Belief-free equilibria in games with incomplete information*

Johannes Hörner† and Stefano Lovo‡

May 2006

Abstract

We define belief-free equilibria in two-player games with incomplete information as sequential equilibria for which players' continuation strategies are best-replies, after every history, independently of their beliefs about the state of nature. We characterize a set of payoffs that includes all belief-free equilibrium payoffs. Conversely, any payoff in the interior of this set is a belief-free equilibrium payoff.

JEL codes: C72, C73

1 Introduction

The purpose of this paper is to characterize the set of payoffs that can be achieved by a particular class of equilibria, belief-free equilibria. The games considered are two-player repeated games with (potentially two-sided) incomplete information, under discounting. The restriction we impose is that the players' equilibrium strategies be optimal independently of their beliefs, from any history on. This concept is not new: it has been introduced in the context of repeated games with imperfect private monitoring by Ely and Välimäki (2002) and further studied in Ely, Hörner and Olszewski (2005). It is also related to the concept of ex post equilibrium that is used in mechanism design (see Crémer and McLean (1985)) and in large finite games (see Kalai (2004)).

To predict players' behavior in games with unknown parameters, a model typically includes the specification of the players' subjective probability distributions over these unknowns, following Harsanyi (1967). This is not necessary when belief-free equilibria are considered. Just as

*We thank Jeff Ely for useful comments.

†KSM-MEDS, Northwestern University, 2001 Sheridan Road, Evanston, IL60201, and HEC School of Management, Paris, and CEPR. j-horner@kellogg.northwestern.edu

‡HEC School of Management, Paris. lovo@hec.fr

ex post equilibria, belief-free equilibria enjoy the desirable property that the beliefs about the underlying uncertainty are irrelevant. Indeed, players' beliefs need not be derived by Bayes' rule from a common prior. Further, the way in which players update their beliefs as the game unfolds is irrelevant.

Therefore, while solving for belief-free equilibria requires the game to be fully specified, it does not require that all players know all the parameters of the model. In this sense, this idea is close to von Neumann and Morgenstern (1944)'s original purpose when they introduced "games of incomplete information", as games in which some parameters remain unspecified. Such equilibria are also consistent with *misperceptions*, as defined by Luce and Raiffa (1957). Nevertheless, our players remain expected utility maximizers, just as in the case of complete information: players are allowed to randomize, and take expectations when evaluating their payoff.

Most importantly, such equilibria are sequential equilibria (for any prior) satisfying any potentially desirable refinement. However, we do not view belief-free equilibrium as an equilibrium refinement per se. In fact, belief-free equilibria need not exist. Rather, our purpose is to characterize which equilibrium payoffs are sensitive to the specification of beliefs, and which are not. The required robustness is extreme in the sense that an equilibrium may be robust to small changes in beliefs, and yet not be belief-free. Also, it could be that some payoff vector is a sequential equilibrium payoff vector for all possible beliefs, but the particular equilibrium achieving this payoff depends on the beliefs; while we characterize equilibrium payoffs, our restriction is one on strategies, not on payoffs.

We provide a set of necessary conditions that belief-free equilibrium payoffs must satisfy which defines a closed convex set, possibly empty. Conversely, we prove that every interior point of this set is a belief-free equilibrium payoff, provided that players are sufficiently patient.

This set of payoffs already plays a prominent role in the literature on Nash equilibria in games with one-sided incomplete information. Shalev (1994) considers the case of private values (the uninformed player knows his own payoffs) and shows that the set of uniform (undiscounted) Nash equilibrium payoffs can be derived from this set. Closest to our analysis is Cripps and Thomas (2003) which considers the one-sided case with private values as well, but with discounting. Most relevant here is their Theorem 2, in which they show that the payoffs in the strict interior of this set are Nash equilibria for all priors. In general, however, the set of Nash equilibrium payoffs is larger, as they demonstrate in their Theorem 3 which establishes a folk theorem. Forges and Minelli (1997) is also related. They show how communication can significantly simplify the construction of strategies that achieve the Nash equilibrium payoffs. These simple strategies also appear in Koren (1988). The most general characterization of Nash equilibrium payoffs remains Hart (1985) for the case of one-sided incomplete information. A survey is provided by Forges (1992).

As mentioned, the concept of belief-free equilibria has already been introduced in the context of games with complete but imperfect information. There, the restriction on the equilibrium pertains to the private history observed by the opponent. In both contexts, the characterization of equilibrium payoffs is very tractable, although these characterizations are quite different.

Non-trivial belief-free equilibria under imperfect monitoring involve players being at least periodically indifferent across continuation strategies. This is no longer true in what follows. In the equilibrium that we construct for the proof, the only time players are potentially indifferent across actions is if minmaxing their opponent calls for randomization.

The next section introduces the two necessary conditions. A leading example is introduced, for which these conditions are explicitly worked out. Section 3 provides the theorem, and gives a relatively short proof using explicit communication. The proof without such communication is given in Appendix. Section 4 applies our logic to another example, a game of bad reputation introduced by Ely and Välimäki.

2 The model

We consider repeated games with (two-sided) incomplete information, as defined by Harsanyi [1967-68] and Aumann and Maschler [1995]. There is an $J \times K$ array of 2-person games in normal form, having the same number of actions for each player. Player 1 is told in which row the true game lies but he is not told which of the games in that row is actually being played. Player 2 is told in which column the true game lies but he is not told which of the games in that column is the true game. Players observe all actions, but not their payoffs. More formally, the stage-game is a finite-action game. Let A_1 and A_2 be the finite sets of actions for player 1 and 2 respectively, where A_i has at least two elements. Let $A = A_1 \times A_2$. When the row is j and the column is k -for short, when the state is (j,k)-, Player i's payoff function is denoted u_i^{jk}, for $i = 1, 2$. We extend the domain of u_i^{jk} from pure action profiles $a \in A$ to mixed actions $\alpha \in \triangle A$ in the standard way. We let $u_1^k := \left\{u_1^{jk}\right\}_{j=1}^{J}$, $u_2^j := \left\{u_2^{jk}\right\}_{k=1}^{K}$. Players select an action in each period $t = 1, 2, \ldots$. Players observe actions, but not payoffs. Let $H^t = (A_1 \times A_2)^{t-1}$ be the set of all possible histories h^t up to and including period t. A (behavioral) strategy for row j, or *type* j, of Player 1 (resp. type k of Player 2) is a sequence of maps $s_1^j := \left(s_1^{j,0}, s_1^{j,1}, \ldots\right)$, $s_1^{j,t} : H^t \to \triangle A_1$ (resp. $s_2^k := \left(s_2^{k,0}, s_2^{k,1}, \ldots\right)$, $s_2^{k,t} : H^t \to \triangle A_2$). We define $s_1 := \left\{s_1^j\right\}_{j=1}^{J}$, $s_2 := \left\{s_2^k\right\}_{k=1}^{K}$. Players use a common factor $\delta < 1$.

Example 1 (Prisoner's dilemma with one-sided incomplete information) *Player* 1 *is informed of the true state (= the row), Player 2 is not, and there is only one column ($J = 2$, $K = 1$). If the true game corresponds to $j = 1$, payoffs are given (in every period) by the prisoner's dilemma payoff matrix in which T is "Cooperate" and B is "Defect". If the true game corresponds to $j = 2$, payoffs are given by the prisoner's dilemma payoff matrix in which B is "Cooperate" and T is "Defect". The payoffs in the first case are*

	T	B
T	$1, 1$	$-L, 1+G$
B	$1+G, -L$	$0, 0$

and in the second state are

	T	B
T	$0,0$	$1+G,-L$
B	$-L,1+G$	$1,1$

As usual, we maintain the assumption that cooperation is efficient: $L-G>-1$.

Our purpose is to characterize the payoffs that can be achieved, with low discounting, by a special class of sequential equilibria. In a *belief-free* equilibrium, each player's continuation strategy, after any history, is a best-reply to his opponent's continuation strategy, independently of his beliefs about the state of the world, and therefore, independently of his opponent's private information. Such equilibria are trivially sequential equilibria that satisfy any belief-based requirement. At the same time, they do not require players to be Bayesian, or to share a common prior. Because they are belief-free, they must in particular induce a subgame-perfect equilibrium in every complete information game that is consistent with the player's private information. Formally, $s:=(s_1,s_2)$ is a belief-free equilibrium if it is the case that, for all (j,k), $\left(s_1^j,s_2^k\right)$ is a subgame-perfect Nash equilibrium of the infinitely repeated game with stage-game payoffs given by $\left(u_1^{jk},u_2^{jk}\right)$.

As mentioned, belief-free equilibria have been previously introduced in and applied to games with imperfect private monitoring. As discussed, with incomplete information, there is no need for randomization on the equilibrium. Indeed, in our construction, along the equilibrium path, players always have a strict preference to play some particular action. Of course, this action potentially depends on a player's private information (and on the history). In our construction, randomization only appears during punishment phases, as is standard in folk theorems that do allow for mixed strategies to determine minmax payoffs, as we do.

A belief-free equilibrium (s_1,s_2) determines, for each Player i, a $J\times K$ array of equilibrium payoffs v_i^{jk}. Consider $i=1$. Conditional on the column k he is being told, Player 2 knows that Player 1's equilibrium payoff is one among the coordinates of the vector $v_1^k=\left(v_1^{1k},\ldots,v_1^{Jk}\right)$. Because the equilibrium is belief-free, Player 1's payoff must be individually rational in the special case in which his beliefs are degenerate on the true column k. It is therefore necessary that for each column k, Player 2 has one *punishment* strategy $\widehat{s}_2^k$ which guarantees that, independently of Player 1's strategy, Player 1 gets no more than v_1^{jk} *for all* j. This ensures that no matter Player 1's information and belief on the row j and the column k respectively, he prefers the equilibrium payoff v_1^{jk} to the payoff received when Player 2 uses strategy $\widehat{s}_2^k$. If $J=1$, so that the game is of one-sided incomplete information, this requirement on Player 1's payoff is weak: for each k, Player 1 must receive at least as much as his minmax payoff (in mixed strategies) in the true game being played. In general however, this is a stringent restriction, as it implies that the set of belief-free equilibria is empty for some games.

