q = x(1 - q) + (1 - q)\underline{x}$, which reflects the martingale property of beliefs.) Similarly, $\bar{y}$ and $\underline{y}$ are given by $\bar{y} = \dfrac{1-p}{p} \times \dfrac{p^* - p}{p + p^* - 1}$ and $\underline{y} = \dfrac{p}{1-p} \times \dfrac{p^* - p}{p + p^* - 1}$.

Next, the parameter values should be adjusted to ensure that a player is indifferent whenever randomizing. For concreteness, consider stage 1. Indifference dictates that the discounted payoff when playing the suboptimal action $a = 0$ be equal to $u_\star(p)$. Similarly, the indifference requirement in stage 3 dictates that the continuation payoff of player 1 following $a = 0$ and $b = 1$ be equal to $u_\star(1 - p)$, which is equal to $u_\star(p)$. The overall payoff when playing $a = 0$ in stage 1 is thus equal to

$$(1 - \delta)(1 - u_\star(p)) + \delta(1 - \delta)u_\star(p) + \delta^2 \left\{ y u_\star(p) + (1 - y)u_\star(p^*) \right\},$$

where the first two terms are the contributions of the first two stages, and the last one is the continuation payoff, which is given by $u_\star(p)$ and $u_\star(p^*)$ with probabilities y and

[7] Note for instance that the likelihood ratio of the two states following $a = 0$ is then given by $\dfrac{q_2}{1 - q_2} = \dfrac{\bar{x}}{\underline{x}} \times \dfrac{q}{1-q} = \dfrac{1-q}{q}$, so that $q_2 = 1 - q$, as claimed.

$1-y$ respectively. Equating this expression with $u_\star(p)=\max\{p,1-p\}$ leads to

$$p^* = p + (2p-1)\frac{1-\delta}{\delta^2+\delta-1}.$$

The same reasoning applied to player 2 yields $q^* = q + (2q-1)\dfrac{1-\delta}{\delta^2+\delta-1}$.

The values of p^* and of q^* must belong to $[0,1]$. This is the case as soon as $\varepsilon_\delta \leq p,q \leq 1-\varepsilon_\delta$, with $\varepsilon_\delta := \dfrac{1-\delta}{\delta^2+\delta-1}$. Not surprisingly, initial beliefs should not be too precise. It is straightforward to check that, as soon as the latter condition is met, these strategies are indeed in equilibrium.

Observe that, while the payoff to player 1 is $u_\star(p)$, the payoff to player 2 is

$$\begin{aligned} f(q) &:= (1-\delta)v_\star(q) + \delta\left(xv_\star(q) + (1-x)v_\star(q^\star)\right) \\ &= v_\star(q) + \frac{1-\delta}{\delta}\left(2v_\star(q)-1\right), \end{aligned}$$

so that $f(q) > v_\star(q)$.

By exchanging the roles of the two players in the construction we obtain an equilibrium with payoff $(f(p), v_\star(q))$.

3.1.2 Further equilibrium payoffs

We here build on the previous section, and explain how to implement many equilibrium payoffs. We proceed in two steps.

We first show how to implement arbitrary payoffs for player 2, while keeping the equilibrium payoff of player 1 equal to $u_\star(p)$. As before, let $p,q\in(0,1)$ be fixed, with $p,q>\frac{1}{2}$. Take any $\underline{q},\bar{q}\in(0,1]$ such that $\underline{q}<q<\bar{q}$, and let δ be high enough so that $p,\underline{q}\in[\varepsilon_\delta, 1-\varepsilon_\delta]$. Consider the following strategy profile. In stage 1, player 2 plays the optimal action, $b=1$, and player 1 randomizes in such a way that the belief of player 2 in stage 2 is equal to $\underline{q}$ and $\bar{q}$ following $a=0$ and $a=1$ respectively. (The existence of such randomizations was first established in Aumann and Maschler (1995), see the so-called splitting lemma.) Following $a=1$, the players switch in stage 2 to indefinite myopic play, with a payoff vector of $(u_\star(p), v_\star(\bar{q}))$. Following $a=0$, the players switch in stage 2 to "the" equilibrium that implements the equilibrium payoff $(f(p), v_\star(\underline{q}))$ in the game with initial beliefs p and $\underline{q}$. Equilibrium properties are derived from the following two insights. Given the first stage behavior, beliefs are either $(p,\underline{q})$ or $(p,\bar{q})$ in stage 2, hence the continuation strategies form indeed an equilibrium in the continuation game. On the other hand, the function $f(\cdot)$ satisfies the identity

$$u_\star(p) = (1-\delta)(1-u_\star(p)) + \delta f(p). \tag{3}$$

Observe that the left-hand side of Eq. (3) is the overall payoff of player 1 when playing $a=1$ in stage 1, while the right-hand side is the overall payoff when playing $a=0$. Hence, by Eq. (3), player 1 is indifferent between his actions in stage 1, as claimed.

The equilibrium payoff of player 2 is given by

$$(1-\delta)v_\star(q) + \delta\left((1-x)v_\star(\underline{q}) + xv_\star(\bar{q})\right),$$

where x is the unconditional probability that player 1 chooses action 1 at stage 1. Since $\varepsilon_\delta \to 0$ as $\delta \to 1$, $\underline{q}$ and $\bar{q}$ may be chosen arbitrarily in $(0,q)$ and $(q,1]$ respectively, and the set of corresponding equilibrium payoffs for player 2 converges to the interval $[v_\star(q),1]$. By symmetry, the limit set of equilibrium payoffs also contains the set $[u_\star(p),1]\times\{v_\star(q)\}$.

We now amend slightly this construction in order to implement equilibrium payoffs that improve upon $(u_\star(p), v_\star(q))$ for *both* players. Take any $\underline{q},\bar{q} \in (0,1)$ such that $\underline{q} < q < \bar{q}$ and let $\underline{\gamma}$ be any payoff in the interval $[u_\star(p),1)$. Given a discount factor δ, we define $\bar{\gamma}_\delta > \underline{\gamma}$ by the identity

$$u_\star(p) + \frac{\delta}{1-\delta}\underline{\gamma} = (1-u_\star(p)) + \frac{\delta}{1-\delta}\bar{\gamma}_\delta. \tag{4}$$

Note that $\bar{\gamma}_\delta \to \underline{\gamma}$ as $\delta \to 1$.

Consider the following strategy profile (σ,τ), which is well-defined for δ high enough. Stage 1 is identical to that in the previous construction, and we consider stage 2. Following $a=1$, the players switch to an equilibrium profile with payoff $(\underline{\gamma}, v_\star(\bar{q}))$ of the game with initial distributions p and $\bar{q}$. Following $a=0$, the players switch to an equilibrium profile with payoff $(\bar{\gamma}_\delta, v_\star(\underline{q}))$ of the game with initial distributions p and $\underline{q}$.

Again, beliefs in stage 2 are either $(p,\underline{q})$ or $(p,\bar{q})$ hence continuation strategies are in equilibrium in the continuation game. On the other hand, the relation (4) between $\underline{\gamma}$ and $\bar{\gamma}_\delta$ ensures that player 1 is indifferent in stage 1. This implies that (σ,τ) is an equilibrium.

Equilibrium payoffs are $(1-\delta)u_\star(p)+\delta\underline{\gamma}$ for player 1, and, as before, $(1-\delta)v_\star(q) + \delta\left((1-x)v_\star(\underline{q}) + xv_\star(\bar{q})\right)$ for player 2. The limit set of such payoffs as $\delta \to 1$ is equal to $[u_\star(p),1]\times[v_\star(q),1]$, as desired.

3.2 Private information on one's own state

We here discuss a more general version of the leading example. We maintain the assumption that each player knows the other state, but assume that each player receives in addition a private signal on his own state. The constructions below are based on the insight that incentives can be provided to report truthfully this private information.

We start with a preliminary digression on the provision of incentives. If $l_1, l'_1 \in L_1$ are two signals following which beliefs of player 1 are the same, then the two signals l_1 and l'_1 are equivalent for all relevant purposes. In particular, merging them into a single signal does not affect the set of equilibrium payoffs. Consequently, we will assume that $p(\cdot \mid l_1) \neq p(\cdot \mid l'_1)$ for every two $l_1 \neq l'_1 \in L_1$ and that a similar condition holds for player 2.[8] Under this assumption, there exists a map $x : L_1 \times S \to (0,1)$ such that for every $l_1 \in L_1$, the map $\lambda \mapsto \mathbf{E}[x(\lambda, \mathbf{s}) \mid l_1]$ is uniquely maximized for $\lambda = l_1$. That is, if player 1 is asked to report a signal, and expects to receive the payoff $x(\lambda, \mathbf{s})$ as a function of his report λ and the state $\mathbf{s}$, then truth-telling is a strictly dominant strategy.

All of our equilibria share a common structure. We describe its main features, focusing on the equilibrium path, that is, at private histories which are consistent with equilibrium play.[9]

Equilibrium play is divided into four phases.

Phase 1 Players report truthfully l_1 and m_2 respectively, by means of encoding them into finite sequences of actions.

Phase 2 is divided into two subphases. Each player sends a message (encoded as a series of actions) in Phase 2.1, and next reacts in Phase 2.2 to the message he received in Phase 2.1. We describe player 1's strategy, the definition for player 2 being symmetric.

In Phase 2.1, player 1 sends a message from the set $T \cup \{\square\}$ (where $\square$ is some extra symbol), which is determined as follows. Player 1 first decides (randomly) whether to disclose *some* information in Phase 2.1, or not. In the latter event, player 1 sends

[8]The notion of equivalent signals has to be adjusted when signalling structures and payoff functions are arbitrary; see the Appendix for more details.

[9]A player reacts to an observable deviation by the other, by repeating the same myopic optimal action afterwards, that is, indefinite punishment. Thus, following one's own observable deviation, it is optimal to play myopically in all following stages. After deviating in an *unobservable* way, a player plays a best reply to the continuation strategy of the other. We make no attempt at describing this best-reply.

the uninformative signal $\square$. In the former event, player 1 draws $\tilde{\mathbf{t}} \in T$, and sends the message $\tilde{\mathbf{t}}$. The variable $\tilde{\mathbf{t}}$ is a noisy, but precise, signal on $\mathbf{t}$. That is, the conditional law of $\tilde{\mathbf{t}}$ given $\mathbf{t}$ has full support and assigns high probability to $\mathbf{t}$.

In Phase 2.2, player 1 plays a deterministic sequence of actions which contains a fraction $x(\tilde{\mathbf{s}}, \lambda)$ of each action $a = \tilde{\mathbf{s}}$, where λ is the report of player 1 in Phase 1.

The durations of Phase 1 and of Phase 2.1 are dictated by the need for encoding: Phase 2.1 lasts two stages, and the length of Phase 1 does not exceed $1 + \log_2 \max\{|L_1|, |M_2|\}$. On the other hand, the length of Phase 2.2 will be adjusted to the discount factor, in such a way that its contribution to the overall discounted payoff is small, yet bounded away from zero.

We make a few remarks. When taking only these first two phases into account, and ruling out deviations in Phase 2.2, it is a dominant strategy to report truthfully in Phase 1, as soon as $\tilde{\mathbf{s}}$ and $\tilde{\mathbf{t}}$ are sufficiently correlated with s and t, and δ is high enough.

We next comment on the design of Phase 2.1. Adding noise in the distribution of $\tilde{\mathbf{t}}$ given $\mathbf{t}$ allows player 1 to retain *private* information on $\mathbf{t}$. This information will be used as a tool to deter deviations in player 2 in Phase 2.2. To be specific, any deviation in Phase 2.2 is observable, and puts an end to the exchange of information, while appropriate behavior in Phase 2.2 is rewarded in later phases by further information.

The main role of the uninformative message $\square$ is to allow this construction to implement arbitrary equilibrium payoffs, and not only those close to the Pareto frontier.

Phase 3 lasts only one stage. Player 1 sends a message $\hat{\mathbf{t}} \in T$. The conditional distribution of $\hat{\mathbf{t}}$ given $\mathbf{t}$ and the message sent by player 1 in Phase 2.1 has full support, and is independent of the initial report of player 2.

We denote by N the first stage of **Phase 4**. Players 1 and 2 compute the beliefs q_N and p_N held by the other player at stage N, assuming that initial reports in **Phase 1** were truthful. At that stage, they switch to an equilibrium of the self-ignorant game with initial distributions p_N and q_N. This continuation equilibrium is chosen to yield a payoff close to the myopic one $(u_\star(p_N), v_\star(p_N))$, and fine-tuned to provide the appropriate incentives for randomizations in the earlier stages. To be precise, fix a public history h_N up to stage N. For each stage $n < N$, excluding all stages from Phase

2.2, player 1 computes the belief q_n held by player 2 at stage[10] n and the opportunity cost incurred by player 2 at that stage, defined as the difference $v_\star(q_n) - v_\star(q_n, b_n)$ between the myopic payoff at stage n, and the payoff of player 2 in that stage. Player 1 adds a "bonus" $b_2(h_N)$ to $v_\star(q_N)$ to compensate for each of these opportunity costs, and to ensure that the discounted sum of payoffs received in Phases 1, 2.1 and 3, and of the continuation payoff $v_*(q_N) + b_2(h_N)$ in Phase 4 is independent of the specific sequence of actions of player 2 along h_N.