Example 2: *(Non-existence of belief free equilibria) Player* 1 *is informed of the true state (= the row), Player 2 is not ($J = 2$, $K = 1$). The payoffs are either*

	L	R
U	$10, -4$	$1, 1$
D	$1, 1$	$0, 0$

or

	L	R
U	$0, 0$	$1, 1$
D	$1, 1$	$10, -4$

For each state, Player 2 must be guaranteed to get at least 0 *in a belief-free equilibrium: his equilibrium strategy must be optimal given any beliefs he may have, including degenerate beliefs on the true state. His payoff must therefore be at least as large as his minimax payoff given the true state, which exceed 0 in both states. This implies that the action profile yielding -4 to Player* 2 *cannot be played more than a fifth of the time in equilibrium. Equivalently, this means that Player* 1 *equilibrium payoff is at most* 14/5 *in each state. However, if Player* 1*randomizes between U and D independently of the state, he is guaranteed to get at least* 3 *in one of the states, a contradiction. [This state will typically depend on Player 2's strategy. However, no strategy of Player 2 can bring down Player 1's payoff below 3 in both states simultaneously.]*

Computing these minmax levels is in general tedious. Note that these levels are vectors, not scalars: punishing severely Player 1 in one row may require leaving him a high payoff in another row. Therefore, different equilibrium payoffs may call for different "punishment" strategies.

For a given $p \in \triangle\{1, \ldots, J\}$(resp. $q \in \triangle\{1, \ldots, K\}$), let $b_1^k(p)$ (resp. $b_2^j(q)$) be the value for Player 1 (resp. Player 2) of the one shot game with payoff matrix $p \cdot u_1^k$ (resp. $q \cdot u_2^j$). We say that a vector $v_1 \in \mathbb{R}^{J \times K}$ is *individually rational* for Player 1 if for any k it is the case that

$$p \cdot v_1^k \geq b_1^k(p), \forall p \in \triangle\{1, \ldots, J\},$$

where $v_1^k := \left\{v_1^{jk}\right\}_{j \in J}$. Similarly, $v_2 \in \mathbb{R}^{J \times K}$ is *individually rational* for Player 2 if for any j it is the case that

$$q \cdot v_2^j \geq b_2^j(q), \forall q \in \triangle\{1, \ldots, K\}$$

where $v_2^j := \left\{v_2^{jk}\right\}_{k \in K}$. Approachability (Blackwell (1968)) can be used to show that, if $v_1 \in \mathbb{R}^{J \times K}$ is individually rational for Player 1, then for any column k, Player 2 has a punishment strategy $\widehat{s}_2^k$such that Player 1's average payoff cannot be larger than v_1^{jk} for all j independently of the strategy he uses. (Obviously, an analogous statement holds for Player 2). [Approachability is usually defined for payoffs evaluated according to the limit of means rather than discounting, but the uniform versions of the results that we will use imply the respective counterparts for discounting, provided the discount factor is close enough to one. See Cripps and Thomas (2003) for details on the issue of discounting.]

Necessary condition 1 (Individual Rationality): *If v_i is a belief-free equilibrium payoff array, then v_i is individually rational.*

We apply approachability to compute these minmax levels in our first example.

Example (Prisoner's dilemma with one-sided incomplete information, continued) *It is shown in Appendix that, in the prisoner's dilemma described above, (v_1^1, v_1^2) is a minmax vector for Player 1 if and only if:*

$$v_1^2 > \max\left\{\frac{1+G}{1+L}\left(1-v_1^1\right), 1-\frac{1+L}{1+G}v_1^1\right\},\ v_1^1 \geq 0,\ v_1^2 \geq 0.$$

when $G < L$, and

$$\begin{aligned} v_1^2 &> \begin{cases} \max\left\{\frac{1+G}{1+L}\left(1-v_1^1\right), 1-\frac{1+L}{1+G}v_1^1\right\}, & \text{for } v_1^1 < \frac{1+G}{2+G+L}, \text{ or } v_1^1 > \frac{(1+G)(1+G+L)}{2+G+L} \\ v_1^1 + \frac{(G+L)(2+G+L)-2\sqrt{(G+L)(2+G+L)}\sqrt{v_1^1(G-L)+L(1+G)}}{G-L} & \text{otherwise} \end{cases} \\ v_1^1 &> 0, v_1^2 > 0. \end{aligned} \tag{1}$$

when $G > L$. See Figure 1.

Note that in this example, individual rationality for Player 2 is straightforward. Because the equilibrium is belief-free, Player 2's payoff must be individually rational in the special case in which his beliefs are degenerate on the true column j. In particular, for each row j, Player 1 has one punishment strategy s_1^j which guarantees that, independently of Player 2's strategy, Player 2's payoff is at most 0. Therefore a payoff $v_2 = (v_2^0, v_2^1)$ is individually rational for Player 2 only if $v_2^j > 0$.

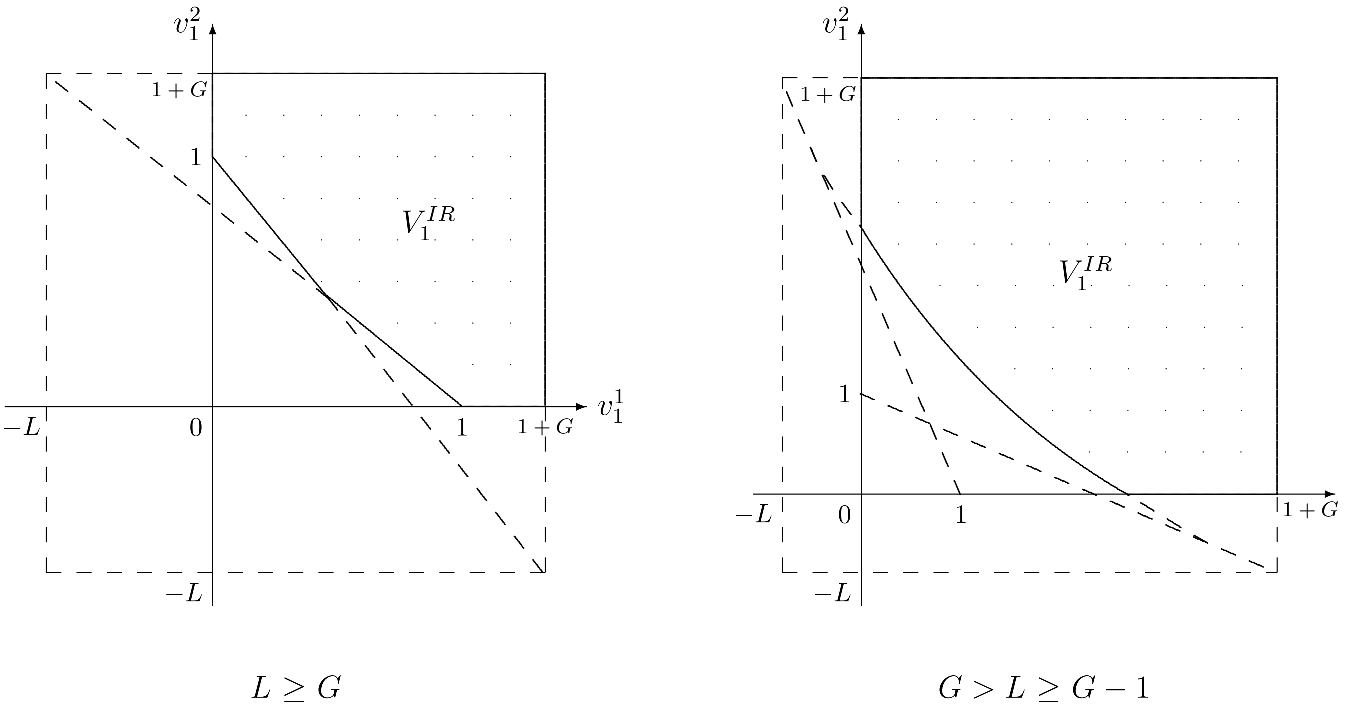


Figure 1: Player 1's individually rational payoffs

In a belief-free equilibrium, play may depend on a player's private information. That is, Player 1's equilibrium strategy s_1^j typically depends on the row j he is told, and Player 2's strategy s_2^k on the row k he is told. Since Player 1's strategy s_1^j must be a best-reply to s_2 independently of his beliefs, it must be a best-reply to s_2^k, that is, when he assigns probability one to the true column k. In particular, s_1^j must be a better-reply to s_2^k than $s_1^{j'}$, $j' \neq j$, when the row is j. This is our second necessary condition. To state it in terms of payoffs, observe that each pair $\left(s_1^j, s_2^k\right)$ induces a distribution $\{\Pr\{a \mid (j,k)\} : a \in A\}_{j,k}$ over action profiles, where:

$$\Pr\{a \mid (j,k)\} = (1-\delta) \sum_{t=1}^{\infty} \delta^{t-1} \Pr\left\{a_t = a \mid \left(s_1^j, s_2^k\right)\right\},$$

and $\Pr\left\{a_t = a \mid \left(s_1^j, s_2^k\right)\right\}$ is the probability that action a is played in period t given the strategy profile $\left(s_1^j, s_2^k\right)$. Therefore:

Necessary Condition 2 (Incentive Compatibility): *If (v_1, v_2) is a pair of belief-free equilibrium payoff arrays, there must exist $\{\Pr\{a \mid (j,k)\} : a \in A\}_{j,k}$ such that, for all (j,k):*

$$v_1^{jk} = \sum_a \Pr\{a \mid (j,k)\} u_1^{jk}(a) \geq \sum_a \Pr\{a \mid (j',k)\} u_1^{jk}(a), \text{ and}$$
$$v_2^{jk} = \sum_a \Pr\{a \mid (j,k)\} u_2^{jk}(a) \geq \sum_a \Pr\{a \mid (j,k')\} u_2^{jk}(a).$$

If such distributions exist, we say that (v_1, v_2) is *incentive compatible.* Incentive compatible payoffs always exist, since the constraints are always satisfied if $\Pr\{a \mid (j,k)\}$ is independent of (j,k). However, not every pair of payoff arrays is incentive compatible.

Example: Prisoner's Dilemma with one-sided incomplete information, continued) *The pair of payoff arrays $\{(v_1^1, v_1^2), (v_2^1, v_2^2)\}$ is incentive compatible if and only if there exists $(\mu_{TT}^1, \mu_{TB}^1, \mu_{BT}^1)$ and $(\mu_{BB}^2, \mu_{BT}^2, \mu_{TB}^2)$ such that $\mu_{TT}^1 \geq 0$, $\mu_{BB}^2 \geq 0$, $\mu_{BT}^j \geq 0$, $\mu_{TB}^j \geq 0$, $j = 1,2$, $\mu_{TT}^1 + \mu_{TB}^1 + \mu_{BT}^1 \leq 1$, $\mu_{BB}^2 + \mu_{BT}^2 + \mu_{TB}^2 \leq 1$, and:*

$$v_1^1 = \mu_{TT}^1 - \mu_{TB}^1 L + \mu_{BT}^1 (1+G) \geq \left(1 - \mu_{BB}^2 - \mu_{BT}^2 - \mu_{TB}^2\right) - \mu_{TB}^2 L + \mu_{BT}^2 (1+G),$$
$$v_1^2 = \mu_{BB}^2 - \mu_{BT}^2 L + \mu_{TB}^2 (1+G) \geq \left(1 - \mu_{TT}^1 - \mu_{TB}^1 - \mu_{BT}^1\right) - \mu_{BT}^1 L + \mu_{TB}^1 (1+G),$$
$$v_2^1 = \mu_{TT}^1 + \mu_{TB}^1 (1+G) - \mu_{BT}^1 L,\ v_2^2 = \mu_{BB}^2 + \mu_{BT}^2 (1+G) - \mu_{TB}^2 L.$$

For each player, we may characterize the set of payoff arrays that are incentive compatible and, in addition, satisfy individual rationality for the other player - a constraint that is necessary for equilibrium. We only describe the resulting set V_i^{IC} in some detail for Player 1, and display the sets for both players in Figures 2 and 3. In either case, the problem is a standard (finite-dimensional) optimization problem. Observe that, when $G > L$, this problem is not linear for Player 2, as the set of Player 1's individual rational payoffs is not a polytope.

Recall that, in the prisoner's dilemma, a player cannot get more than $1 + G/(1+L)$ if his opponent is guaranteed at least 0. Therefore, $v_1^j \leq 1 + G/(1+L)$, $j = 1,2$. The point $(1 + G/(1+L), 1 + G/(1+L))$ is obtained by setting $\mu_{TT}^1 = 1 - \mu_{TB}^1 = \mu_{BB}^2 = 1 - \mu_{BT}^2 = L/(1+L)$. The point $(-L/(1+L), 1 + G/(1+L))$ is obtained by setting $\mu_{TT}^1 = \mu_{BT}^1 = 0$, $\mu_{TB}^1 = \mu_{TB}^2 = 1 - \mu_{BB}^2 = 1/(1+L)$. Further extreme points depend on the value of $L - G$:

(i) if $L - G \geq 1$, we get the point $\left(\frac{3}{2-G+L} - 1, \frac{3}{2-G+L} - 1\right)$ by picking:

$$\left(\mu_{TT}^1, \mu_{TB}^1, \mu_{BT}^1\right) = \left(\mu_{BB}^2, \mu_{BT}^2, \mu_{TB}^2\right) =$$
$$\left(1 - \frac{3}{2-G+L}, \frac{1}{2}\left(1 - \frac{L-(1+G)}{2-G+L}\frac{2+G+L}{1+G+L}\right), \frac{1}{2}\left(1 - \frac{L-(1+G)}{2-G+L}\frac{G+L}{1+G+L}\right)\right);$$

(ii) if $1 > L - G \geq 0$, we get $((1+G-L)/2, (1+G-L)/2)$ by picking $(\mu_{TT}^1, \mu_{TB}^1, \mu_{BT}^1) = (\mu_{BB}^2, \mu_{BT}^2, \mu_{TB}^2) = (0, 1/2, 1/2)$;

(iii) if $0 > L - G \geq -1$*, we get the point* $(0,1)$ *(and analogously the point* $(0,1)$*) by picking* $(\mu^1_{TT}, \mu^1_{TB}, \mu^1_{BT}) = (0,0,0)$ *and* $(\mu^2_{BB}, \mu^2_{BT}, \mu^2_{TB}) = (1,0,0)$.

The three cases are illustrated below, as are the two cases for Player 2.

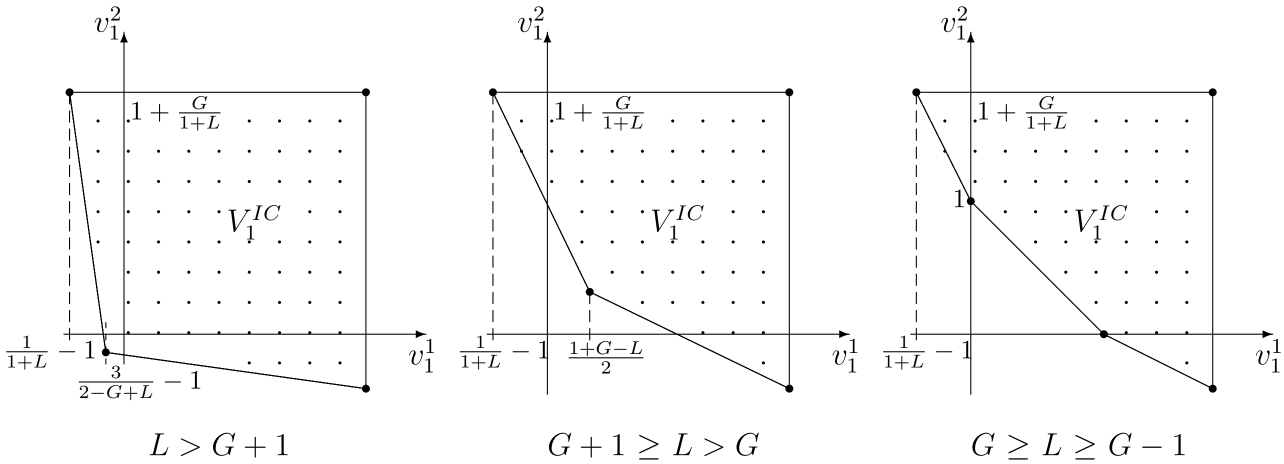


Figure 2: Incentive compatible payoffs for Player 1
(for some individually rational payoffs for Player 2)

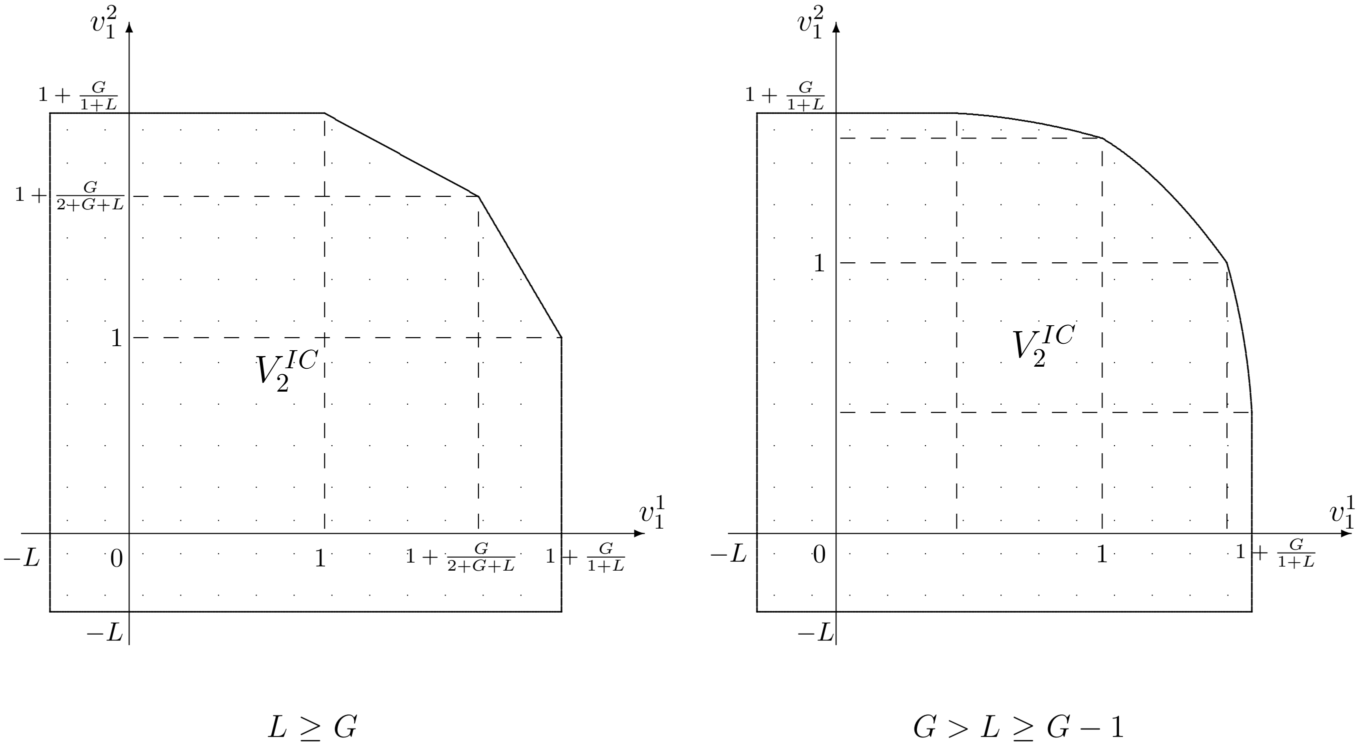


Figure 3: Set of Player 2's payoffs
(for some individually rational and incentive compatible payoffs for Player 1)

3 Theorem

Let V^* denote the feasible set of pairs of payoff arrays satisfying conditions 1 and 2. It is clear that V^* is convex. We prove that:

Theorem Fix some v in the interior of V^*. The pair of payoff arrays v is achieved in some belief-free equilibrium if players are sufficiently patient.