We now provide some insights into why equilibrium properties hold. By construction, no strategy that first deviates in Phase 4 can be a profitable deviation, and we accordingly focus on deviations in earlier stages.

Assume that player 1 makes an observable deviation in some stage $n < N$. Assume first that n does not belong to Phase 2.2. Since observable deviations are triggered by myopic play, the payoff of player 1 when deviating[11] is at most $(1-\delta)u_\star(p_n) + \delta\mathbf{E}[u_\star(\mathbf{p}_{n+1}) \mid h_n, l_1]$. This continuation payoff will not exceed the "equilibrium" continuation payoff, thanks to the bonus in Phase 4, the convexity of $u_\star$, and martingale properties of the sequence of beliefs.

If n instead belongs to Phase 2.2, the overall continuation payoff of player 1 when deviating is at most $u_\star(p_n)$. We have little control on the incentivizing payoff $x(\lambda, s)$ in Phase 2.2, and the actual payoff of player 1 in that phase may well be below $u_\star(p_n)$. As explained above, such deviations are deterred by the threat of no further information. For this threat to be effective, Phase 2.2 should not be too long. This puts a (mild) constraint on the length of Phase 2.2, as a function of the residual information held by player 2.

We now discuss unobservable deviations. Assume first that player 1 reports truthfully in Phase 1, but deviates in either Phase 2.1 or Phase 3. Thanks to the bonus added in Phase 4, player 1 is indifferent between his two actions in any of the corresponding stages.

Assume instead that player 1 chooses to misreport his private information in Phase 1. By design, the information received from player 1 in Phases 2.1 and 3 is *independent* of player 1's report. Next, by design also, the bonuses $b_1(h_N)$ and $b_2(h_N)$ in Phase 4 are of the order of $(1-\delta)$, hence continuation strategies in Phase 4 entail little information

[10] Assuming truthful reporting in Phase 1

[11] It might be the case that player 2 discloses information in stage n, so that p_{n+1} need not be equal to p_n.

disclosure. In the Appendix, we will show that the continuation payoff of player 1 in Phase 4 following his untruthful report is then of the order of $u_\star(p_N)+(1-\delta)C$, where p_N is the *actual* belief of player 1 at stage N, and C is some constant independent of δ. Thus, the total gain in Phases 1, 2.1, 3 and 4 from untruthful reporting is at most of the order of $1-\delta$. On the other hand, the *loss* in Phase 2.2 is bounded away from zero, independently of δ. Such deviations therefore fail to be profitable, provided δ is high enough.

The previous discussion highlighted the need for a careful choice of the various parameters. We here add some partial explanations on this issue. Let $\gamma_1, \gamma_2 < 1$ be the desired equilibrium payoffs. First, the level of noise in the conditional distribution of $\tilde{s}$ and $\tilde{t}$ is set low enough, so that truthful reporting in Phase 1 is the unique way to maximize payoffs in Phase 2.2. Second, the weight of Phase 2.2 is set small enough, so that the value of the residual information held by the players after phase 2.1 is valuable enough to deter observable deviations in Phase 2.2. Third, the discount factor should be large enough so that the loss incurred in Phase 2.2 when misreporting exceeds the potential gains in Phase 4. Finally, the randomizations in Phase 3 (and the weight assigned to $\square$ in phase 2.1) are fine-tuned so as to induce the desired payoffs γ_1 and γ_2.

References

[1] Admati, A.R. and M. Perry (1987) Strategic Delay in Bargaining, *The Review of Economic Studies*, **54**, 345–364.

[2] Amitai, M. (1996) Cheap Talk with Incomplete Information on Both Sides, *DP 90, Center for Rationality, The Hebrew University of Jerusalem.*

[3] Aumann, R.J. and S. Hart (1992) *Handbook of Game Theory with Economic Applications, vol. 1,* Handbook of Game Theory with Economic Applications, Elsevier Science Publishers, North-Holland.

[4] Aumann, R.J. and S. Hart (2003) Long Cheap Talk, *Econometrica*, **71**, 1619–1660.

[5] Aumann, R.J. and M.B. Maschler (1966) Game Theoretic Aspects of Gradual Disarmament, Chapter V in Report to the US Arms Central and Disarmament Agency, Washington, DC.

[6] Aumann, R.J. and M.B. Maschler (1995) *Repeated Games with Incomplete Information*, The MIT Press.

[7] D. Bergemann and J. Valimaki. Bandit Problems. In *The New Palgrave Dictionary of Economics*, 2nd edition, editors S. Durlauf and L. Blume. Basingstoke and New York: Palgrave Macmillan Ltd, 2008.

[8] Che, Y.-K. and J. Sakovics (2004) A Dynamic Theory of Holdup, *Econometrica*, **72**, 1063–1103.

[9] Compte, O. and P. Jehiel (2004) Gradualism in Bargaining and Contribution Games, *The Review of Economic Studies*, **71**, 975–1000.

[10] Crawford, V.P. and J. Sobel (1982) Strategic Information Transmission, *Econometrica*, **50**, 1431–1451.

[11] Cripps, M.W. and J.P. Thomas (2003) Some Asymptotic Results in Discounted Repeated Games of One-Sided Incomplete Information, *Mathematics of Operations Research*, **28**, 433–462.

[12] Forges, F. (1990) Equilibria with Communication in a Job Market Example, *the Quarterly Journal of Economics*, **105**, 375–398.

[13] Forges, F. (1992) Repeated Games of Incomplete Information : Non-Zero-Sum, in *Handbook of Game Theory with Economic Applications, vol. 1,* R. Aumann et S. Hart , Handbook of Game Theory with Economic Applications, Elsevier Science Publishers, North-Holland, 155–177.

[14] Hart, S. (1985) Non Zero-Sum Two-Person Repeated Games with Incomplete Information, *Mathematics of Operations Research*, **10**(1), 117–153.

[15] Hörner, J. and S. Lovo (2009) Belief-Free Equilibria in Games with Incomplete Information, *Econometrica*, **77**, 453–487.

[16] Hörner, J. and S. Lovo and T. Tomala (2011) Belief-free Equilibria in Games with Incomplete Information: Characterization and Existence, *Journal of Economic Theory*,**146**, 1770–1795.

[17] Mailath, G. and L. Samuelson (2006) *Repeated Games and Reputations: Long-Run Relationships*, Oxford University Press.

[18] Marx, L.M. and S.A. Matthews (2000) Dynamic Voluntary Contribution to a Public Project, *The Review of Economic Studies*, **67**, 327–358.

[19] Mayberry, J.P. (1967) Discounted Repeated Games with Incomplete Information, in *Report for the U.S. Arms Control and Disarmament Agency, ST-116*, **Chapter V**, 435–461.

[20] Peski, M. (2008) Repeated Games with Incomplete Information on One Side, *Theoretical Economics*, **3**, 29–84.

[21] Pitchford, R. and C. Snyder (2004) A Solution to the Hold-Up Problem involving Gradual Investment, *Journal of Economic Theory*, **14**, 88–103.

[22] Tomala T. (2012) Belief-Free Communication Equilibria in Repeated Games. Preprint.

[23] Wiseman, T. (2005) A Partial Folk Theorem for Games with Unknown Payoff Distributions, *Econometrica*, **73**, 629–645.

[24] Zamir, S. (1992) Repeated Games with Incomplete Information: Zero-Sum, in *Handbook of Game Theory, Vol. 1*, (Chapter 5), Edited by R.J. Aumann and S. Hart, Elsevier Science Publishers B.V., 109–154.

Appendix

We follow as closely as possible the structure of the text. We devote the first two sections to the proof of Theorem 1. We start in Section A with the analysis of self-ignorant games, in which players receive no private information on their own state. In Section B, we build on this analysis to deal with general games. In Section C, we show how Theorem 2 follows from the proof of Theorem 1.

A Self-Ignorant Games

We assume throughout the Appendix that all payoffs are in $[0,1]$. We here deal with self-ignorant games. Equivalently, both sets L_1 and M_2 are singletons. For simplicity, we here write L and M instead of L_2 and M_1. We recall that, being the value function of a decision problem, $u_\star(p)$ is convex in p.

Let initial distributions $p \in \Delta(S \times M)$ and $q \in \Delta(T \times L)$ be given. We denote the corresponding game by $\Gamma(p, q)$. W.l.o.g. we assume that $p(m) > 0$ for each $m \in M$.[12]

We assume that the information of each player i is valuable to player j. Since $\Gamma(p, q)$ is a self-ignorant game, this is equivalent to assuming that there is no action $a \in A$ that is optimal at *all* distributions $p_m := p(\cdot \mid m)$, $m \in M$. Equivalently, one has

$$u_\star = u_\star(p) < u_{\star\star} = \sum_{m \in M} p(m) u_\star(p_m) . \tag{5}$$

By Bayes' rule, the belief $\mathbf{p}_n$ on $(\mathbf{s}, \mathbf{m})$ of player 1 at stage n is the weighted average of $\{p_m \otimes 1_m, m \in M\}$, where the weight of $p_m \otimes 1_m$ is equal to the probability that the signal of player 2 is m, given player 1's information at stage n.[13] Thus, the set of possible values of $\mathbf{p}_n$ is

$$\Delta^\dagger(S \times M) := \mathrm{conv}\{p_m \otimes 1_m, m \in M\}. \tag{6}$$

Because $p(m) > 0$, p lies in the (relative) interior of $\Delta^\dagger(S \times M)$, which we denote by $\overset{\circ}{\Delta}{}^\dagger(S \times M)$. We set $q_l := q(\cdot \mid l)$ for $l \in L$, and the set $\Delta^\dagger(T \times L)$ is defined in a symmetric way.

It is convenient to allow the initial distribution to vary, to account for the fact that beliefs may change along the play. Since all beliefs lie in $\Delta^\dagger(S \times M)$ and $\Delta^\dagger(T \times L)$, we will only consider initial distributions in these sets. We still denote arbitrary such distributions by p and q.

In this section, we prove the two propositions below.

Proposition 2 *Let $p \in \overset{\circ}{\Delta}{}^\dagger(S \times M)$ and $q \in \overset{\circ}{\Delta}{}^\dagger(T \times L)$ be given. There exists $\varepsilon > 0$ and $\bar{\delta} < 1$ such that the following holds. For every $\delta \geq \bar{\delta}$, every payoff vector in $[u_\star(p), u_\star(p) + \varepsilon] \times [v_\star(q), v_\star(q) + \varepsilon]$ is a sequential equilibrium payoff of $\Gamma(p, q)$.*

Given p, q, a discount factor δ and a payoff vector γ that satisfy the conditions of Proposition 2, we will construct an equilibrium profile $(\sigma_{p,q,\gamma}, \tau_{p,q,\gamma})$ in $\Gamma(p, q)$, with payoff γ. Proposition 3 bounds the possible gain of player 1 if player 2 follows $\tau_{p,q,\gamma}$ but has an incorrect belief on p.

[12] And that $q(l) > 0$ for each $l \in L$; to avoid useless repetitions we sometimes state properties for player 1, with the implicit understanding that analogous properties hold for player 2 as well.

[13] This is a way of stating that, as the play proceeds, the belief of player 1 on $\mathbf{m}$ evolves, but the distribution of $\mathbf{s}$ conditional on $\mathbf{m}$ remains equal to $p(\cdot \mid \mathbf{m})$.

Proposition 3 *Let $p \in \overset{\circ}{\Delta}{}^{\dagger}(S \times M)$, $q \in \overset{\circ}{\Delta}{}^{\dagger}(T \times L)$ and $c > 0$ be given. There exists a constant $C > 0$ and $\bar{\delta} < 1$ with the following property. For every discount factor $\delta \geq \bar{\delta}$, every payoff vector γ such that $\|\gamma - (u_\star(p), v_\star(q))\|_\infty \leq (1-\delta)c$, and every $p' \in \Delta^\dagger(S \times M)$, one has*[14]

$$\gamma_\delta^1(p', q, \sigma, \tau_{p,q,\gamma}) \leq u_\star(p') + (1-\delta)C, \text{ for every strategy } \sigma.$$

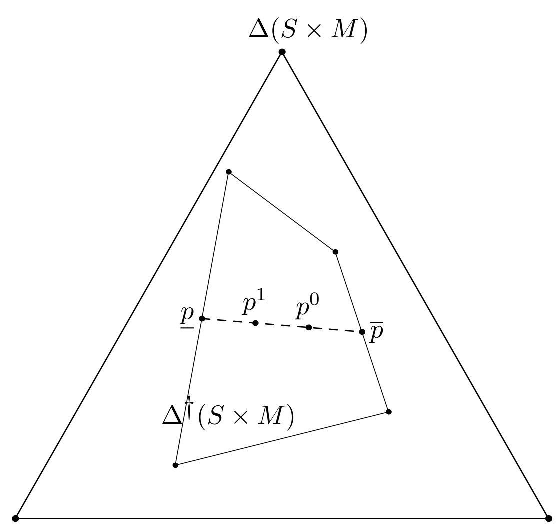
In Propositions 2 and 3 as just stated, ε, C and $\bar{\delta}$ may depend a priori on the choice of (p, q). We will prove in addition that they can be chosen in such a way that the conclusions hold uniformly throughout some neighborhoods of p and q.

A.1 Notations and Preliminaries

We here start with the proof of Propositions 2 and 3, which closely mimics the analysis of the binary case in Section 3.1. We let $[p^0, p^1]$ be any segment in the interior of $\Delta^\dagger(S \times M)$ such that $u_\star$ is not affine on the segment $[p^0, p^1]$. The beliefs p^0 and p^1 take the role of p and $1-p$ in the binary case.

An optimal action a at p^0 is not optimal at p^1 (and vice-versa). Otherwise, a would be optimal throughout the segment $[p^0, p^1]$, and then $u_\star$ would coincide with the affine map $u(\cdot, a)$ on that segment.

For $k = 0, 1$, we let $a^k \in A$ be an optimal action at p^k. We denote by D^1 the straight line spanned by p^0 and p^1 in $\mathbf{R}^{S \times M}$, and we denote by $\underline{p}$ and $\bar{p}$ the endpoints of the segment $D^1 \cap \Delta^\dagger(S \times M)$, with the convention of Figure 3.



[14]Recall that γ_δ^1 is the δ-discounted payoff of player 1 in the whole game.

Figure 3

Let $\pi \in [p_0, \bar{p}]$, and assume that player 1 receives information that changes his belief from p^0 to either p^1 (with probability y) or π (with probability $1-y$). For the martingale property of beliefs to hold, we must have $p^0 = yp^1 + (1-y)\pi$. Assume moreover that from the next stage on player 1 receives his myopically optimal payoff. The marginal gain of player 1 from the information that is revealed to him (relative to his myopically optimal payoff at p^0), is then $h_{p^0}(\pi) = (yu_\star(p^1) + (1-y)u_\star(\pi)) - u_\star(p^0)$. Since $u_\star$ is convex, $h_{p^0}(\pi) \geq 0$ for each π. Since $u_\star$ is not affine on the interval $[p^0, p^1]$, one also has $h_{p^0}(\pi) > 0$ for $\pi \in (p^0, \bar{p}]$, see Figure 4. In addition, h_{p^0} is piecewise affine, and non-decreasing as π moves away from p^0 towards $\bar{p}$.

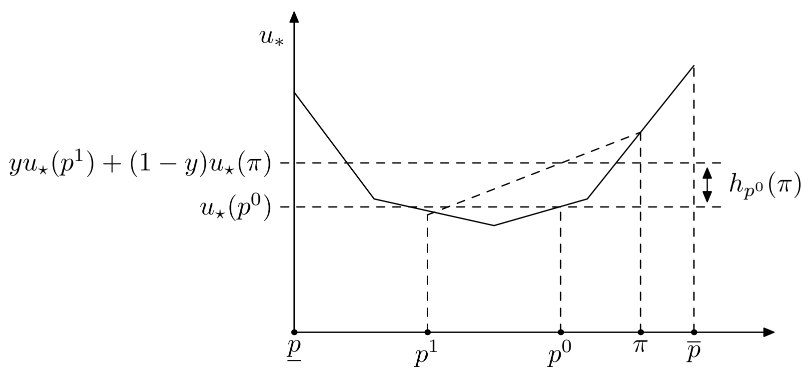


Figure 4

Similarly, define $h_{p^1} : [p^1, \underline{p}] \to \mathbf{R}^+$ by $h_{p^1}(\pi) = (yu_\star(p^0) + (1-y)u_\star(\pi)) - u_\star(p^1)$, where y solves $p^1 = yp^0 + (1-y)\pi$.

We proceed in a symmetric way with player 2. We let $[q^0, q^1]$ be an arbitrary segment in the interior of $\Delta^\dagger(T \times L)$ such that the restriction of $v_\star$ to the segment $[q^0, q^1]$ is not an affine map. We denote by D^2 the straight line in $\mathbf{R}^{T\times L}$ spanned by q^0 and q^1, and by $\underline{q}, \bar{q}$ the endpoints of the segment $D^2 \cap \Delta^\dagger(T \times L)$. Finally, we define $h_{q^0} : [q^0, \bar{q}] \to \mathbf{R}^+$ and $h_{q^1} : [q^1, \underline{q}] \to \mathbf{R}^+$ by adapting the definitions of h_{p^0} and h_{p^1}.

Given a belief $\pi \in \Delta(S \times M)$, and an action $a \in A$, the *cost* of a at π is defined as the loss incurred when playing a instead of the optimal action at π:

$$c(\pi, a) := u_\star(\pi) - u(\pi, a).$$

The cost $c(\pi, b)$, for $\pi \in \Delta(T \times L)$ and $b \in B$, is defined analogously.

The proof of Propositions 2 and 3 relies on Lemmas 1 and 2 below.

Lemma 1 *Let $\delta < 1$ be such that $\frac{1-\delta}{\delta}c(p^i, a^j) < \max h_{p^i}$ and $\frac{1-\delta}{\delta}c(q^i, b^j) < \max h_{q^i}$, for $i, j = 0, 1$. Then the vector $(u_\star(p^0), v_\star(q^0) + \frac{1-\delta}{\delta}c(q^0, b^1))$ is a sequential equilibrium payoff of $\Gamma(p^0, q^0)$.*

Lemma 2 *Let $\varepsilon > 0$ be such that $\varepsilon < \max h_{p^i}$, and $\varepsilon < \max h_{q^j}$ for $i, j = 0, 1$. There is $\bar{\delta} < 1$, such that for every discount factor $\delta \geq \bar{\delta}$, every payoff in $[u_\star(p^0), u_\star(p^0) + \varepsilon] \times [v_\star(q^0), v_\star(q^0) + \varepsilon]$ is a sequential equilibrium payoff of $\Gamma(p^0, q^0)$.*

We will prove that the conclusion holds uniformly for all initial distributions $\tilde{p}^i$, $\tilde{q}^j$ close to p^i and q^j. To be precise, there exist a neighborhood $V(p^i)$ of p^i, and a neighborhood $V(q^j)$ of q^j ($i, j \in \{0, 1\}$) such that, for every $\delta \geq \bar{\delta}$, and every $\tilde{p}^i \in V(p^i), \tilde{q}^j \in V(q^j)$, all vectors in $[u_\star(\tilde{p}^i), u_\star(\tilde{p}^i)+\varepsilon] \times [v_\star(\tilde{q}^j), v_\star(\tilde{q}^j)+\varepsilon]$ are sequential equilibrium payoffs of $\Gamma(\tilde{p}^i, \tilde{q}^j)$.

A.2 Proof of Lemma 1

In the construction of Section 3, the probabilities x and y assigned to suboptimal actions were pinned down by equilibrium requirements. The construction here is slightly more involved because the number of signals may be larger than 2.

We let δ be as stated. Define first $\bar{p}^1 \in [p^0, \bar{p})$ by the condition $h_{p^0}(\bar{p}^1) = \frac{1-\delta}{\delta}c(p^0, a^1)$, and $y^1 \in [0, 1)$ by the equality $p^0 = y^1 p^1 + (1 - y^1)\bar{p}^1$. The value of the information being revealed just offsets the cost to player 1 of playing the suboptimal action a^1 when the belief is p^0. The belief $\bar{p}^1$ plays the role of $p^\star$ in the binary case. For $m \in M$, we set $y^1_m = \frac{p^1(m)}{p^0(m)} y^1$. Because p^0 is in the (relative) interior of $\Delta^\dagger(S \times M)$, one has $p^0(m) > 0$ for each m, and $y^1_m \in (0, 1)$. Observe that $y^1 = \sum_{m \in M} p^0(m) y^1_m$, and that the following Bayesian updating property holds. If player 1's belief is p^0, and if player 2 plays two different actions b and b' with respective probabilities $y^1_{\mathbf{m}}$ and $1 - y^1_{\mathbf{m}}$, then the posterior belief of player 1 is equal to p^1 following b, and it is equal to $\bar{p}^1$ following b'.

Similarly, we let $\bar{p}^0 \in (p^1, \underline{p})$ be defined by $h_{p^1}(\bar{p}^0) = \frac{1-\delta}{\delta}c(p^1, a^0)$, and we set $y^0_m = \frac{p^0(m)}{p^1(m)} y^0$ for $m \in M$, where y^0 solves $p^1 = y^0 p^0 + (1 - y^0)\bar{p}^0$.

We next exchange the roles of the two players, and proceed in a slightly asymmetric way. We let $\bar{q}^1 \in (q^0, \bar{q})$ be defined by $h_{q^0}(\bar{q}^1) = \frac{1-\delta}{\delta}c(q^0, b^1)$, we let x^0 be defined by $q^0 = x^0 q^1 + (1 - x^0)\bar{q}^1$, and we set $x^0_l = \frac{p^1(l)}{p^0(l)} x^0$ for $l \in L$.

We finally define $\bar{q}^0 \in (q^1, \underline{q})$, $x^1 \in (0,1)$, and $x^1_l = \dfrac{q^0(l)}{q^1(l)} x^0$ for $l \in L$ in a similar way.

We are now in a position to define strategies $\sigma_\star$ and $\tau_\star$. As long as players alternate in playing their suboptimal action, player 1 (resp. player 2) randomizes in each *odd* (resp. in each *even*) stage, and beliefs evolve cyclically:

$$p^0, q^0 \to p^0, q^1 \to p^1, q^1 \to p^1, q^0 \to p^0, q^0 \to \cdots$$

Along this cycle, player 1 assigns a probability $x^1_{\mathbf{l}}$ to his suboptimal action, a^1, when player 2's belief is q^0, and a probability $x^0_{\mathbf{l}}$ to his suboptimal action, a^0, when player 2's belief is q^1. Analogous properties hold for player 2. This is summarized in Figure 5 below.

Stage	player 1	player 2	belief	Suboptimal action
1 mod 4	$[x^1_{\mathbf{l}}(a^1), (1-x^1_{\mathbf{l}})(a^0)]$	b^0	p^0, q^0	a^1
2 mod 4	a^0	$[y^1_{\mathbf{m}}(b^0), (1-y^1_{\mathbf{m}})(b^1)]$	p^0, q^1	b^0
3 mod 4	$[x^0_{\mathbf{l}}(a^0), (1-x^0_{\mathbf{l}})(a^1)]$	b^1	p^1, q^1	a^0
0 mod 4	a^1	$[y^0_{\mathbf{m}}(b^1), (1-y^0_{\mathbf{m}})(b^0)]$	p^1, q^0	b^1

Figure 5: the first phase of play: information exchange.

As soon as either player 1 plays his optimal action in some odd stage, or player 2 plays his optimal action in some even stage, the players switch to myopic play forever, as described in columns 3 and 4 in Figure 6. Here and later, $o(p)$ (resp. $o(q)$) stands for an optimal action of player 1 at p (resp. of player 2 at q).

First stage in which myopically optimal action is played	new belief	player 1	player 2
1 mod 4	$p^0, \bar{q}^1$	a^0	$o(\bar{q}^1)$
2 mod 4	$\bar{p}^1, q^1$	$o(\bar{p}^1)$	b^1
3 mod 4	$p^1, \bar{q}^0$	a^1	$o(\bar{q}^0)$
0 mod 4	$\bar{p}^0, q^0$	$o(\bar{p}^0)$	b^0

Figure 6: the second phase of play: myopic play.

We complete the definition of $(\sigma_\star, \tau_\star)$ by specifying actions and beliefs at information sets that are ruled out by $(\sigma_\star, \tau_\star)$. For concreteness, we focus on player 1.

An information set of player 1 contains all histories of the form (l, h), for a fixed signal $l \in L$, and a fixed sequence $h \in H$ of actions. Fix an information set that is reached with probability 0 under $(\sigma_\star, \tau_\star)$. We denote it by $I^1_{l,h}$, with $h \in H$. Write $h = (h', (\bar{a}, \bar{b}))$, so that h' is the longest prefix of h, and assume that $I^1_{l,h'}$ is reached with positive probability.

We distinguish two cases. Assume first that the action $\bar{b}$ has probability zero conditional on h'. That is, player 2 deviates in an observable way at h'. We let the belief of player 1 at $I^1_{l,h}$ be equal to the belief held at $I^1_{l,h'}$ – the deviation by player 2 is interpreted as being non-informative about $\mathbf{m}$. Assume now that $\bar{b}$ is played with positive probability at h'. In that case, the belief of player 1 at $I^1_{l,h}$ can be computed by Bayes' rule, from the belief held at $I^1_{l,h'}$.

In both cases, we let the belief at all subsequent information sets be equal to the belief at $I^1_{l,h}$, and we let $\sigma_\star$ repeat forever any action that is optimal at $I^1_{l,h}$.

Observe that, following any history in $I^1_{l,h}$, under $\tau_\star$ player 2 repeats forever the same action.[15] Indeed, either the sequence h of actions has probability 0, or it has positive probability. In the former case, the claim follows from the definition of $\tau_\star$ at zero probability information sets. In the latter case, this implies that the information set $I^1_{l',h}$ has positive probability, for some $l' \neq l$. Since the support of player 1's mixed actions in the information phase does not depend on his signal, this implies that $I^1_{l,h}$ must belong to the myopic play phase. Using this observation, one can check that beliefs are consistent with the strategy profile $(\sigma_\star, \tau_\star)$. We omit the proof.

Note that the strategy $\sigma_\star$ is sequentially rational at any $I^1_{l,h}$ that is reached with probability 0. Indeed, since the belief of player 1 is the same at $I^1_{l,h}$ and at all subsequent information sets, it is a best reply to repeat any action that is optimal at $I^1_{l,h}$.