This theorem establishes that the necessary conditions are 'almost' sufficient. It is then natural to ask whether we can get an exact characterization. The strict inequalities corresponding to individual rationality cannot be weakened, in general. One (but not the only) reason for this is that our optimality criterion involves discounting, while Blackwell's result is for the undiscounted case. The strict inequalities corresponding to incentive compatibility may be weakened when V^*

has nonempty interior. It then suffices that, if $v \in V^*$ and some incentive compatibility for Player 1, say, binds at v (and nothing else binds), we can find $v' \in V^*$ such that $v'_1 = v_1$ but $v'^{jk}_2 \neq v^{jk}_2$ for all (j,k). However, for the interesting case in which V^* has empty interior, this may not be possible. Consider for instance the case of one-sided incomplete information; Player 1 knows the row, but his payoff does not depend on the row, so that the incentive compatibility constraints necessarily bind. A difficulty is then to induce Player 1 to play the minmaxing strategy after a deviation by Player 2 that is appropriate given the true row. Because there is no possibility to provide strict incentives for 'truthtelling' after the punishment phase, it is then necessary that the punishment strategy itself be incentive compatible, which reduces the scope for punishment, and changes the relevant individual rationality constraints.

Figure 4 and 5 display the resulting equilibrium payoffs in the prisoner's dilemma. Observe however, that these are the projections of belief-free equilibrium payoff pairs onto each player's payoff space. It is not true that any pair of vectors selected from these projections is a pair of belief-free equilibrium payoff vectors: incentive compatibility imposes some restrictions on the possible pairs.

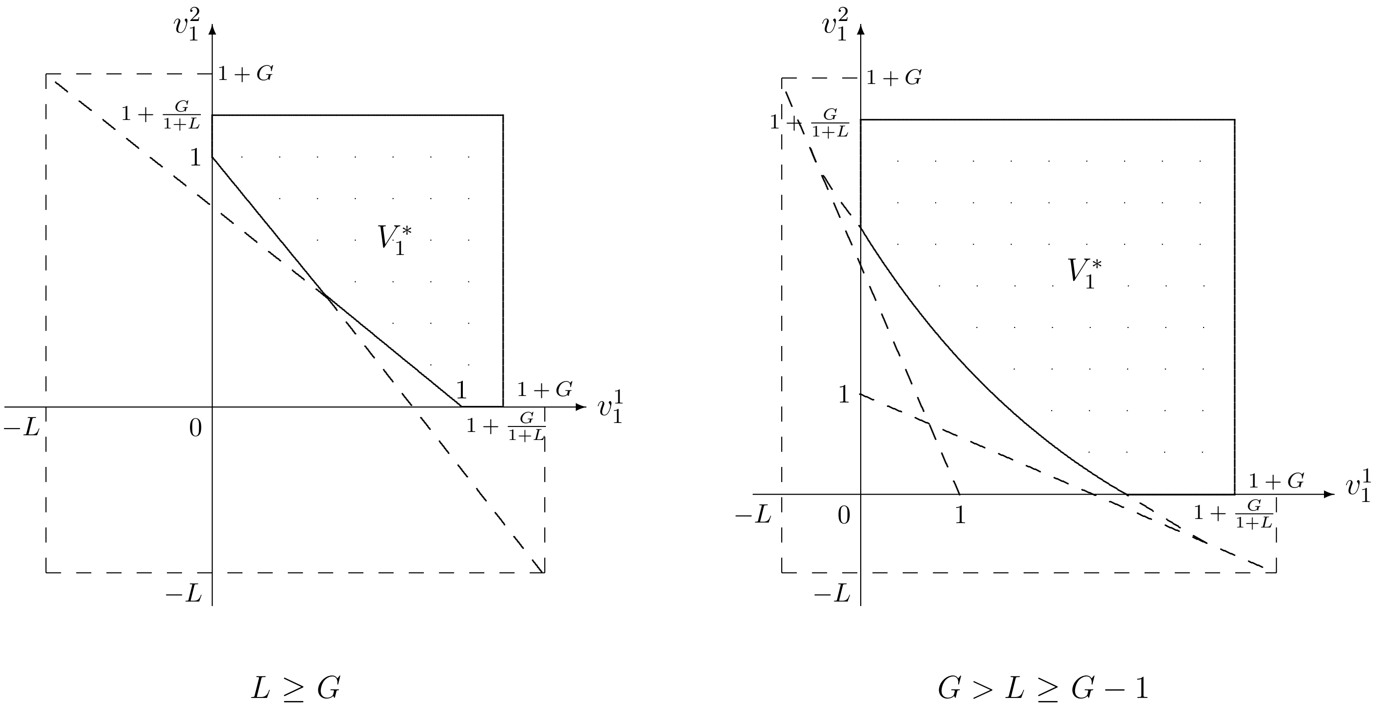


Figure 4: Belief-free equilibrium payoffs for Player 1 as $\delta \to 1$

The proof of the theorem is constructive. As a first attempt, one may want to try to construct an equilibrium as follows: initially, players signal (through their choices of actions) what their private information is, and then play according to the probability distribution over action profiles corresponding to their messages; further, any deviation from this play is punished by minmaxing (in the sense of Blackwell). Indeed, this is essentially the construction of Koren (1988). The problem is that it is not belief-free: if a player is convinced that his opponent's type is not the one that he has signalled, the corresponding play need no longer be individually rational.

The actual construction is therefore more involved, to guarantee that beliefs are irrelevant, after every possible history. In the simple proof presented below, we assume that there is a public randomization device and that players can communicate at no cost in every period. Both assumptions can be dropped, as proved in Appendix.

At the end of every period each player announces his private information. Then players play a correlated action profile that only depends on the last announcement made by the players. These correlated action profiles are such that Player i obtains an expected payoff of v_i^{jk} whenever the true state is jk and players announce jk, and that this payoff is higher than what each player can secure by an unilateral deviation in the message he sends, independently of the beliefs he holds, which is possible by condition (2), since $v \in V^*$. This ensures that each player is willing to signal his information truthfully, regardless of his beliefs. In case a player chooses an action that is not consistent with the correlated strategy corresponding to the last announcement, he is then punished during T periods (T long but finite, with $\delta^T \simeq 1$). As $v \in V^*$, and in particular because of condition 1, each player has a punishment strategy that forces the other player's payoff strictly below its equilibrium level in each state of the nature, provided it is used sufficiently long. Of course, we also need to make sure that play during such a punishment phase is also belief-free, and this introduces additional complications.

By repeating announcements in every period, we avoid the problem in the first paragraph, in case the belief of a player about his opponent's private information does not coincide with the signal the opponent has just sent. Since equilibrium strategies specify that players communicate truthfully their information at the end of each period, the player's expected payoff is hardly affected by such a situation (when δ is sufficiently close to one), since he expects his opponent to revert to what he believes is the true signal within one period.

It is simplest to start proving the proposition under the assumption that players can communicate at no cost, and that they have access to a public randomization device: at the beginning of each period, the outcome of a random variable, independent across time and of the state of the world, is publicly observed. For concreteness, assume that it is uniformly distributed on $[0, 1]$. The proof for the general case is in Appendix.

So, assume for now that, at the end of each period (including at the end of period "0", that is, at the beginning of the game), players simultaneously make an announcement that is publicly observable. The set of messages is the set of rows and columns, respectively. That is, Player 1 announces $j' = 1, \ldots, J$, while Player 2 announces $k' = 1, \ldots, K$.

We first describe the equilibrium strategies, and then check that the strategies yield the desired payoff arrays, are belief-free, and that no deviation is profitable.

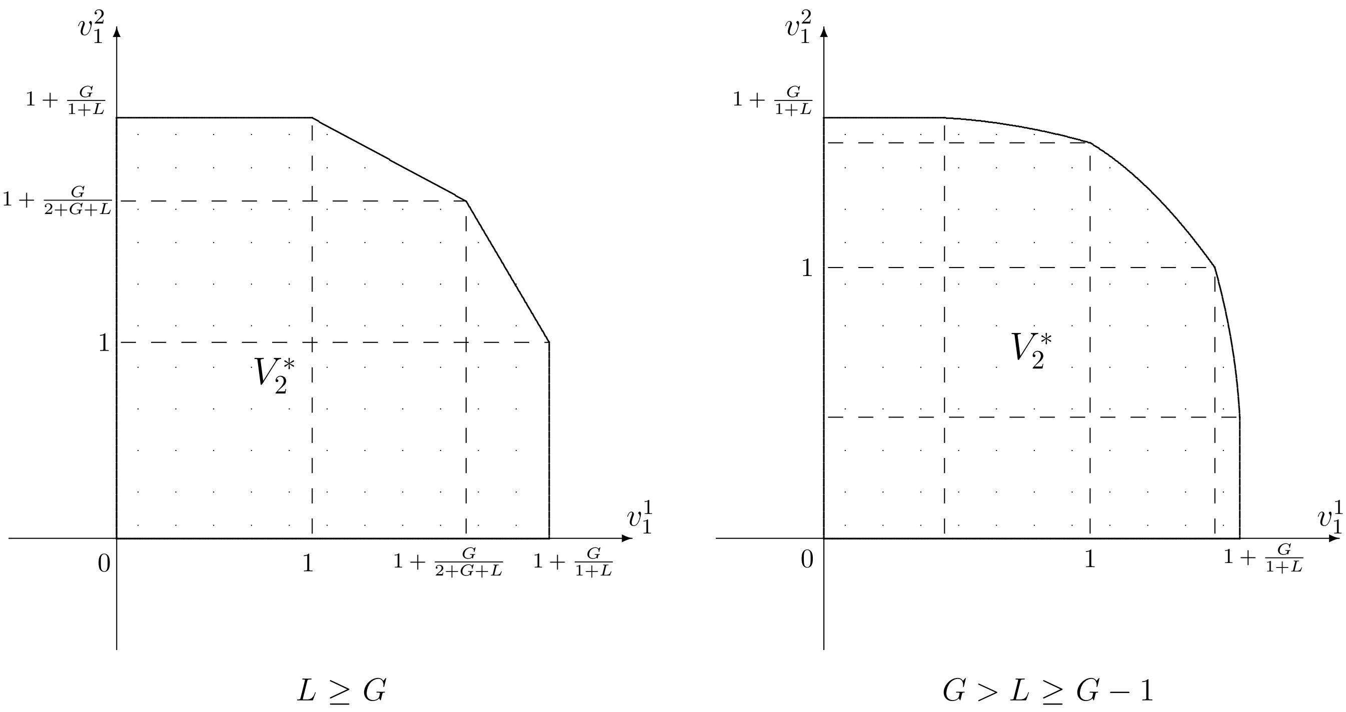


Figure 5: Belief-free equilibrium payoffs for Player 2 as $\delta \to 1$

Equilibrium Strategies:

The play can be divided in *phases*, which are similar to, but not to be confused with states of an automaton.