Lemma 3 *The profile $(\sigma_\star, \tau_\star)$ is a sequential equilibrium of $\Gamma(p^0, q^0)$, with payoff $(u_\star(p^0), v_\star(q^0) + \frac{1-\delta}{\delta} c(q^0, b^1))$.*

We will use this lemma for various distributions p^0, q^0. To avoid confusion, we will then denote the profile $(\sigma_\star^{p^0,q^0}, \tau_\star^{p^0,q^0})$.

Proof. Each of the strategies $\sigma_\star$ and $\tau_\star$ can be described by an automaton with 8 states: four states that implement the periodic play in Figure 5, and four states that implement the myopic play in Figure 6.

[15] To be precise, player 2 plays the same action at $I^1_{m,h}$ and in all subsequent information sets.

In addition, transitions between (automaton) states are deterministic and depend only on the public history of actions. Hence, player i can always compute the current state of player j's automaton. Moreover, as can be verified inductively, the belief of player i following any public history h of actions only depends on the current state of player j's automaton.

It follows that player i has a best response that can be implemented by an automaton that has the same (or smaller) number of states as the automaton of player j. The dynamic programming principle may be used to identify such a best response. Using this principle, it is routine to verify that $\tau_\star$ is a best response against $\sigma_\star$, and vice versa. Indeed, denoting the 8 states of the automata by $\Omega = \{(1,\texttt{periodic}), (2,\texttt{periodic}), (3,\texttt{periodic}), (0,\texttt{periodic}), (1,\texttt{myopic}), (2,\texttt{myopic}), (3,\texttt{myopic}), (0,\texttt{myopic})\}$, the expected payoff to player 2 starting at any given $\omega \in \Omega$ is:

$$\begin{array}{ll}
\mathcal{V}(1,\texttt{periodic}) = v_\star(q^0) + \frac{1-\delta}{\delta}c(q^0,b^1); & \mathcal{V}(1,\texttt{myopic}) = v_\star(q^0);\\
\mathcal{V}(2,\texttt{periodic}) = v_\star(q^1); & \mathcal{V}(2,\texttt{myopic}) = v_\star(\bar{q}^1);\\
\mathcal{V}(3,\texttt{periodic}) = v_\star(q^*) + \frac{1-\delta}{\delta}c(q^1,b^0); & \mathcal{V}(3,\texttt{myopic}) = v_\star(q^1);\\
\mathcal{V}(0,\texttt{periodic}) = v_\star(q^0); & \mathcal{V}(0,\texttt{myopic}) = v_\star(\bar{q}^0).
\end{array}$$

One may verify that for every $\omega \in \Omega$, V solves

$$\mathcal{V}(\omega) = \max_{b\in B}\left\{(1-\delta)r(\omega,b) + \delta\sum_{\omega'\in\Omega}\mathcal{V}(\omega,b)[\omega']\right\}; \tag{7}$$

here $r(\omega,b)$ stands for the expected payoff of player 2 when playing b in the (automaton) state ω, where the expectation is taken w.r.t. the belief held at state ω. ■

A.3 Proof of Lemma 2

Let a payoff vector $\gamma \in [u_\star(p^0), u_\star(p^0)+\varepsilon] \times [v_\star(q^0), v_\star(q^0)+\varepsilon]$ be given. For δ high enough, we will define a sequential equilibrium profile in $\Gamma(p^0,q^0)$ with payoff γ, using the ideas in Section 3.1.2. We need some preparations.

Define γ_s^1 and γ_o^1 by the equations

$$\begin{aligned}
\gamma^1 &= (1-\delta)u_\star(p^0) + \delta(1-\delta)u_\star(p^0) + \delta^2\gamma_o^1,\\
\gamma^1 &= (1-\delta)u(p^0,a^1) + \delta(1-\delta)u_\star(p^0) + \delta^2\gamma_s^1.
\end{aligned}$$

γ_s^1 (resp. γ_o^1) are the continuation payoffs of player 1 at stage 2, which ensure that the expected payoff of player 1 is γ^1, if player 1 plays the myopically suboptimal (resp. optimal) action at stage 1, and the myopically optimal action[16] at stage 2.

Define γ_o^2 be the equality

$$\gamma^2 = (1-\delta)v_\star(q^0) + \delta\gamma_o^2.$$

Because $\gamma^1 > u_\star(p^0) > u(p^0, a^1)$ and $\gamma^2 > v_\star(q^0)$ it follows that $\gamma_s^1 > \gamma_o^1 \geq u_\star(p^0)$ while $\gamma_o^2 \geq v_\star(q^0)$. For δ high enough, and by definition of ε, one has

$$\gamma_s^1 < u_\star(p^0) + \max h_{p_0} \text{ and } \gamma_o^2 < v_\star(q^0) + \max h_{q^0}.$$

Hence, there exist $p_s, p_o \in [p^0, \bar{p})$, and $q_o \in [q^0, \bar{q})$ such that

$$\begin{aligned} h_{p^0}(p_o) &= \gamma_o^1 - u_\star(p^0), \\ h_{p^0}(p_s) &= \gamma_s^1 - u_\star(p^0), \\ h_{q^0}(q_o) &= \gamma_o^1 - v_\star(q^0). \end{aligned}$$

Mimicking the previous section, we define

- $y_{o,m} = \dfrac{p^1(m)}{p^0(m)} y_o$, for $m \in M$, where y_o solves $p^0 = y_o p^1 + (1-y_o)p_o$.
- $y_{s,m} = \dfrac{p^1(m)}{p^0(m)} y_s$ $(m \in M)$, where y_s solves $p^0 = y_s p^1 + (1-y_s)p_s$.
- $x_l = \dfrac{q^1(m)}{q^0(m)} x$ for $l \in L$, where x solves $q^0 = xq^1 + (1-x)q_o$.

We are now in a position to define a profile as follows (see also Figure 7).

Stage 1: Player 2 plays b^0, while player 1 plays the two actions a^1 and a^0 with probabilities $x_{\mathbf{l}}$ and $1 - x_{\mathbf{l}}$. By Bayesian updating, the belief of player 2 in stage 2 is equal to q^1 following a^1, and it is equal to q_o following a^0 (while the belief of player 1 is still p^0).

Stage 2: Player 2 randomizes. Following a^1, player 2 plays the two actions b^0 and b^1 with probabilities $y_{s,\mathbf{m}}$ and $1 - y_{s,\mathbf{m}}$ respectively. Following a^0, he plays the two

[16] The letters s, o remind that γ_o^1 and γ_s^1 are continuation payoffs following an optimal and a suboptimal action respectively.

actions b^1 and $o(q_o)$ with probabilities $y_{o,\mathbf{m}}$ and $1 - y_{o,\mathbf{m}}$ respectively. Meanwhile, player 1 plays a^0. By Bayesian updating, the belief of player 1 is equal to (i) p^1 following either (a^0, b^1) or (a^1, b^0), (ii) to p_s following $(a^0, o(q_o))$ and (iii) to p_o following (a^1, b^1).

Stage 3 and on: If player 2 played his optimal action in stage 2, players repeat their optimal action. The continuation payoff is then $(u_\star(p^1), v_\star(q^1))$ following (a^1, b^1) and is $(u_\star(p_s^1), v_\star(q_o))$ following $(a^0, o(q_o))$. Assume now that player 2 played b^1 in stage 2, following a^0. Beliefs are then (p_o^1, q_c) and players switch to the equilibrium profile $(\sigma_\star^{p_o^1,q_o}, \tau_\star^{p_o^1,q_o})$ of $\Gamma(p_o^1, q_o)$, with payoff $\left(u_\star(p_o^1), v_\star(q_o) + \frac{1-\delta}{\delta} c(q_o, b^1)\right)$. Finally, assume that player 2 played b^0 in stage 2, following a^1. Beliefs are then (p^1, q^1), and players switch to the profile $(\sigma_\star^{p^1,q^1}, \tau_\star^{p^1,q^1})$.

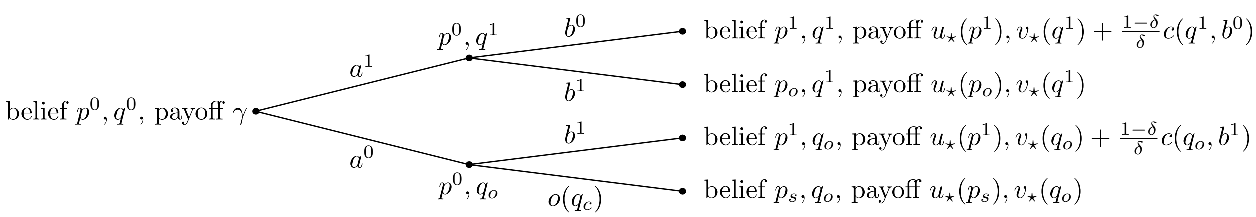


Figure 7: The evolution of beliefs and of continuation payoffs.

Beliefs and actions at information sets that are ruled out by this description are defined as in the proof of Lemma 1. The equations defining γ_s^1, γ_o^1 (resp. γ_o^2) ensure that player 1 is indifferent in stage 1 (resp. player 2 in stage 2) between the two actions that are assigned positive probability. This implies the equilibrium property. Details are standard and omitted.

A.4 Proofs of Propositions 2 and 3

We start with the proof of Proposition 2. The construction we provide here is more complex than needed for Proposition 2. However, it will facilitate the proof of Proposition 3. We let initial distributions p and q be given in the interiors of $\Delta^\dagger(S \times M)$ and $\Delta^\dagger(T \times L)$. Choose a segment $[p^0, p^1]$ included in the interior of $\Delta^\dagger(S \times M)$, such that (i) $u_\star$ is not affine on $[p^0, p^1]$, and (ii) $p \in (p^0, p^1)$.

By (i) and (ii), one has $u_\star(p) < yu_\star(p^0) + (1 - y)u_\star(p^1)$, where y solves $yp^0 + (1 - y)p^1 = p$. Observe also that the quantity $\tilde{y}u_\star(p^0)+(1-\tilde{y})u_\star(\tilde{p}^1)$ (with $\tilde{y}p^0+(1-\tilde{y})\tilde{p}^1 = p$)

is strictly decreasing in the neighborhood of p^1, as $\tilde{p}^1 \in [p^0, p^1]$ moves away from p^1 and towards p^0.

By Lemma 2, there exists $\varepsilon_0 > 0$, $\bar{\delta} < 1$, and neighborhoods $V(p^i)$ and $V(q^j)$ of p^i and q^j ($i, j \in \{0, 1\}$), such that any payoff in $[u_\star(\tilde{p}^i), u_\star(\tilde{p}^i) + \varepsilon_0] \times [v_\star(\tilde{q}^j), v_\star(\tilde{q}^j) + \varepsilon_0]$ is a sequential equilibrium payoff of the game $\Gamma(\tilde{p}^i, \tilde{q}^j)$, as soon as $\delta \geq \bar{\delta}$, $\tilde{p}^i \in V(p^i)$ and $\tilde{q}^j \in V(q^j)$.

We now prove that the conclusion of Proposition 2 holds with $\varepsilon = \varepsilon_0$. Let $\gamma \in [u_\star(p), u_\star(p) + \varepsilon_0] \times [v_\star(q), v_\star(q) + \varepsilon_0]$ be given. We describe an equilibrium profile that implements γ.

One main feature of this profile is the following. As a result of information disclosure by player 2, player 1's belief will move in *one* stage from p to a belief $\tilde{p}^i$ close to either p^0 or p^1. Similarly, player 2's belief will change to a belief $\tilde{q}^j$ close to either q^0 or q^1 in exactly one stage. From that point on, players implement an equilibrium of $\Gamma(\tilde{p}^i, \tilde{q}^j)$ with the appropriate payoff. There is however one minor difference with previously defined equilibria. If $u_\star(p) < \gamma^1 < y u_\star(p^0) + (1 - y) u_\star(p^1)$, then the expected payoff of player 1 if we follow the previous construction will be higher than γ^1, which is the target payoff. There are two ways to overcome this difficulty. One way is to choose in this case p^0 and p^1 which are closer to p, thereby lowering the expected continuation payoff $y u_\star(p^0) + (1 - y) u_\star(p^1)$. A second way, which we adopt here, is to delay information revelation, so that the discounted payoff is lower than $y u_\star(p^0) + (1 - y) u_\star(p^1)$.

Define $N_1 \geq 1$ to be the least integer[17] such that $\gamma^1_c \geq y u_\star(p^0) + (1-y) u_\star(p^1)$, where γ^1_c is defined by $\gamma^1 = (1 - \delta^{N_1}) u_\star(p) + \delta^{N_1} \gamma^1_c$. The inequality $\gamma^1_c \geq y u_\star(p^0) + (1-y) u_\star(p^1)$ ensures that if player 2 starts revealing information at stage N_1, then one can support γ^1_c as a continuation payoff of player 1 at that stage. Define N_2 in a similar way for player 2, and assume w.l.o.g. that $N_1 \leq N_2$. Information is first disclosed at stage N_1. The choice of N_1 implies

$$\gamma^1_c - \left(y u_\star(p^0) + (1 - y) u_\star(p^1) \right) \leq \frac{1 - \delta}{\delta} \left(y u_\star(p^0) + (1 - y) u_\star(p^1) - u_\star(p) \right),$$

provided δ is high enough.

This implies that for δ high enough, there is $\tilde{p}^1 \in V(p^1) \cap [p^0, p^1]$ such that $\gamma^1_c = \tilde{y} u_\star(p^0) + (1 - \tilde{y}) u_\star(\tilde{p}^1)$, and $\tilde{y} p^0 + (1 - \tilde{y}) \tilde{p}^1 = p$.

[17] $N_1 = \infty$ if $\gamma^1 = u_\star(p)$.

We first define a strategy pair (σ, τ) up to stage N_1+1. Player 1 repeats an optimal action $o(p)$ at all stages $1, \dots, N_1$. Player 2 plays $o(q)$ at all stages $1, \dots, N_1 - 1$. In stage N_1, player 2 plays both actions $o(q)$ and $b' \neq o(q)$ with probabilities such that beliefs in stage $N_1 + 1$ are (p^0, q) following b', and $(\tilde{p}^1, q)$ following $o(q)$.

We now define the continuation of (σ, τ) following $o(q)$. Define γ^2_c by the equality $\gamma^2 = (1 - \delta^{N_1})v_\star(q) + \delta^{N_1}\gamma^2_c$. The continuation of (σ, τ) in the other case is defined in an analogous way, except that γ^2_c has to be replaced by $\gamma^2_c + \dfrac{1-\delta}{\delta}c(q, b')$, and the equations that describe equilibrium constraints have to be adjusted.

Let $\tilde{N}_2$ be the least integer (possibly infinite) such that $\tilde{\gamma}^2 \geq xv_\star(q^0)+(1-x)v_\star(q^1)$, where $\tilde{\gamma}^2$ is defined by $\gamma^2_c = (1 - \delta^{\tilde{N}_2})v_\star(q) + \delta^{\tilde{N}_2}\tilde{\gamma}^2$. The choice of $\tilde{N}_2$ implies

$$\tilde{\gamma}^2 - \left(xv_\star(q^0) + (1-x)v_\star(q^1)\right) \leq \frac{1-\delta}{\delta}\left(xu_\star(q^0) + (1-x)u_\star(q^1) - v_\star(q)\right),$$

provided δ is high enough. This implies that for δ high enough, there is $\tilde{q}^1 \in V(q^1) \cap [q^0, q^1]$ such that $\tilde{\gamma}^2 = \tilde{x}v_\star(q^0) + (1-\tilde{x})v_\star(\tilde{q}^1)$, and $\tilde{x}q^0 + (1-\tilde{x})\tilde{q}^1 = q$.

The continuation profile is defined as follows. Player 2 repeats $o(q)$ in all stages $N_1+1, \dots, N_1+\tilde{N}_2$. Player 1 repeats $o(\tilde{p}^1)$ in all stages $N_1, \dots, N_1 + \dots \tilde{N}_2 - 1$. In stage $N_1 + \tilde{N}_2$, player 1 plays both actions $o(\tilde{p}^1)$ and $a \neq o(\tilde{p}^1)$ with probabilities such that the belief of player 2 is equal to $\tilde{q}^1$ following $o(\tilde{p}^1)$, and to q^0 following a.

Following $o(\tilde{p}^1)$, players switch to an equilibrium of the game $\Gamma(\tilde{p}^1, \tilde{q}^1)$ with payoff $(u_\star(\tilde{p}^1), v_\star(\tilde{q}^1))$. Following a, players switch to an equilibrium of the game $\Gamma(\tilde{p}^1, q^0)$ with payoff $(u_\star(\tilde{p}^1) + \frac{1-\delta}{\delta}c(\tilde{p}^1, a), v_\star(q^0))$.

Beliefs and actions off-the-equilibrium-path are defined as in the proof of Lemma 1. The definition of beliefs and continuation payoffs ensures that players are indifferent whenever randomizing, and that the overall payoff is exactly γ.

Observe also that there exists a neighborhood $V(p)$ of p, with the following property. The two beliefs p'^0 and p'^1 associated with $p' \in V(p)$ can be chosen to be continuous in p' and $x'u_\star(p'^0) + (1-x')u_\star(p'^1)$ (with $x'p'^0 + (1-x')p'^1 = p'$) is bounded away from $u_\star(p')$ over $V(p)$. Together with the symmetric property for player 2, this ensures that the robustness result mentioned after Proposition 2 holds. This concludes the proof of Proposition 2.

We next proceed to the proof of Proposition 3. Let $p \in \overset{\circ}{\Delta}{}^{\dagger}(S \times M)$, $q \in \overset{\circ}{\Delta}{}^{\dagger}(T \times L)$, and $c > 0$ be given. Let γ be such that $|\gamma^1 - u_\star(p)| \leq (1-\delta)c$ and $|\gamma^2 - v_\star(q)| \leq (1-\delta)c$.

Let $[p^0, p^1]$ be the segment associated with p in the proof of Proposition 2, and let y solve the equation $p = yp^0 + (1-y)p^1$. Set

$$\eta := \big(yu_\star(p^0) + (1-y)u_\star(p^1)\big) - u_\star(p) > 0,$$

and let N_1 be defined as in the proof of Proposition 2. By construction, one has

$$(1-\delta^{N_1-1})u_\star(p) + \delta^{N_1-1}(u_\star(p)+\eta) < \gamma^1 \leq u_\star(p) + (1-\delta)c,$$

hence $\eta\delta^{N_1-1} \leq (1-\delta)c$. Similarly, one has $\eta\delta^{N_2-1} \leq (1-\delta)c$ (possibly for a lower value of η). In the construction of Proposition 2, players repeat the same action until stage $\min\{N_1, N_2\}$. Therefore, for any $p' \in \Delta^\dagger(S \times M)$ and every strategy σ, one has

$$\begin{aligned}\gamma^1(p', q, \sigma, \tau_{p,q,\gamma}) &\leq (1-\delta^{\min\{N_1,N_2\}})u_\star(p') + \delta^{\min\{N_1,N_2\}} \\ &\leq u_\star(p') + (1-\delta)\frac{\delta c}{\eta}.\end{aligned}$$

The result follows, with $C = c/\eta$.

B General games

We here complete the proof of Theorem 1. We find it more convenient to relabel here L_1 and M_1 as L_S and M_S, and L_2 and M_2 as L_T and M_T. Although slightly more cumbersome, this label is more transparent: the label S (resp. T serves as reminder that the signals l_S and m_S provide information on $\mathbf{s}$.

We start with a few notations and remarks in the spirit of Section A. We let initial distributions $p \in \Delta(S \times L_S \times M_S)$ and $q \in \Delta(T \times L_T \times M_T)$ be given. W.l.o.g., we also assume that $p(l_S) > 0$ and $p(l_T) > 0$ for each $l_S \in L_S$ and $l_T \in L_T$.[18] We assume that the information of each player i is valuable for the other player. This is equivalent to assuming that, for each $l_S \in L_S$, there is no action $a \in A$ that is optimal at all beliefs $p_{l_S,m_S} := p(\cdot \mid l_S, m_S)$, $m_S \in M_S$.

As the play proceeds, player 2 may disclose information relative to $\mathbf{m}_S$, and player 1's belief about $\mathbf{m}_S$ may change. Analogously to the case of self-ignorant games (see Eq. (6)), the belief $\mathbf{p}_n$ of player 1 given $\mathbf{l}_S = l_s$ is always in the set

$$\Delta^\dagger_{l_S}(S \times M_S) = \mathrm{conv}\{p_{l_S,m_S} \otimes 1_{l_S} \otimes 1_{m_S}, m_S \in M_S\}.$$

[18] And we make the symmetric assumption for player 2.

Note that $p(\cdot \mid l_S)$ lies in the relative interior of the set $\Delta^\dagger(S \times M_S)$.

For $m_T \in M_T$, we define $\Delta^\dagger_{m_T}(T \times L_T)$ in a symmetric way. The results of Section A will be applied to the different sets $\Delta^\dagger_{l_S}(S \times M_S)$ and $\Delta^\dagger_{m_T}(T \times L_T)$ of initial distributions.

B.1 Providing Incentives

For simplicity, we focus here on player 1. Analogous properties hold for player 2 as well. We first define an equivalence relation $\sim$ over L_S. As we will see, two signals $\underline{l}_S$ and $\bar{l}_S$ such that $\underline{l}_S \sim \bar{l}_S$ may be merged, and treated as a single signal. Given $l_S \in L_S$, we define a vector $\vec{Z}^{l_S}$ of size $M_S \times A \times A$ by

$$\vec{Z}^{l_S}_{m_S,a,a'} := p(m_S \mid l_S)\left(u(p_{l_S,m_S}, a) - u(p_{l_S,m_S}, a')\right), \text{ for } m_S \in M_S, \text{ and } a, a' \in A.$$

Because the information held by player 2 is valuable for player 1, $\vec{Z}^{l_S} \neq \vec{0}$, for each $l_S \in L_S$.[19]

Definition 2 *Let $\underline{l}_S, \bar{l}_S \in L_S$ be given. The two signals $\underline{l}_S$ and $\bar{l}_S$ are* equivalent*, written $\underline{l}_S \sim \bar{l}_S$, if the two vectors $\vec{Z}^{\underline{l}_S}$ and $\vec{Z}^{\bar{l}_S}$ are positively collinear, that is, if*

$$\vec{Z}^{\bar{l}_S} = \alpha \vec{Z}^{\underline{l}_S}, \text{ for some } \alpha > 0. \tag{8}$$

Plainly, if the two distributions $p(\cdot \mid \underline{l}_S)$ and $p(\cdot \mid \bar{l}_S)$ in $\Delta(S \times M_S)$ coincide, then $\underline{l}_S \sim \bar{l}_S$. However, the converse implication does not hold.

We wish to prove here that one can safely assume that no two signals are equivalent. This is done by proving that any equilibrium of the game in which L_S and M_T are replaced by the set of equivalence classes of signals, is an equilibrium of the original game. Formally, this follows from the fact that, given any strategy of player 2, player 1 has a best-reply which only depends on the equivalence class of l_S, see Lemma 4 below.

Observe that, if $\underline{l}_S \sim \bar{l}_S$ and $\vec{Z}^{\bar{l}_S} = \alpha \vec{Z}^{\underline{l}_S}$, then for every two mixed actions $x, x' \in \Delta(A)$ we have:

$$p(m_S \mid \bar{l}_S)\left(u(p_{\bar{l}_S,m_S}, x) - u\left(p_{\bar{l}_S,m_S}, x'\right)\right) = \alpha p(m_S \mid \underline{l}_S)\left(u(p_{\underline{l}_S,m_S}, x) - u\left(p_{\underline{l}_S,m_S}, x'\right)\right). \tag{9}$$

As a preparation for Lemma 4 below, observe that a strategy σ may be viewed as a collection $(\sigma_{l_S})_{l_S \in L_S}$, with the interpretation that $\sigma_{l_S} : L_T \times H \to \Delta(A)$ is the 'interim' strategy used if $\mathbf{l}_S = l_S$.[20]

[19] Indeed, if $\vec{Z}^{l_S} = \vec{0}$, then *any* action $a \in A$ is optimal at p_{l_S,m_S}, for *each* m_S.

[20] To be formal, $\sigma_{l_S}(l_T, h)$ is defined to be $\sigma(l_S, l_T, h)$.

Lemma 4 *Let τ be any strategy of player 2. Then there exists a best reply σ of player 1 to τ such that $\sigma_{\bar{l}_S} = \sigma_{\underline{l}_S}$ whenever $\bar{l}_S \sim \underline{l}_S$.*

According to Lemma 4, player 1 has a best reply that depends only on the equivalence class of $\mathbf{l}_S$.

Proof. Let a strategy τ of player 2 be fixed throughout. Given $f : H \to \Delta(A)$, and $(l_S, l_T) \in L_S \times L_T$, we denote by $\gamma^1(f, \tau \mid l_S, l_T)$ the interim expected payoff of player 1, when getting $\mathbf{l}_S = l_S$, $\mathbf{l}_T = l_T$, and when playing according to f thereafter. Given $n \geq 1$, we also denote by $g^1_n(f, \tau \mid l_S, l_T)$ the corresponding payoff at stage n.

We let $\bar{l}_S \sim \underline{l}_S$ be any two equivalent signals, so that $\vec{Z}^{\bar{l}_S} = \alpha \vec{Z}^{\underline{l}_S}$ for some $\alpha > 0$. We will prove that, for every two "interim strategies" $f : H \to \Delta(A)$ and $f' : H \to \Delta(A)$, for every $l_T \in L_T$ and every stage $n \geq 1$, one has

$$g^1_n(f, \tau \mid \bar{l}_S, l_T) - g^1_n(f', \tau \mid \bar{l}_S, l_T) = \alpha \left(g^1_n(f, \tau \mid \underline{l}_S, l_T) - g^1_n(f', \tau \mid \underline{l}_S, l_T) \right). \tag{10}$$

Equation (10) will imply that

$$\gamma^1(f, \tau \mid \bar{l}_S, l_T) - \gamma^1(f', \tau \mid \bar{l}_S, l_T) = \alpha \left(\gamma^1(f, \tau \mid \underline{l}_S, l_T) - \gamma^1(f', \tau \mid \underline{l}_S, l_T) \right),$$

from which the result follows. Indeed, if f is better than f' when the signal is $\bar{l}_S$, then it is also the case when the signal is $\underline{l}_S$. Therefore if f is a best response when the signal is $\bar{l}_S$, then it is also a best response when the signal is $\underline{l}_S$.