Phases: There are two kinds of phases: regular phases last only one period, while punishment phases last at most T periods, where T is to be specified. Regular phases are denoted $R^{jk}(\varepsilon_1, \varepsilon_2)$ where $\varepsilon_1, \varepsilon_2 \in \mathbb{R}$, $j = 1, \ldots, J$, $k = 1, \ldots, K$. Punishment phases are denoted P_1^k, P_2^j.

Actions:

(i) *Regular Phase*: In each regular phase $R^{jk}(\varepsilon_1, \varepsilon_2)$, actions are determined by the outcome of the public randomization device. Each action profile a is selected with probability

$\Pr\left\{a \mid R^{jk}(\varepsilon_1, \varepsilon_2)\right\}$. Let:

$$v_i^{jk}\left(R^{j'k'}(\varepsilon_1, \varepsilon_2)\right) := \sum_{a \in A} \Pr\left\{a \mid R^{j'k'}(\varepsilon_1, \varepsilon_2)\right\} u_i^{jk}(a),$$

and let $v_i(R(\varepsilon_1, \varepsilon_2)) = \left\{v_i^{jk}\left(R^{jk}(\varepsilon_1, \varepsilon_2)\right)\right\}_{j,k}$ denote the corresponding $J \times K$ array. These probabilities (and a real number $\bar{\varepsilon} > 0$) are chosen such that:

$$v_i(R(\varepsilon_1, \varepsilon_2)) = v_i + \varepsilon_i, \tag{2}$$

and:

$$v_1^{jk}\left(R^{jk}(\varepsilon_1, \varepsilon_2)\right) > v_1^{jk}\left(R^{j'k}(\varepsilon_1', \varepsilon_2')\right),\ v_2^{jk}\left(R^{jk}(\varepsilon_1, \varepsilon_2)\right) > v_2^{jk}\left(R^{jk'}(\varepsilon_1', \varepsilon_2')\right), \tag{3}$$

for all $i = 1, 2$, $\varepsilon_i, \varepsilon_i' \in [-\bar{\varepsilon}, \bar{\varepsilon}]$, $j = 1, \ldots, J$, $k = 1, \ldots, K$, $j' \neq j$, $k' \neq k$. This is possible since v is in the interior of V^*.

At the end of the regular phase, messages are reported truthfully independently of the specific regular phase, of the outcome of the public randomization device and of the realized action profile.

(ii) *Punishment phase*: The punishment phase lasts at most T periods.[1] Without loss of generality, we describe here the actions/messages in phase P_1^k. The (behavior) strategy $\widehat{s}_2^k$ of Player 2 during the punishment phase P_1^k is such that the average discounted payoff of Player 1 in the T periods of the phase conditional on state (j, k) is not larger than $v_1^{jk} - 2\bar{\varepsilon}$, which is possible since v is in the interior of V^*.

We now define T, $\bar{\delta} < 1$, if necessary decrease $\bar{\varepsilon} > 0$ (introduced above), and specify strategies in the punishment phase so as to satisfy the following inequalities, for all j, k, and $i = 1, 2$:

$$-(1-\delta) M + \delta\left(v_1^{jk} - \bar{\varepsilon}\right) > (1-\delta) M + \delta\left(\left(1-\delta^T\right)\left(v_1^{jk} - 2\bar{\varepsilon}\right) + \delta^T\left(v_1^{jk} - \bar{\varepsilon}\right)\right), \tag{4}$$

(along with the corresponding inequality for Player 2 in phase P_2^j)

$$-\left(1-\delta^T\right) M + \delta^T v_i^{jk} > \left(1-\delta^T\right) M + \delta^T\left(v_i^{jk} - 2\bar{\varepsilon}/3\right). \tag{5}$$

To see that such T, $\bar{\delta}$, $\bar{\varepsilon}$ and strategies exist, observe that for a fixed but small enough $\bar{\varepsilon} > 0$, (4) can be satisfied for all T large enough and $\delta > \bar{\delta}$ for $\bar{\delta}$ close enough to one. Increasing the value of $\bar{\delta}$ if necessary, (5) can be satisfied as well.

Returning to the specification of actions and message, as long as the punishment phase P_1^k lasts (i.e. for at most T periods), Player 2 plays according to $\widehat{s}_2^k$ (given k and the history starting in the initial period of P_1^k). Observe that $\widehat{s}_2^k$ need not be pure. Player 1 plays a best-reply $s_1^{j,k}$ to

[1] If we described the equilibrium strategies by a formal automaton, we would introduce as many states of the automaton as possible histories within each punishment phase. We feel that this would needlessly clutter the exposition.

$\widehat{s}_2^k$, conditional on the true column being k. Without loss of generality, we pick s_1^{jk} to be pure. Observe that s_1^{jk} may depend on j.

Players report truthfully states in all periods of the punishment phase.

Initial phase: As mentioned, players send messages before the beginning of the game. These initial announcements are made truthfully. The initial phase is $R^{jk}(0,0)$, where j, k are the messages sent.

Transitions:

(i) *From a regular phase* $R^{jk}(\varepsilon_1, \varepsilon_2)$: if the action of Player 1 (respectively, Player 2) is different from the action determined by the outcome of the randomization device and the action of Player 2 (resp. 1) is equal to the action determined by the outcome of the randomization device, then the next state is $P_1^{k'}$ (resp. $P_2^{j'}$), where k' (resp. j') is the message sent at the end of the period by the corresponding player. [Observe that the message of the deviator plays no role here.] Otherwise, (i) if $(j', k') = (j, k)$ or both $j \neq j'$and $k \neq k'$, the next state is $R^{j'k'}(\varepsilon_1, \varepsilon_2)$, where (j', k') is the pair of messages in the period, (ii) if $j \neq j'$and $k = k'$ (resp. $j = j'$and $k \neq k'$), the next state is $R^{j'k'}(-\bar{\varepsilon}, \varepsilon_2)$ (resp. $R^{j'k'}(\varepsilon_1, -\bar{\varepsilon})$).

(ii) *From a punishment phase*: without loss of generality, consider P_1^k. All statements to histories here refer to the partial history that starts with the beginning of the punishment phase. Given $\widehat{s}_2^k$, define $H^k \subseteq H^T$ as the set of histories of length at most T for which there exists an (arbitrary) strategy s_1 of Player 1 such that this history is on the equilibrium path for s_1 and $\widehat{s}_2^k$, as far as actions are concerned. That is, a history is not in H^k if at some point during the punishment phase the action of Player 2 is inconsistent with $\widehat{s}_2^k$.

As soon as the history $h \in H^T$ is not in H^k, the punishment phase stops and the next state is $P_2^{j'}$, where j' is Player 1's announcement in the last period of the punishment phase. If $h \in H^k$, the punishment phase continues up to the T-th period, and we let henceforth h denote such a history of length T. Let (j', k') denote the pair of messages in the last period of the punishment phase.

The next state is then $R^{j'k'}\left(\varepsilon_1\left(h; P_1^k\right), \varepsilon_2\left(h; P_1^k\right)\right)$, with $\varepsilon_1\left(h; P_1^k\right) \in [-\bar{\varepsilon}, 0]$, $\varepsilon_1\left(h; P_1^k\right) = -\bar{\varepsilon}$ if $k' = k$, and (6): $\varepsilon_1\left(h; P_1^k\right)$ is such that, if $k' \neq k$, Player 1 is indifferent between $s_1^{j'k}$ and playing a best-reply to $\widehat{s}_2^k$ assuming that the state of the world is (j', k') along every history $h \in H^k$ within the punishment phase (recall that h specifies (j', k')).[2] Inequality (4) guarantees that the variation of $\varepsilon_1\left(h; P_1^k\right)$ across histories h that is required is less than $2\bar{\varepsilon}/3$, so that this can be done with $\varepsilon_1\left(h; P_1^k\right)$ in $[-\bar{\varepsilon}, 0]$ for all histories h. As for $\varepsilon_2\left(h; P_1^k\right)$, it is in $[\bar{\varepsilon}/3, \bar{\varepsilon}]$ if $k' = k$, and it is in $[-\bar{\varepsilon}, -\bar{\varepsilon}/3]$ otherwise; further, (7): $\varepsilon_2\left(\cdot; P_1^k\right)$ is such that, conditional on state (j', k') and after every history $h' \in H^k$ within the punishment phase, Player 2 is indifferent over all pure continuation strategies (within the punishment phase) consistent with H^k, and prefer those to all others; given (4), this is possible whether $k' = k$ or not.

[2] See Hörner and Olszewski for the details of such a specification.

It is clear that the strategy profile yields the pair of payoff arrays $v=(v_1, v_2)$. It is equally clear that play is specified in a way that is independent of beliefs.

Verification that the described strategy profile is a Perfect Bayesian Equilibrium[3]

Regular phase $R^{jk}(\varepsilon_1, \varepsilon_2)$: (i) actions: suppose that one player, say Player 1, unilaterally deviates from the action profile determined by the public randomization device, for some realization of this device. Then the punishment phase $P_1^{k'}$ starts, where k' is the announcement by Player 2. Accordingly, the payoff from deviating is at most equal to the right-hand side of (4), while the payoff from playing according to the strategy profile is at least the left-hand side of (4). The result follows. (ii) messages: (a) assume first that Player 1 has deviated from the recommended action profile, while Player 2 has not. Because Player 2 will correctly report the column k at the end of the punishment phase $P_1^{k'}$ that starts, he will get at most $\left(1-\delta^T\right) M + \delta^T \left(v_i^{jk} - \bar{\varepsilon}/3\right)$ by announcing $k' \neq k$, while he gets at least $-\left(1-\delta^T\right) M + \delta^T \left(v_i^{jk} + \bar{\varepsilon}/3\right)$ if he announces $k' = k$, so that Player 2 has a strict incentive to report truthfully given (5). Given that Player 1 has deviated, Player 1's message plays no role in future play, and so it is also optimal for Player 1 to report truthfully; (b) otherwise, if Player i (say Player 2) reports the true state he gets at least $v_i^{jk} - \bar{\varepsilon}$, while if he misreports, he gets at most $(1-\delta) \max_{k'} v_i^{jk}\left(R^{jk'}(\varepsilon_1, \bar{\varepsilon})\right) + \delta \left(v_i^{jk} - \bar{\varepsilon}\right)$. Therefore, (3) guarantees that neither player has an incentive to deviate.