We let a stage $n \geq 1$ be given. We fix $(l_S, l_T) \in L_S \times L_T$, and we decompose the payoff $g^1_n(f, \tau \mid l_S, l_T)$ as follows. For a given sequence of actions $h \in H_n := (A \times B)^{n-1}$, we denote by $\mathbf{P}_{f,\tau}(h \mid l_S, l_T)$ the probability that h occurs, when $(\mathbf{l}_S, \mathbf{l}_T) = (l_S, l_T)$ and players play according to f and τ. We denote by $\mathbf{P}_{f,\tau}(\cdot \mid h, l_S, l_T) \in \Delta(S \times M_S)$ the belief which is then held by player 1.

With these notations, one has

$$g^1_n(f, \tau \mid l_S, l_T) = \sum_{h \in H_n} \mathbf{P}_{f,\tau}(h \mid l_S, l_T) u \left(\mathbf{P}_{f,\tau}(\cdot \mid h, l_S, l_T)), f(h) \right). \tag{11}$$

The belief of player 1 following h is given by

$$\mathbf{P}_{f,\tau}(s \mid h, l_S, l_T) = \frac{1}{\mathbf{P}_{f,\tau}(h, l_S, l_T)} \sum_{m_S \in M_S} \mathbf{P}_{f,\tau}(s, h, l_S, l_T, m_S), \ s \in S,$$

where $\mathbf{P}_{f,\tau}(h, l_S, l_T) = \mathbf{P}_{f,\tau}(h \mid l_S, l_T)p(l_S)q(l_T)$. Because the state $\mathbf{s}$ and the history of actions until stage n are conditionally independent given $(\mathbf{l}_S, \mathbf{l}_T, \mathbf{m}_S)$, this belief is equal to

$$\mathbf{P}_{f,\tau}(s \mid h, l_S, l_T) = \frac{1}{\mathbf{P}_{f,\tau}(h, l_S, l_T)} \sum_{m_S \in M_S} \mathbf{P}_{f,\tau}(h \mid l_S, l_T, m_S) p_{l_S,m_S}(s) p(l_S, m_S) q(l_T). \tag{12}$$

Plugging (12) into (11), and using the linearity of u, one gets

$$g_n^1(f, \tau \mid l_S, l_T) = \sum_{h \in H_n} \sum_{m_S \in M_S} p(m_S \mid l_S) \mathbf{P}_{f,\tau}(h \mid l_S, l_T, m_S) u(p_{l_S,m_S}, f(h)). \tag{13}$$

Using (13) for both f and f', and because $\sum_{h \in H_n} \mathbf{P}_{\sigma,\tau}(h \mid l_S, l_T, m_S) = \sum_{h' \in H_n} \mathbf{P}_{\sigma',\tau}(h' \mid l_S, l_T, m_S) = 1$, we obtain:

$$\begin{aligned} & g_n^1(f, \tau \mid l_S, l_T) - g_n^1(f', \tau \mid l_S, l_T) \\ &= \sum_{m_S \in M_S} p(m_S \mid l_S) \left(\sum_{h \in H_n} \mathbf{P}_{f,\tau}(h \mid l_S, l_T, m_S) u(p_{l_S,m_S}, f(h)) - \sum_{h' \in H_n} \mathbf{P}_{f',\tau}(h' \mid l_S, l_T, m_S) u(p_{l_S,m_S}, f'(h')) \right) \\ &= \sum_{m_S \in M_S} \sum_{h \in H_n} \sum_{h' \in H_n} \mathbf{P}_{f,\tau}(h \mid l_S, l_T, m_S) \mathbf{P}_{f',\tau}(h' \mid l_S, l_T, m_S) \times p(m_S \mid l_S) \left(u(p_{l_S,m_S}, f(h)) - u(p_{l_S,m_S}, f'(h')) \right). \end{aligned} \tag{14}$$

Because $\underline{l}_S$ and $\bar{l}_S$ are equivalent, Eq. (10) follows by (9), and (14) applied to both $l_S = \underline{l}_S$ and $l_S = \bar{l}_S$. ■

Lemma 4 implies that any equilibrium of the modified game in which player 1 only observes the equivalence class of $\mathbf{l}_S$ is an equilibrium of the original game. Put it differently, the set of equilibrium payoffs of the game in which players do not distinguish between equivalent signals is a subset of the set of equilibrium payoffs of the game we started with. Besides, the values of $u_\star$ and $v_\star$ (resp. of $u_{\star\star}$ and $v_{\star\star}$) are the same for both games.

Therefore, it is sufficient to prove that the conclusion of the Main Theorem holds for the modified game. In particular, we may and will assume from here on that, for every two signals $\bar{l}_S \neq \underline{l}_S$, the vectors $\vec{Z}^{\bar{l}_S}$ and $\vec{Z}^{\underline{l}_S}$ are *not* positively collinear. We also make the symmetric assumption for player 2. A direct consequence of this assumption is Corollary 3 below.

Corollary 3 *Let $\bar{a} \in A$ be arbitrary. For $l_S \in L_S$, define the vector $\vec{Y}^{l_S}$ of size $M_S \times A$ by*

$$\vec{Y}^{l_S}_{m_S,a} := p(m_S \mid l_S) \left(u(p_{l_S,m_S}, a) - u(p_{l_S,m_S}, \bar{a}) \right), \quad m_S \in M_S, a \in A.$$

Then for every two signals $\bar{l}_S \neq \underline{l}_S$, the two vectors $\vec{Y}^{\bar{l}_S}$ and $\vec{Y}^{\underline{l}_S}$ are not positively collinear.

The vector $\vec{Y}^{l_S}$ is equal to the projection of $\vec{Z}^{l_S}$ on a lower-dimensional space. Hence, linear independence of $\vec{Y}^{\bar{l}_S}$ and of $\vec{Y}^{\underline{l}_S}$ does not follow in general from linear independence of $\vec{Z}^{\bar{l}_S}$ and $\vec{Z}^{\underline{l}_S}$, and an *ad hoc* proof is needed.

Proof. We argue by contradiction, and assume that $\vec{Y}^{\bar{l}_S} = \alpha \vec{Y}^{\underline{l}_S}$ for some $\alpha > 0$. Let $m_S \in M_S, a, a' \in A$ be arbitrary. Observe that $\vec{Z}^{l_S}_{m_S,a,a'} = \vec{Y}^{l_S}_{m_S,a} - \vec{Y}^{l_S}_{m_S,a'}$, for $l_S = \bar{l}_S, \underline{l}_S$. Hence $\vec{Z}^{\bar{l}_S} = \alpha \vec{Z}^{\underline{l}_S}$, a contradiction. ■

The next lemma is central to the provision of incentives (phase 2 of the equilibrium play). Given $x : L_S \times M_S \to \Delta(A)$, and for every $l_S, k \in L_S$, we define

$$\mathbf{E}_x[l_S \to k] = \sum_{m_S \in M_S} p(m_S \mid l_S) u(p_{l_S,m_S}, x_{k,m_S}),$$

with the following interpretation. The expression $\mathbf{E}_x[l_S \to k]$ is the expected stage payoff when player 1 gets $\mathbf{l}_S = l_S \in L_S$, 'reports' $k \in L_S$, is told $\mathbf{m}_S$, and plays the mixed action $x_{k,\mathbf{m}_S}$ that depends on player 1's report, and on player 2's signal.[21] According to Lemma 5 below, the map x can be chosen in a way that this expected payoff is highest when reporting truthfully.

Lemma 5 *There exists $x^\star : L_S \times M_S \to \Delta(A)$ such that*

$$\mathbf{E}_{x^\star}[l_S \to k] < \mathbf{E}_{x^\star}[l_S \to l_S], \text{ for every } l_S, k \in L_S, l_S \neq k.$$

Proof. Let $\bar{a} \in A$ be arbitrary and let $x : L_S \times M_S \to \Delta(A)$ be given. For $k \in L_S$, we define a vector $\vec{X}^k$ of size $M_S \times A$ by $\vec{X}^k_{m_S,a} := x_{k,m_S}(a)$, $m_S \in M_S, a \in A$. Observe that $x_{k,m_S}(\bar{a}) = 1 - \sum_{a \neq \bar{a}} x_{k,m_S}(a)$. Hence $\mathbf{E}_x[l_S \to k]$ may be rewritten as

$$\begin{aligned}
\mathbf{E}_x[l_S \to k] &= \sum_{m_S \in M_S} p(m_S \mid l_S) u(p_{l_S,m_S}, \bar{a}) \\
&\quad + \sum_{m_S \in M_S} \sum_{a \in A} p(m_S \mid l_S) x_{k,m_S}(a) \left(u\left(p_{l_S,m_S}, a\right) - u\left(p_{l_S,m_S}, \bar{a}\right)\right) \\
&= \vec{Y}^{l_S} \cdot \vec{X}^k + \sum_{m_S \in M_S} p(m_S \mid l_S) u(p_{l_S,m_S}, \bar{a}).
\end{aligned}$$

[21] We use the different letter k to distinguish between a signal and a report, although both belong to the same set L_S.

Because the second term in the last displayed equation does not depend on k, it is sufficient to construct x such that

$$\vec{Y}^{l_S} \cdot \vec{X}^k < \vec{Y}^{l_S} \cdot \vec{X}^{l_S} \text{ for every } l_S, k \in L_S, l_S \neq k. \tag{15}$$

For $l_S \in L_S$ define

$$\tilde{X}^{l_S} := \frac{1}{\|\vec{Y}^{l_S}\|_2} \vec{Y}^{l_S},$$

and let $l_S \neq k$ be arbitrary in L_S. Then by the Cauchy-Schwartz inequality,

$$\tilde{X}^k \cdot \vec{Y}^{l_S} = \frac{\vec{Y}^k}{\|\vec{Y}^k\|_2} \cdot \vec{Y}^{l_S} < \|\vec{Y}^{l_S}\|_2 = \frac{\vec{Y}^{l_S}}{\|\vec{Y}^{l_S}\|_2} \cdot \vec{Y}^{l_S} = \tilde{X}^{l_S} \cdot \vec{Y}^{l_S},$$

where the strict inequality holds since $\vec{Y}^k$ and $\vec{Y}^{l_S}$ are not positively collinear. Therefore, Eq. 15) holds with $(\tilde{X}^{l_S})_{l_S \in L_S}$. Note that (15) still holds when the same constant is added to all components, and/or when all components are multiplied by the same constant $\phi > 0$. Choose $\beta \in \mathbf{R}$ and $\phi > 0$ such that all components of $\phi \tilde{X}^{l_S} + \beta$ lie in $(0, \frac{1}{|M_S \times A|})$, for all l_S. Because $Y^{l_S}_{m_S, \bar{a}} = 0$, it suffices to set

$$x^\star_{l_S, m_S}(a) = \phi \tilde{X}^{l_S}_{m_S, a} + \beta \text{ for } a \neq \bar{a},$$

and $x^\star_{l_S, m_S}(\bar{a}) = 1 - \sum_{a \neq \bar{a}} x_{l_S, m_S}(a)$. ■

Given $\varepsilon_2 : M_S \to \Delta(M_S)$ and $l_S, k \in L_S$ we define

$$\mathbf{E}_{\varepsilon_2, x^\star}[l_S \to k] = \sum_{m_S, \mu \in M_S} p(m_S \mid l_S) \varepsilon_2(\mu \mid m_S) u(p_{l_S, \mu}, x^\star_{k, \mu}).$$

This is the expected stage payoff of player 1 when (i) player 1 gets $\mathbf{l}_S = l_S$ and 'reports' k, (ii) player 2 draws $\mu \in M_S$ according to $\varepsilon_2(\cdot \mid \mathbf{m}_S)$, and (iii) player 1 plays $x^\star_{k,\mu}$. We here abuse notation and write $p_{l_S,\mu}$ for the belief of player 1, given l_S and μ.[22]

Observe that the expectation $\mathbf{E}_{\varepsilon_2, x^\star}[l_S \to k]$ is continuous w.r.t. ε_2, and that $\mathbf{E}_{\varepsilon_2, x^\star}[l_S \to k]$ is equal to $\mathbf{E}_{x^\star}[l_S \to k]$ when $\varepsilon_2(\cdot \mid m_S)$ assigns probability 1 to m_S, for each m_S. Continuity allows us to pick $\varepsilon_2(\cdot \mid m_S)$ with full support, while keeping the conclusions of Lemma 5.

Corollary 4 *There exists* $\varepsilon_2 : M_S \to \mathring{\Delta}(M_S)$ *such that*

$$\mathbf{E}_{\varepsilon_2, x^\star}[l_S \to k] < \mathbf{E}_{\varepsilon_2, x^\star}[l_S \to l_S], \text{ for every } l_S, k \in L_S, l_S \neq k. \tag{16}$$

[22]Note that, for fixed μ, the belief $p_{l_S,\mu}$ depends on ε_2, although this is not emphasized in the notation.

We fix ε_2 and $x^\star$ for the rest of the paper. Because the distribution $\varepsilon_2(\cdot \mid m_S)$ has full support, the conditional distribution $p_{l_S,\mu}$ lies in the relative interior of $\Delta^\dagger_{l_S}(S \times M_S)$ (for each $\mu \in M_S$). Define ε_1 analogously.