Punishment phase: without loss of generality, consider P_1^k. (i) messages: Observe first that all the messages in the punishment phase are irrelevant, except in the last period of this punishment phase, whether this occurs after T periods or before. If such a history belongs to H^k, then truthful announcements are optimal because of (3), as in case (ii-b) above; if such a history does not belong to H^k, then truthful announcements are also optimal as the situation is identical to the one described just above (case (ii-a)). (ii) actions: the inequality corresponding to (4) for Player 2 guarantees that he has no incentive to take an action outside of the support of the (possibly mixed) action specified by $\widehat{s}_2^k$ after every history $h \in H^k$; within this support, (7) guarantees that he is indifferent over all the actions (whether k is the true column or not); as for Player 1, by definition his strategy is optimal in case k is the true column, and (6) guarantees that it remains optimal to play according to s_1^{jk} in state of the world (j, k'), for all j, k'.

4 An example from Ely and Välimäki (2004)

Consider the example of Ely and Välimäki (2004) with two-long run players. Player 1 is informed or the row, his *type* G or B, at the beginning of the game. In every period, there are two possible states of the world, not to be confused with the row: these states, θ_e and θ_t are realizations of random variables drawn independently and identically over time. In every period, both states are

[3] Given that the public randomization device is not finitely-valued, sequential equilibrium is not well-defined.

equally likely. The realizations are observed by Player 1, but not by Player 2. Player 1 has two actions, e and t, which stand for *engine replacement* and *tune-up*, respectively. In every period, Player 2 can choose to stay out, in which case both players get a payoff of zero, independently of the state and row, or trade, in which case the payoff depends both on the action of Player 1 and the state of the world, according to the following matrices, for some $w > u > 0$:

Type G:

	θ_e	θ_t
e	(u, u)	$(-w, -w)$
t	$(-w, -w)$	(u, u)

Type B:

	θ_e	θ_t
e	(u, u)	$(u, -w)$
t	$(-w, -w)$	$(-w, u)$

That is, Player 1's type G and Player 2 have the same preferences: matching action and state, while Player 2's type B prefers one action to the other independently of the state.

We can easily adapt our proof to encompass such a set-up. We restrict attention to the case in which trade takes place in virtually all periods. Our purpose is to study which payoffs are belief-free equilibria *with respect to* the type of Player 1. We have in mind a situation in which there is sufficient statistical evidence for Player 2 to treat the law of the *i.i.d.* state as objective uncertainty. In other applications, it may make more sense to require that the equilibrium be belief-free with respect to the evolving state (or both). In this application, there is no belief-free equilibrium payoff if we insist that the restriction be relative to both kinds of uncertainties simultaneously, as long as $w > u$.

We define two probabilities π^B and π^G, that correspond (approximately) to the fraction of time Player 1's bad type and good type perform engine replacements (Player 2 trades almost always, as we will argue). Incentive compatibility requires that:

$$V_1^G := u - \left|\pi^G - \frac{1}{2}\right|(u+w) > u - \left|\pi^B - \frac{1}{2}\right|(u+w) \text{ (Player 1's type } G)$$
$$V_1^B := \pi^B(u+w) - w > \pi^G(u+w) - w \text{ (Player 1's type } B).$$

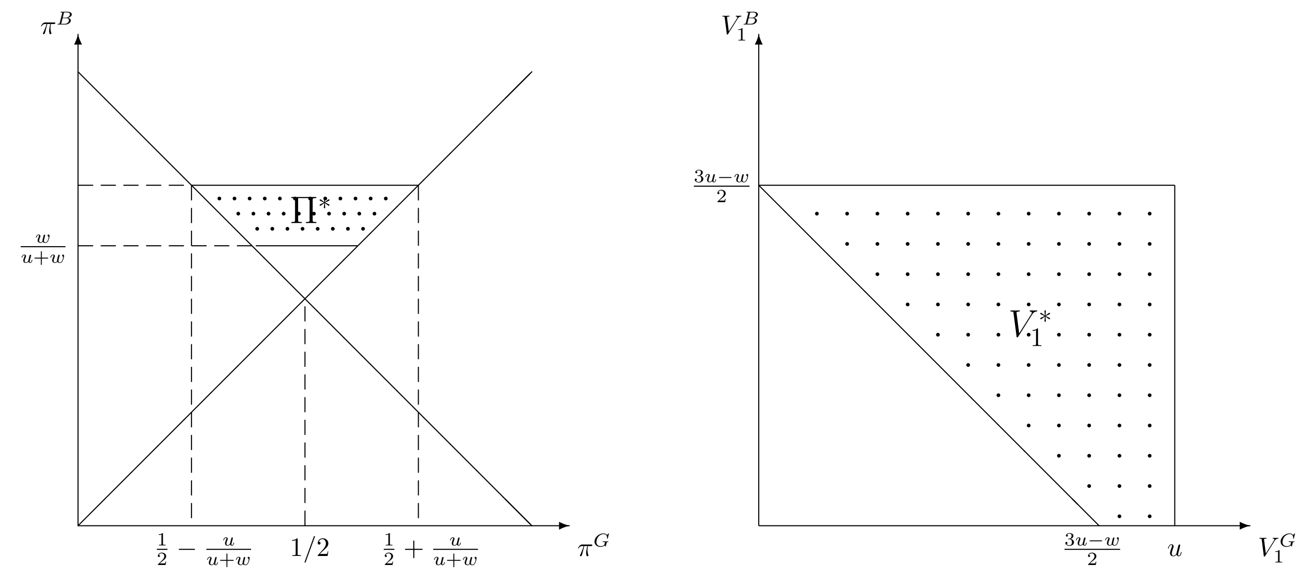
Assuming that Player 1's bad type matches action and state whenever possible given π^B (as does the good type), individual rationality further requires:

$$u - \left|\pi^B - \frac{1}{2}\right|(u+w) > 0 \text{ and } \pi^B(u+w) - w > 0.$$

The first inequality guarantees that it is individually rational for Player 2 to trade with a Player 1's bad type, and implies (given incentive compatibility) that it is also individually rational for him to do so with a good type (and that it is also individually rational for Player 1's good type to follow his equilibrium strategy). The second inequality guarantees that it is individually rational for Player 1's bad type to follow his equilibrium strategy. These inequalities reduce to:

$$\frac{1}{2}-\frac{u}{u+w}<\pi^G,\pi^B<\frac{1}{2}+\frac{u}{u+w},\ \text{and } \pi^B>\max\left\{\pi^G,1-\pi^G\right\}.$$

The corresponding set $\Pi^*$of probabilities $\left\{\pi^G,\pi^B\right\}$ is non-empty if and only if $3u>w$, which we assume from now on (observe that in this example, non-existence in the case $3u>w$ is not driven by individual rationality or incentive compatibility *per se*, but by their conjunction). It is then immediate to characterize the belief-free equilibrium payoffs. The probabilities and payoffs are represented in the Figure below.



The reader can probably guess how a full construction in this example (which does not fit the assumptions of the theorem, because of the changing state) would go, along the lines of the construction underpinning the proof (we omit the details): the play is divided in phases of length T, at the beginning of which Player 1 signals his type, and in which he is then supposed to replace engine a number of times equal (to the nearest integer close) to $\pi^B T$, or $\pi^G T$, depending on the signal he sent. If he does so in a way that matches the state of the world as often as possible given this constraint, then indeed the fraction of mismatches within such a phase will tend to $\left|\pi^B-1/2\right|$ or $\left|\pi^G-1/2\right|$ as $T\to\infty$. In order to guarantee that Player 1's bad type is indeed willing to match the state of the world rather than replace engines as soon and as often as he is allowed within a phase, it is necessary that he be punished at the end of the phase by an amount proportional to the timing of the observed replacement: as $\delta \to 1$, the maximal necessary punishment tends to zero. To enforce the punishment, players can agree, using a public randomization device, on a period in which no trade takes place, at the end of the phase, in a way that gives exactly the right punishment (the device can be dispensed with). It is also clear how each player can secure zero and drive down his opponent's payoff to zero, so there is no need to elaborate on the way punishments for observable deviations are enforced.

5 Conclusion

We have studied belief-free equilibria in two-player repeated games with two-sided incomplete information under discounting. In a belief-free equilibrium, players' strategies are best-replies after every history regardless of a player's belief about his opponent's type. Hence, these equilibria are robust to all specifications of prior beliefs and updating rules. We show that, when players are sufficiently patient, any payoff that is (strictly) incentive compatible and individually rational can by achieved with a belief free equilibrium. Conversely, the payoff in a belief free equilibrium must be incentive compatible and individually rational. One question that remains open is a precise characterization of those games for which belief-free equilibria do not exist. We provide examples to show how to actually determine these payoff sets, and describe an extension to a game with *i.i.d.* shocks.

Bibliography:

Aumann R.J. and M.B. Maschler, 1995. *Repeated Games with Incomplete Information*, The MIT Press.

Blackwell, D., 1956. "An analog of the minmax theorem for vector payoffs," *Pacific Journal of Mathematics*, **6**, 1–8.

Crémer, J. and R. McLean, 1985. "Optimal Selling Strategies under Uncertainty for a Discriminating Monopolist when Demands are Interdependent," *Econometrica*, **53** (2), 345-361.

Cripps, M. and J. Thomas, 2003. "Some Asymptotic Results in Discounted Repeated Games of One-Sided Incomplete Information," *Mathematics of Operations Research*, **28** (3), 433-462.

Ely, J. and J. Välimäki, 2003. "Bad reputation," *Quarterly Journal of Economics*, **118**, 785-814

Ely, J. and J. Välimäki, 2002. "A Robust Folk Theorem for the Prisoner's Dilemma," *Journal of Economic Theory*, **102**, 84–105.

Ely, J., J. Hörner and W. Olszewski, 2005. "Belief-free Equilibria in Repeated Games," *Econometrica*, **73**, 377-415.

Hörner, J. and W. Olszewski, 2006. "The Folk Theorem with Almost-Perfect Monitoring," *Econometrica*, forthcoming.

Forges, F. 1992. "Non-zero-sum repeated games of incomplete information". R. J. Aumann, S. Hart, eds. *Handbook of Game Theory*, Vol. 1. North Holland, Amsterdam, The Netherlands.

Forges, F. and E. Minelli, 1997. "A Property of Nash Equilibria in Repeated Games with Incomplete Information," *Games and Economic Behavior*, **18**, 159-175.

Harsanyi, J.C., 1967-1968. "Games with incomplete information played by Bayesian players," *Management Science* **14**, 159-182, 320-334, 486-502.

Hart, S., 1985. "Nonzero-sum two-person repeated games with incomplete information," *Mathematics of Operations Research* **10**, 117–153.

Kalai, E., 2004. "Large Robust Games," *Econometrica*, **72** (6), 1631-1665.

Koren, G. 1988. *Two-Person Repeated Games with Incomplete Information and Observable Payoffs.* M.Sc. thesis, Tel Aviv University, Tel Aviv, Israel.

Luce, D. and H. Raiffa, 1957. *Games and Decisions.* Wiley.

von Neumann, J. and O. Morgenstern, 1944. *Theory of Games and Economic Behavior.* Princeton University Press.

Shalev, J. 1994. "Nonzero-sum two-person repeated games with incomplete information and known-own payoffs," *Games and Economic Behavior* **7**, 246–259.

Appendix: Proof of the main theorem

We first explain the construction without explicit communication, but with a randomization device. Naturally, communication is replaced by choices of actions, but since the set of actions may be more limited than the set of states (j,k), it is typically necessary to use several periods to 'communicate' a state. We let $c-1$ denote the smallest such number given the number of states and actions $(c-1 \leq \ln_2(\max\{J,K\})+1)$. Players will regularly communicate their private information in rounds of c periods. In the last of these c periods, players have the opportunity, through the choice of a specific action, to communicate that the report they have just sent is incorrect.

Equilibrium Strategies:

The play can again by divided in phases. To guarantee that players' best-replies are independent of their beliefs, even within a round of communication (especially if a player's own deviation during that round already prevents him from truthfully reporting his information), the construction must be considerably refined. For each player, we pick two specific actions from A_i, henceforth referred to as B and U. The pair of payoff arrays $v \in V^*$ is fixed throughout.

There are two kinds of phases: regular phases last at most n periods, and punishments phases, that last at most T periods, where n, T are to be specified. Regular phases are denoted $R^{jk}(\varepsilon_1, \varepsilon_2)$, where $j \in \{1,\ldots,J\}$, $k \in \{1,\ldots,K\}$, or R^{xy}, where $x \in \{1,\ldots,J,(L,n_1^U)\}$ and $x \in \{1,\ldots,K,(L,n_2^U)\}$, with $n_i^U \in \{1,\ldots,c\}$, and either $x=(L,n_1^U)$, or $y=(L,n_2^U)$, or both (L stands for "Lie"). Punishment phases are denoted P_i, $i=1,2$. We let $\widehat{s}_2^k$ (resp. $\widehat{s}_1^j$) denote a (behavior) strategy of Player 2 (resp. 1) such that Player 1's (resp. Player 2's) payoff be less than $v_1^{jk}-3\bar{\varepsilon}$ for all j and all strategies of Player 1 (resp. $v_2^{jk}-3\bar{\varepsilon}$ for all k and all strategies of Player 2), for $\bar{\varepsilon}$ small enough to be specified. Such strategies exists since $v \in V^*$. Further, we let s_1^{jk} (resp. s_2^{jk}) denote a fixed, pure-strategy best-reply to $\widehat{s}_2^k$ (resp. $\widehat{s}_1^j$) given row j (resp. column k).

In several places of the construction, a *communication round* of c periods takes place (within a phase). The integer c is chosen to be the smallest integer such that both $|A_1|^{c-1} \geq J > |A_1|^{c-2}$ and $|A_2|^{c-1} \geq K > |A_2|^{c-2}$. We fix a mapping from states J to $|J|$ sequences $\{a_1^t\}_{t=1}^{c-1}$ of length $c-1$ ($a_1^t \in A_1$) and similarly a mapping from states K to $|K|$ sequences $\{a_2^t\}_{t=1}^{c-1}$ of length $c-1$ ($a_2^t \in A_2$). If the play of Player 1 during the first $c-1$ periods equals such a sequence, and his action in period c equals B, we say that Player 1 (or his play) communicates the corresponding row j. Similarly, if the play of Player 2 during the first $c-1$ periods equals such a sequence, and his action in period c equals B, we say that Player 2 (or his play) communicates the corresponding column k. Otherwise, we say that Player i (or his play) communicates (L,n_i^U), where U is the number of periods during these c periods in which Player i chose action U. We shall provide incentives for Player i to rather report the true row or column, rather than communicate (L,n_i^U) for any n_i^U, and prefer communicating any such (L,n_i^U) rather than the incorrect row or column.

Further, we provide incentives for Player i to maximize the number n_i^U once his play does not coincide with any of the aforementioned particular sequences that communicate rows or columns (to avoid having his beliefs enter his choice of action after such out-of-equilibrium histories).

Actions:

(i) *Regular phase*: A regular phase lasts at most $n > c$ periods, the last c of which is a communication round. During the first $n - c$ periods, play proceed as follows, for all regular phases indexed by j, k and true column k':

Phase:	Player 1	Player 2
$R^{j\left(L,n_2^U\right)}$	s_1^j	$s_2^{jk'}$
$R^{jk}(\varepsilon_1,\varepsilon_2)$	$a_1^{jk}(\varepsilon_1,\varepsilon_2)$	$a_2^{jk}(\varepsilon_1,\varepsilon_2)$
$R^{\left(L,n_1^U\right)\left(L,n_2^U\right)}$	$(U,\ldots,U)$	$(U,\ldots,U)$

The specification for $R^{\left(L,n_1^U\right)k}$ is the obvious analogue to the case $R^{j\left(L,n_2^U\right)}$. The action $a^{jk}(\varepsilon_1,\varepsilon_2)$ is to be specified. Observe that the expression $s_2^{jk'}$ refers to the regular phase through j, which need not be the true row, while k' refers to the true column. This specification of actions is valid as long as (in case of $R^{jk}(\varepsilon_1,\varepsilon_2)$ or $R^{\left(L,n_1^U\right)\left(L,n_2^U\right)}$) the history within the phase is consistent with these actions, or if all deviations from the specified actions during this phase were simultaneous, and as long as (in case of $R^{j\left(L,n_2^U\right)}$) the history within the phase is consistent with $s_1^j$for some arbitrary s_2: as will be specified, a punishment phase is immediately entered otherwise. During the periods $n-c,\ldots,n-1$ of this phase, Player 1 (resp. Player 2) communicates the true row j (resp. true column k); if his play since period $n-c$ makes this impossible, he chooses U in every remaining period.

(ii) *Punishment phase*: Without loss of generality, consider P_1, where $T > 2c$ is to be specified. In the first c periods of this phase, Player 1 plays U repeatedly while Player 2 communicates the true column. As in the regular phase, if Player 2's previous action makes this impossible, he chooses U in every remaining period of this communication round. In the table below, we refer to the case in which the column communicated is k as the case k, while $\left(L,\left(n_1^U,n_2^U\right)\right)$ refers to any other case, where n_i^U is the number of times Player i chose action U in periods $1,\ldots,c$. Play in periods $c+1,\ldots,T-c$ is then as follows:

Phase P_1^T	Player 1	Player 2
k	$s_1^{j'k}$	$\widehat{s}_2^k$
$\left(L,\left(n_1^U,n_2^U\right)\right)$	U	U

This specification is valid (up to period $T-c$) as long as (in case $\left(L,\left(n_1^U,n_2^U\right)\right)$) both players have played U in all periods since period $c+1$ or all deviations have been simultaneous, or (in case k) as long as the history since period $c+1$ is consistent with $\widehat{s}_2^k$ for some strategy s_1, for otherwise a punishment phase is immediately entered. Here, j' refers to the true row privately known to Player 1.

In the last c periods of a punishment phase, a communication round takes place, i.e. players communicate the true row and column, and as soon as they fail to do so, play U repeatedly.

Initial Phase: In the first c periods of the game, a communication round takes place, i.e. players communicate the true row and column, and as soon as they fail to do so, play U repeatedly. In period $c+1$, the regular phase $R^{jk}(\varepsilon_1, \varepsilon_2)$ is entered if row j and column k are communicated, where $\varepsilon_i \in [-\bar{\varepsilon}, \bar{\varepsilon}]$ is chosen so that the ex ante payoff in period 1 is exactly v^{jk} conditional on j and k being the true row and column. If Player 1 communicates j and Player 2 communicates $\left(L, n_2^U\right)$ in the first c periods, the regular phase $R^{j\left(L,n_2^U\right)}$ is entered. Similarly, if Player 1 communicates $\left(L, n_1^U\right)$, whereas Player 2 communicates k, regular phase $R^{\left(L,n_2^U\right)k}$ is entered. Regular phase $R^{\left(L,n_1^U\right),\left(L,n_2^U\right)}$ is entered in the remaining case.

Transitions:

From a regular phase: We have already mentioned what happens if there is a deviation during the first $n-c$ periods of such a phase: if a player makes a unilateral deviation during the first $n-c$ periods of a regular phase $R^{jk}(\varepsilon_1, \varepsilon_2)$ or $R^{\left(L,n_1^U\right)\left(L,n_2^U\right)}$, a punishment phase starts: if Player 1 (Player 2) unilaterally deviates, punishment phase P_1 (resp. P_2) is immediately entered. Similarly, if Player 1 (resp. Player 2) deviates from s_1^j (resp. s_2^k) during the first $n-c$ periods of a regular phase $R^{j\left(L,n_2^U\right)}$ (resp. $R^{\left(L,n_2^U\right)k}$), the punishment phase P_1 (resp. P_2) is immediately entered. From now on, we assume without repeating it that no such deviation occurs.