B.2 Equilibrium strategies – Structure

We let a payoff vector $\gamma = (\gamma^1, \gamma^2)$ be given, with $u_\star < \gamma^1 < u_{\star\star}$ and $v_\star < \gamma^2 < v_{\star\star}$. We will construct a sequential equilibrium with payoff γ. We let the discount factor δ be given. In the construction we add one additional message, $\square$, to each player.

Given $x \in \Delta(A)$, and given a number N of stages, we denote by $\vec{a}^N(x) \in A^N$, a sequence of actions of length N that provides the best approximation of the mixed action x in terms of discounted frequencies. That is, $\vec{a}^N(x) = (a_n)_{1 \le n \le N}$ is chosen to minimize $\|x_\delta(\vec{a}^N) - x\|_\infty$, where

$$x_\delta(\vec{a}^N)[a] := \frac{1-\delta}{1-\delta^N} \sum_{n=1}^{N} \delta^{n-1} 1_{\{a_n = a\}},\ a \in A.$$

The sequence $\vec{a}^N(x^\star_{k,\mu_1})$ will be the sequence of actions required from player 1 in phase 2.2, when player 1 reports $k \in L_S$ and player 2 sends the message $\mu_1 \in M_S \cup \{\square\}$. For $\mu_1 = \square$, we let $\vec{a}^N(x^\star_{k,\mu_1})$ be an arbitrary sequence of actions, that does not depend on $k \in L_S$.

Similarly, $\vec{b}^N(y) \in B^N$ is a vector that approximates the mixed action y in terms of discounted frequencies.

We set $K_1 := \max\{|L_S|, |M_T|\}$, and we let $\alpha_1 : L_S \to A^{K_1}$ and $\beta_1 : M_T \to B^{K_1}$ be arbitrary one-to-one maps. Similarly, we set $K_2 := 1 + \max\{|L_T|, |M_S|\}$, and we let $\alpha_2 : L_T \cup \{\square\} \to A^{K_2}$ and $\beta_2 : M_S \cup \{\square\} \to B^{K_2}$ be arbitrary one-to-one maps. The maps α_1 and β_1 are used to encode reports on one's own state into sequences of actions, while the maps α_2 and β_2 are used to encode messages on the other player's state into sequences of actions.

We let $\pi^1 \in \mathring{\Delta}(L_T)$ and $\pi^2 \in \mathring{\Delta}(M_S)$ be arbitrary distributions with full support.

We now proceed to the definition of a strategy profile $(\sigma_\delta, \tau_\delta)$. The definition involves additional parameters θ, ζ, and $\psi^i, \psi^i_\square$ $(i = 1, 2)$, all in $(0, 1)$, which will be chosen later. We first define the profile only at information sets that are not ruled out by the definition of $(\sigma_\delta, \tau_\delta)$ at earlier information sets. The definition of $(\sigma_\delta, \tau_\delta)$ at information sets that are reached with probability zero will be provided after.

Phase 1 It lasts K_1 stages. Player 1 plays the sequence $\alpha_1(\mathbf{l}_S)$ of actions, and player 2 plays the sequence $\beta_1(\mathbf{m}_T)$ of actions.

Phase 2 It is divided into two subphases, **Phase 2.1** and **Phase 2.2**.

Phase 2.1 It lasts K_2 stages. Player 1 first draws a message $\lambda_1 \in L_T \cup \{\square\}$. The probability assigned to $\square$, (resp. to each $l'_T \in L_T$), is equal to $1 - \zeta$ (resp. $\zeta \times \varepsilon_1(l'_T \mid \mathbf{l}_T)$). Symmetrically, player 2 draws a message $\mu_1 \in M_S \cup \{\square\}$. The probability assigned to $\square$, (resp. to each $m'_S \in M_S$) is equal to $1 - \zeta$, (resp. $\zeta \times \varepsilon_2(m'_S \mid \mathbf{m}_S)$).

In that phase, the players play the sequences $\alpha_2(\lambda_1)$ and $\beta_2(\mu_1)$ of actions.

Phase 2.2 It lasts $\nu := \lfloor \frac{\ln(1-\theta)}{\ln \delta} \rfloor$ stages. Player 1 infers μ_1 from the actions played by player 2 in Phase 2.1, and plays the sequence $\vec{a}^{\nu}(x^{\star}_{\mathbf{l}_S,\mu_1})$ of actions. Meanwhile, player 2 infers λ_1 from the actions played by player 1 in Phase 2.1, and plays the sequence $\vec{b}(y^{\star}_{\mathbf{m}_T,\lambda_1})$ of actions.

Phase 3 It lasts K_2 stages. Player 1 draws a message $\lambda_2 \in L_T$. The distribution of λ_2 depends on λ_1. If $\lambda_1 = \square$, the probability assigned to $\mathbf{l}_T$ (resp. to each $l'_T \neq \mathbf{l}_T$), is equal $(1 - \psi^1_{\square}) + \psi^1_{\square} \times \pi^1(\mathbf{l}_T)$ (resp. $\psi^1_{\square} \times \pi^1(l'_T)$). If $\lambda_1 \neq \square$, the probability assigned to λ_2 is equal to $(1 - \psi^1) + \psi^1 \times \pi^1(\lambda_2)$ if $\lambda_2 = \mathbf{l}_T$, and it is equal $\psi^1 \times \pi^1(\lambda_2)$ otherwise. Player 2 draws a message $\mu_2 \in M_S$. The distribution of μ_2 depends on μ_1, and is obtained as for player 1.

In this phase, the players play the sequences $\alpha_2(\lambda_2)$ and $\beta_2(\mu_2)$ of actions.

Phase 4 It contains all remaining stages. We denote by $N = K_1 + 2K_2 + \nu + 1$ its first stage. Let $h = (a_n(h), b_n(h))_{n<N} \in (A \times B)^{N-1}$ be the history of actions up to stage N. Player 2 infers from h the belief $p_n(h)$ held by player 1 in each stage $n < N$ along h. In this computation, the report of player 1 in **Phase 1** is assumed to be truthful. For $n < N$, the belief $q_n(h) \in \Delta(T \times L_T)$ is defined in a symmetric way. The players compute

$$c_1(h) = \delta^{-N} \sum_n (1-\delta)\delta^{n-1} c(p_n(h), a_n(h)) \text{ and } c_2(h) = \delta^{-N} \sum_n (1-\delta)\delta^{n-1} c(q_n(h), b_n(h)),$$

where the sum is taken over all stages n of **Phases 1**, **2.1** and **3**. Players then start playing according to the equilibrium profile of the self-ignorant game $\Gamma(p_n(h), q_n(h))$, with payoff $(u_\star(p_n(h)) + c_1(h), v_\star(q_n(h)) + c_2(h))$.

Some interpretation may be helpful. In Phase 2.1, the message $\Box$ is uninformative[23] and is sent with high probability. In Phase 3, the level noise in the message sent by player 1 depend on player 1's first message, and is either ψ^1 if the first message was informative, or $\psi^1_\Box$ otherwise.

B.3 Equilibrium Strategies – Parameter values

We now fix the parameter values, starting with θ. As $\delta \to 1$, the discounted weight of the $\lfloor \frac{\ln(1-\theta)}{\ln \delta} \rfloor$ stages of Phase 2.2 converges to θ. Thus, θ is a measure of the contribution of the checking phase 2.2 to the total payoff. We choose $\theta \in (0,1)$ to be small enough so that the following set of inequalities is satisfied:

$$(1-\theta)\mathbf{E}[u_\star(p_{l_S,\mathbf{m}_S}) \mid \mathbf{l}_S = l_S, \mu_1 = m_S] \quad > \quad u_\star(p(\cdot \mid \mathbf{l}_S = l_S, \mu_1 = m_S)), \ \forall m_S \in M_S \tag{17}$$

$$(1-\theta)u_{\star\star} \quad > \quad \gamma^1, \tag{18}$$

together with the symmetric conditions for player 2.

By construction, the conditional distribution of $\mathbf{m}_S$ given $(\mathbf{l}_S, \mu_1) = (l_S, m_S)$ is independent of ζ, and only depends on the fixed map ε_2. Since $\varepsilon_2(\cdot \mid m'_S)$ has full support for each m'_S, this conditional distribution has full support. Therefore, the residual information held by player 2 is still valuable to player 1, whatever be $\mu^1 \in \{\Box\} \cup M_S$. In particular, (17) holds with $\theta = 0$, and thus also for $\theta > 0$ small enough. Because $\gamma^1 < u_{\star\star}$, condition (18) is also satisfied for small θ.

Condition (17) ensures that, even if payoffs in phase 2.2 are very low, the weight θ of phase 2.2 is so small, that the residual value of the information held by player 2 can still offset the cost incurred when playing the prescribed sequence in phase 2.2. Condition (17) is designed to make sure that, when in phase 2.2, player 1 will rather play the prescribed sequence of actions, than switch to an optimal action.

Observe that with probability $1-\zeta$, player 1 receives no information prior to phase 3. Therefore, for $\zeta > 0$ small the bulk of information exchange takes place in phase 3. Condition (18) ensures that, even if all information exchange is postponed to phase 3, payoffs as high as γ^1 can be implemented.

Choose $\zeta \in (0,1)$ to be small enough so that the two inequalities

$$(1-\zeta)u_\star + \zeta u_{\star\star} < \gamma^1 < (1-\zeta)(1-\theta)u_{\star\star} \tag{19}$$

[23] Since its probability does not depend on signals.

hold, together with the analogous inequalities for player 2.

In phase 2.2, the (conditional) optimal payoff of player 1 is $u_\star$ if $\mu_1 = \Box$, and does not exceed $u_{\star\star}$ if $\mu_1 \neq \Box$. The first inequality ensures that the probability $1-\zeta$ of *not* disclosing information in phase 2.1 ($\mu_1 = \Box$) is so high that the expectation of the optimal payoff given μ_1 does not exceed γ^1. That is, additional information must be disclosed in phase 3 in order to implement γ. This inequality, together with (18), will allow us to adjust other parameter values in a way that the overall payoff is γ. The second inequality in (19) does not play a critical role.

We now choose the value of $\psi^2 \in (0,1)$ small enough so that, for every $l_S \in L_S, m_S \in M_S$,

$$(1-\theta)\mathbf{E}[u_\star(p(\cdot \mid \mathbf{l}_S, \mu_1, \mu_2)) \mid \mathbf{l}_S = l_S, \mu_1 = m_S] > u_\star(p(\cdot \mid \mathbf{l}_S = l_S, \mu_1 = m_S)). \tag{20}$$

In this expression, $p(\cdot \mid \mathbf{l}_S, \mu_1, \mu_2))$ is the belief held by player 1 at the beginning of phase 4 after having received the two messages μ_1, μ_2 of player 2. The left-hand side of (20) is continuous w.r.t. ψ^2. For $\psi^2 = 0$, μ_2 is equal to $\mathbf{m}_S$ with probability 1, and (20) therefore holds by (17). Hence (20) holds for $\psi^2 > 0$ small enough.

Observe that all parameters values ζ, θ, ψ^2 are independent of the discount factor. The last parameter, $\psi^2_\Box$ is chosen such that the expected payoff of player 1 is γ^1. We first argue that for a given $\psi^2_\Box$, the limit discounted payoff of player 1, as $\delta \to 1$, is equal to[24]

$$\theta\mathbf{E}[u(p(\cdot \mid \mathbf{l}_S, \mu_1), x^\star_{\mathbf{l}_S,\mu_1})] + (1-\theta)\mathbf{E}[u_\star(\mathbf{p}_N)]. \tag{21}$$

Here is why. The contribution of Phases 1, 2.1 and 3 vanishes, as the length of these phases is fixed independently of δ. The expected payoff in phase 2.2 converges[25] to $\mathbf{E}[u(p(\cdot \mid \mathbf{l}_S, \mu_1), x^\star_{\mathbf{l}_S,\mu_1})]$. Finally, for a fixed δ, the expected continuation payoff from stage N is equal to $\mathbf{E}[u_\star(\mathbf{p}_N) + c_1(\mathbf{h}_N)]$. As Lemma 6 will show, $\mathbf{E}[c_1(\mathbf{h}_N)]$ will converge to 0.

Observe that for $\psi^2_\Box = 0$, and following $\mu_1 = \Box$, the message μ_2 of player 2 is non-informative. Thus, conditional on the event that $\mu_1 = \Box$, player 2 does not disclose information prior to phase 4. Thus, for $\psi^2_\Box = 0$, the left-hand side of (21) does not exceed $(1-\zeta)u_\star + \zeta u_{\star\star}$ which by (19) is less than γ^1. If $\psi^2_\Box = 1$, following $\mu_1 = \Box$ the message μ_2 is fully informative, and the left-hand side of (21) is at least equal to

[24] We here abuse notation, since $N \to +\infty$ as $\delta \to 1$. However, the limit of $\mathbf{E}[u_\star(\mathbf{p}_N)]$ is well-defined.

[25] Because the approximation of $x^\star$ by $x_\delta(\vec{a}(x^\star))$ becomes perfectly accurate as $\delta \to 1$.

$\zeta u_\star + (1-\zeta)\left(\theta u_\star + (1-\theta)u_{\star\star}\right)$, which exceeds γ^1 by (19). It follows that for δ high enough, say $\delta \geq \bar{\delta}_1$, there exists $\psi^2_\square(\delta) \in (0,1)$, such that the discounted payoff of player 1 is equal to γ^1, and such that $\psi^2_\square(1) := \lim_{\delta\to 1} \psi^2_\square(\delta) \in (0,1)$.

We conclude this section by discussing how high should δ be, for the profile $(\sigma_\delta, \tau_\delta)$ to be well-defined, and by discussing beliefs and actions off-equilibrium.

We first argue that the costs $c_1(h)$ and $c_2(h)$ are small.

Lemma 6 *There is $c > 0$ such that for every $\delta \geq \bar{\delta}_1$ and every $h \in H_N$, one has*

$$c_1(h) \leq (1-\delta)c.$$

Proof. Because payoffs are bounded by 1, one has

$$\begin{aligned} c_1(h) &\leq (K_1+2K_2)(1-\delta)\delta^{-N} = (K_1+2K_2)\frac{(1-\delta)}{\delta^{K_1+2K_2}}\delta^{-\lfloor\frac{\ln(1-\theta)}{\ln\delta}\rfloor} \\ &\leq (K_1+2K_2)\frac{(1-\delta)}{\delta^{K_1+2K_2+1}}\delta^{\frac{\ln(1-\theta)}{\ln\delta}} = (K_1+2K_2)\frac{(1-\delta)}{\delta^{K_1+2K_2+1}}\frac{1}{\ln(1-\theta)}, \end{aligned}$$

and the result follows. ■

For $\delta \leq 1$ (including $\delta = 1$), denote by $\mathcal{P}(\delta)$ the support of $\mathbf{p}_N$ when $\psi^2_\square$ is set to $\psi^2_\square(\delta)$, and define $\mathcal{Q}(\delta)$ in a symmetric way. Since π_2 and $\varepsilon_2(\cdot \mid m_S)$ have full support, and since $\psi^2, \psi^2_\square(1) \in (0,1)$, one has $\mathbf{p}_N \in \overset{\circ\,\dagger}{\Delta}_{\mathbf{l}_S}(S \times M_S)$, with probability 1.

Because $\mathcal{P}(1)$ and $\mathcal{Q}(1)$ are finite sets, and by Proposition 2, there is $\bar{\delta}_2 < 1$, $\varepsilon > 0$, and neighborhoods $V(p)$ of $p \in \mathcal{P}$, $V(q)$ of $q \in \mathcal{Q}$, such that any payoff in $[u_\star(p'), u_\star(p')+\varepsilon] \times [v_\star(q'), v_\star(q')+\varepsilon]$ is a sequential equilibrium of $\Gamma(p', q')$, for every $p \in \mathcal{P}, p' \in V(p)$, and $q \in \mathcal{Q}, q' \in V(q)$.

In addition, we choose the neighborhoods $V(p)$, $V(q)$ to be small enough, and $C > 0$ so that the conclusion of Proposition 3 holds for every $p \in \mathcal{P}, p' \in V(p)$, and $q \in \mathcal{Q}, q' \in V(q)$.

We choose $\bar{\delta}_3 < 1$ to be high enough so that the following conditions are met for each $\delta \geq \bar{\delta}_3$: (i) every $p' \in \mathcal{P}(\delta)$ belongs to $V(p)$ for some $p \in \mathcal{P}$; (ii) $(1-\delta)c \leq \varepsilon$.

For $\delta \geq \bar{\delta}_3$, the profile $(\sigma_\delta, \tau_\delta)$ is then well-defined, at any information set that is not ruled out by the definition of $(\sigma_\delta, \tau_\delta)$ at earlier stages.

Consider now an information set $I^1_{l,h}$ that is reached with probability 0, and assume that the information set $I^1_{l,h'}$ is reached with positive probability, where h' is the longest prefix of h.

If the sequence h of actions has probability zero, then we let beliefs at $I^1_{l,h}$ and at all subsequent information sets coincide with the belief held at $I^1_{l,h'}$. Player 1 repeats the action that is optimal at $I^1_{l,h'}$.

Assume now that the sequence h has positive probability. This corresponds to the case where player 1 misreported in Phase 1, and played consistently with his report afterwards. Then the belief of player 1 at $I^1_{l,h}$ is well-defined by Bayes' rule (and is independent of player 1's strategy), and only assigns a positive probability to information sets $I^2_{m,h}$ that are reached with positive probability under $\tau_\star$. We let $\sigma_\star$ play at $I^1_{l,h}$ a best reply to $\tau_\star$.

By construction, sequential rationality holds at any information set $I^1_{l,h}$ that is reached with probability zero. One can verify that beliefs are consistent with $(\sigma_\star, \tau_\star)$. We omit the proof.

B.4 Equilibrium properties

We claim that the profile $(\sigma_\delta, \tau_\delta)$ is a sequential equilibrium profile for $\delta < 1$ high enough.

Let $\eta > 0$ be small enough so that

$$\mathbf{E}_{\varepsilon_2, x^\star}[l_S \to k] < \mathbf{E}_{\varepsilon_2, x^\star}[l_S \to l_S] - 2\eta \text{ for every } l_S, k \in L_S, l_S \neq k,$$

and we choose $\bar{\delta}_4 < 1$ such that

$$\mathbf{E}_{\varepsilon_2, x_\delta(\vec{a}(x^\star))}[l_S \to k] < \mathbf{E}_{\varepsilon_2, x_\delta(\vec{a}(x^\star))}[l_S \to l_S] - \eta \text{ for every } l_S, k \in L_S, l_S \neq k \text{ and } \delta \geq \bar{\delta}_4.$$

We finally choose $\delta_5 < 1$ to be such that $1 - \delta^{K_1+2K_2} + (1-\delta)C < \eta\delta^{K_1+2K_2}$ for each $\delta \geq \bar{\delta}_5$.

We now verify that $(\sigma_\star, \tau_\star)$ is a sequential equilibrium as soon as $\delta \geq \max\{\bar{\delta}_4, \bar{\delta}_5\}$. It is sufficient to check that sequential rationality holds at any information set that is reached with positive probability. Let such an information set $I_{l,h}$ be given, and let n be the stage to which $I_{l,h}$ belongs. If stage n belongs to phase 4, then sequential rationality at $I_{l,h}$ follows because continuation strategies in phase 4 form a sequential equilibrium of the associated self-ignorant game. Assume then that $n < N$.

We will make use of the following observation that holds because $\varepsilon_1(\cdot)$, $\varepsilon_2(\cdot)$, π^1 and π^2 have full support: if $I_{l_S,l_T,h}$ is reached with positive probability, then the set of actions that are played with positive probability at $I_{l_S,l_T,h}$ does not depend on l_T, and, therefore, the information set $I_{l_S,l'_T,h}$ is also reached with positive probability, for every $l'_S \in L_S$. We note that the compensation made in phase 4 implies that player 1 is indifferent at $I_{l_S,l_T,h}$ between all actions that are played with positive probability. One thus simply needs to check that player 1 cannot increase his continuation payoff by playing some other action, a.

Assume first that n belongs to either phase 2.1, 2.2 or to phase 3. In that case, the set of actions that are played at $I_{l,h}$ does not depend on l. Hence, when playing a, player 1 triggers a myopic play by player 2, and player 1's overall payoff in that case does not exceed

$$(1-\delta)u_\star(p_n) + \delta \mathbf{E}[u_\star(\mathbf{p}_{n+1}) \mid l, h].$$

On the other hand, the expected continuation payoff of player 1 at $I_{l,h}$ is at least $\delta^N \mathbf{E}[u_\star(\mathbf{p}_N) \mid l, h]$. Sequential rationality then follow from the choice of parameters.

Assume finally that stage n belongs to phase 1. Again, it is not profitable to switch to an action that triggers a myopic play from player 2. What if player 1, instead of reporting l_S, chooses to report $k \neq l_S$? Then, as above, the choice of parameters ensures that it is optimal for player 1 to play consistently with k, at least until phase 4. Such a deviation yields a payoff (discounted back to h) of at most

$$\delta^{-n}\left(\delta^{K_1+K_2}\mathbf{E}_{\varepsilon_2,x_\delta(\vec{a}(x^\star))}[l_S \to k] + (1-\delta)\left(1+\cdots+\delta^{K_1+2K_2-1}\right) + \delta^N \mathbf{E}[u_\star(\mathbf{p}_N) + (1-\delta)C]\right).$$

On the other hand, player 1's continuation payoff when reporting truthfully is at least

$$\delta^{-n}\left(\delta^{K_1+K_2}\mathbf{E}_{\varepsilon_2,x_\delta(\vec{a}(x^\star))}[l_S \to l_S] + \delta^N \mathbf{E}[u_\star(\mathbf{p}_N) + (1-\delta)C]\right).$$

We stress that the distribution of $\mathbf{p}_N$ is the same in both expressions, because the distribution of (μ_1, μ_2) does not depend on player 1's report. The result follows, by the choice of $\bar{\delta}_4$ and $\bar{\delta}_5$.

Assume now that h is longer than K_1 and that player 1 misreported in Phase 1. By definition, the continuation strategy of player 1 is defined to be a best reply to the continuation strategy of player 2. Hence no deviation is profitable. Before we conclude, let us bound the highest continuation payoff at the beginning of Phase 2, conditional on receiving l_S and reporting $i^1 \neq l_S$.

- If player 1 decides to fulfill the requirement of Phase 3, his continuation payoff is at most (roughly, up to terms of the order $(1-\delta)\zeta^2(\varepsilon\mathbf{E}^{\varepsilon^2,x}[l_S \to i^1] + (1-\varepsilon)\mathbf{E}[u_\star(p^0(\cdot \mid l_S, \mu^1, \nu^1)) + (1-\zeta^2)u_\star(p_{B_2})$, which is strictly less than his continuation payoff, had he reported truthfully l_S;

- If player 1 decides to fail the requirement of Phase 3, his continuation payoff is at most $\zeta^2\mathbf{E}[u_\star(p^0(\cdot \mid l_S, \mu^1))] + (1-\zeta^2)u_\star(p_{B_2})$, which is also strictly less that his continuation payoff, had he decided to report truthfully.

Assume finally that h is shorter than K_1.

- If player 1 chooses an action inconsistent with α^1, his continuation payoff is (at most) $u_\star(p^0(\cdot \mid l_S))$, which is less than his payoff along the profile.
- If player 1 chooses to misreport l_S, and say, to report i^1, then there are two cases. If player 1's report in phase 3 is consistent with his deviation in phase 1, then he loses. If player 1's report in phase 3 is inconsistent with his deviation in phase 1, then he loses by the construction in phase 4.

C The proof of Theorem 2

We here briefly show how to deduce Theorem 2 from the proof of Theorem 1. We will assume that with p-probability 1, player 1 has a unique myopically optimal action at $\mathbf{p}_1$, and that the symmetric property[26] holds for player 2.

Consider the following class of strategy profiles. In the first two stages, each player i 'tells' player j whether the information held by j has positive value to i or not. This is done as follows. In stage 1, player i plays his myopically optimal action. In stage 2, player 1, say, repeats this action if $\mathbf{l}_S \notin \tilde{L}_S$, and switches to a different (suboptimal) action if $\mathbf{l}_S \in \tilde{L}_S$ to signal his willingness to disclose/acquire information. If *both* players switched in stage 2, they implement from stage 3 on an equilibrium such as we designed in the proof of the Main Theorem. Otherwise, players repeat their stage 1 action. The sole role of stage 1 is to instruct the other player *how* to interpret the action played in stage 2.

If $\mathbf{l}_s \notin \tilde{L}_S$, it is strictly dominant for player 1 to repeat his optimal action throughout, as required. Indeed, playing a different action in stage 2 would only lower player 1's payoff, with no benefit since player 2's information is valueless.

If $\mathbf{l}_s \in \tilde{L}_S$, player 1's overall payoff is $u_\star(\mathbf{p}_1)$ if he pretends that the information held by player 2 is valueless. However, because there is a positive q-probability that

[26]If a player has two myopically optimal actions at $\mathbf{p}_1$, he can costlessly reveal information to the other player.

$\mathbf{m}_T \in \tilde{M}_T$, it is a best reply for player 1 to switch to a suboptimal action in stage 2 as soon as the value of the information disclosed by player 2 exceeds on average the cost incurred[27] in stage 2.

In such an equilibrium, conditional on $\mathbf{l}_S$, player 1's payoff is $u_\star(\mathbf{p}_1)$, which is then also equal to $\mathbf{E}[u_\star(\tilde{\mathbf{p}}) \mid \mathbf{l}_S]$, if $\mathbf{l}_S \notin \tilde{L}_S$. If instead $\mathbf{l}_S \in \tilde{L}_S$, then with probability $q(\tilde{M}_T)$, player 1's payoff may be as high as $\mathbf{E}[u_\star(\tilde{\mathbf{p}}) \mid \mathbf{l}_S]$. Otherwise, player 1's payoff will be (approximately) $u_\star(\mathbf{p}_1)$. The *ex ante* expected payoff can therefore be as high as

$$u_\star(1 - q(\tilde{M}_T)) + q(\tilde{M}_T) u_{\star\star}.$$

[27] It cannot be optimal for player 1 to pretend that $\mathbf{l}_S \in \tilde{L}_S$, yet to lie about his optimal action.