(i) from $R^{jk}(\varepsilon_1, \varepsilon_2)$: the new phase depends on the last c periods of the phase. Define also $\rho := 2(1-\delta)\delta^{-T}M$. The quantity $\tilde{\varepsilon}_i^{jk}$ will be defined shortly. In all tables that follow, $j' \neq j$, $k' \neq k$. We have:

Regular Phase:	During periods $n-c, \ldots, n$ of the phase, Player 1 and 2 communicate:	Next Regular Phase:
$R^{jk}(\varepsilon_1, \varepsilon_2)$	$\left(L, n_1^U\right), \left(L, n_2^U\right)$	$R^{\left(L,n_1^U\right)\left(L,n_2^U\right)}$
$R^{jk}(\varepsilon_1, \varepsilon_2)$	$\left(L, n_2^U\right), k$	$R^{\left(L,n_2^U\right)k}$
$R^{jk}(\varepsilon_1, \varepsilon_2)$	$\left(L, n_2^U\right), k'$	$R^{\left(L,n_2^U\right)k'}$
$R^{jk}(\varepsilon_1, \varepsilon_2)$	j, k'	$R^{jk'}(\varepsilon_1, -\bar{\varepsilon})$
$R^{jk}(\varepsilon_1, \varepsilon_2)$	j', k'	$R^{j'k'}(\varepsilon_1, \varepsilon_2)$
$R^{jk}(\varepsilon_1, \varepsilon_2)$	j, k	$R^{jk}(\varepsilon_1, \varepsilon_2)$

(ii) from $R^{\left(L,n_1^U\right)k}$:

Regular Phase:	During periods $n-c,\ldots,n$ of the phase, Player 1 and 2 communicate:	Next Regular Phase:
$R^{\left(L,n_1^U\right)k}$	$\left(L,n_1'^U\right),\left(L,n_2'^U\right)$	$R^{\left(L,n_1'^U\right)\left(L,n_2'^U\right)}$
$R^{\left(L,n_1^U\right)k}$	$\left(L,n_1'^U\right),k$	$R^{\left(L,n_1'^U\right)k}$
$R^{\left(L,n_1^U\right)k}$	$\left(L,n_1^U\right),k'$	$R^{\left(L,n_1'^U\right)k'}$
$R^{\left(L,n_1^U\right)k}$	$j,\left(L,n_2'^U\right)$	$R^{j\left(L,n_2'^U\right)}$
$R^{\left(L,n_1^U\right)k}$	j,k	$R^{jk}\left(\tilde{\varepsilon}_1^{jk}+\rho n_1^U,\bar{\varepsilon}\right)$
$R^{\left(L,n_1^U\right)k}$	j,k'	$R^{jk'}\left(\tilde{\varepsilon}_1^{jk'}+\rho n_1^U,-\bar{\varepsilon}\right)$

and symmetrically from $R^{j\left(L,n_2^U\right)}$;

(iii) finally, from $R^{\left(L,n_1^U\right)\left(L,n_2^U\right)}$:

Regular Phase:	During periods $n-c,\ldots,n$ of the phase, Player 1 and 2 communicate:	Next Regular Phase:
$R^{\left(L,n_1^U\right)\left(L,n_2^U\right)}$	$\left(L,n_1'^U\right),\left(L,n_2'^U\right)$	$R^{\left(L,n_1'^U\right)\left(L,n_2'^U\right)}$
$R^{\left(L,n_1^U\right)\left(L,n_2^U\right)}$	$\left(L,n_1'^U\right),k$	$R^{\left(L,n_1'^U\right)k}$
$R^{\left(L,n_1^U\right)\left(L,n_2^U\right)}$	j,k	$R^{jk}\left(\rho n_1^U,\rho n_2^U\right)$

From a punishment phase: Without loss of generality, consider P_1. We have already mentioned what happens if there is a deviation during the periods $c+1,\ldots,T-c$ of such a phase. In case $\left(L,\left(n_1^U,n_2^U\right)\right)$, if Player i unilaterally deviates from the play of U, the punishment phase P_i is immediately entered. In case k, if Player 2 deviates from the support of the (possibly mixed) action specified by $\widehat{s}_2^k$, punishment phase P_2 is entered (no matter how Player 1 has played). From now on, we assume without repeating it that no such deviation occurs. In case k, let h denote the history during the periods $c+1,\ldots,T-c$.

(i) In case k:

Punishment Phase P_1:	During periods $T-c,\ldots,T$ of the phase, Player 1 and 2 communicate:	Next Regular Phase:
case k	$\left(L,n_1'^U\right),\left(L,n_2'^U\right)$	$R^{\left(L,n_1'^U\right)\left(L,n_2'^U\right)}$
case k	$\left(L,n_1'^U\right),k$	$R^{\left(L,n_1'^U\right)k}$
case k	$\left(L,n_1'^U\right),k'$	$R^{\left(L,n_1'^U\right)k'}$
case k	j,k	$R^{jk}\left(\rho n_1^U-\bar{\varepsilon},\varepsilon_2^{k;k}(h)\right)$
case k	j,k'	$R^{jk'}\left(\varepsilon_1^{k;k'}(h),\varepsilon_2^{k;k'}(h)\right)$
case k	$j,\left(L,n_2'^U\right)$	$R^{j\left(L,n_2'^U\right)}$

where $\varepsilon_2^{k;k}(h) \in [3\bar{\varepsilon}/4, \bar{\varepsilon}]$, $\varepsilon_2^{k;k'}(h) \in [-\bar{\varepsilon}/2, -\bar{\varepsilon}/4]$, and $\varepsilon_1^{k;k'}(h) \in [-\bar{\varepsilon}, \bar{\varepsilon}]$ are computed as follows: $\varepsilon_2^{k;k}(\cdot)$ makes Player 2 precisely indifferent over all histories h that are consistent with $\widehat{s}_2^k$, conditional on the true column being k; $\varepsilon_2^{k;k'}(\cdot)$ makes Player 2 precisely indifferent over all histories h that are consistent with $\widehat{s}_2^k$, conditional on the true column being k'; finally, $\varepsilon_1^{k;k'}(h)$ compensates Player 1 for every period along h in which the action he took is the action specified by s_1^{jk}, so as to make sure that playing this action is optimal, conditional on the true state being (j, k') (communicated in the last c periods).

(ii) In case $\left(L, \left(n_1^U, n_2^U\right)\right)$:

Punishment Phase P_1:	During periods $T-c, \ldots, T$ of the phase, Player 1 and 2 communicate:	Next Regular Phase:
case $\left(L, \left(n_1^U, n_2^U\right)\right)$	$\left(L, n_1'^U\right), \left(L, n_2'^U\right)$	$R^{\left(L, n_1'^U\right)\left(L, n_2'^U\right)}$
case $\left(L, \left(n_1^U, n_2^U\right)\right)$	$\left(L, n_1'^U\right), k$	$R^{\left(L, n_1'^U\right)k}$
case $\left(L, \left(n_1^U, n_2^U\right)\right)$	$j, \left(L, n_2'^U\right)$	$R^{j\left(L, n_2'^U\right)}$
case $\left(L, \left(n_1^U, n_2^U\right)\right)$	j, k	$R^{jk}\left(\rho n_1^U, \rho n_2^U\right)$

It is clear from this specification that the strategy profile described here is belief-free, since actions are always determined by the history and possibly by a player's own private information (in case he is minmaxed), but not on his beliefs about his opponent's information.

Specification of $\bar{\varepsilon}$, $a_1^{jk}(\varepsilon_1, \varepsilon_2)$, δ, T, n, $\tilde{\varepsilon}_i^{jk}$:

Since $v \in V^*$, it is possible to find $\bar{\varepsilon} > 0$, for all $(\varepsilon_1, \varepsilon_2), (\varepsilon_1', \varepsilon_2') \in [-2\bar{\varepsilon}, 2\bar{\varepsilon}]$, there exists probability distributions over A, $\Pr\left\{\cdot \mid R^{jk}(\varepsilon_1, \varepsilon_2)\right\}$ such that for all j, k, j', k', and $i = 1, 2$, defining:

$$v_i^{jk}\left(R^{j',k'}(\varepsilon_1, \varepsilon_2)\right) := \sum_{a \in A} \Pr\left\{a \mid R^{j'k'}(\varepsilon_1, \varepsilon_2)\right\} u_i^{jk}(a),$$

it is the case that, for $j' \neq j$, $k' \neq k$,

$$v_1^{jk}\left(R^{j,k}(\varepsilon_1, \varepsilon_2)\right) > v_1^{jk}\left(R^{j'k}(\varepsilon_1', \varepsilon_2')\right) \text{ and } v_2^{jk}\left(R^{j,k}(\varepsilon_1, \varepsilon_2)\right) > v_2^{jk}\left(R^{jk'}(\varepsilon_1', \varepsilon_2')\right); \qquad \text{(1A)}$$

further, if $\{a_1^t\}_{t=1}^c$, $\{a_2^t\}_{t=1}^c$ is the sequence that communicates j and k, for all δ close enough to one and n large enough, we can pick those distributions so that Player i's average discounted payoff under state (j, k) from the sequence $\{a_1^t, a_2^t\}_{t=1}^c$ followed by n repetitions of the action profile determined by $\Pr\left\{a \mid R^{jk}(\varepsilon_1, \varepsilon_2)\right\}$ is exactly equal to $v_i^{jk} + \varepsilon_i$. Observe that in the equilibrium described above, all values of ε_i are in $[-\bar{\varepsilon}, \bar{\varepsilon}]$. Further, since $v \in V^*$, we may assume that Player 1's (resp. Player 2's) average discounted payoff under state (j, k) given that Player 2 uses $\widehat{s}_2^k$ (resp. $\widehat{s}_1^j$) [that is, given that his opponent minmaxes him in the sense of Blackwell] for $n - 2c$ periods, followed by any arbitrary play during c periods, is at most $v_1^{jk} - 2\bar{\varepsilon}$ (resp. $v_2^{jk} - 2\bar{\varepsilon}$).

Consider the following inequalities:

$$v_1^{jk} + \varepsilon_1 > (1-\delta^c) M + \delta^c (1-\delta^n)(v - 2\bar{\varepsilon}) + \delta^{n+c} \left(v_1^{jk} + \tilde{\varepsilon}_1^{jk} + c\rho \right), \tag{2A}$$

$$v_1^{jk} + \varepsilon_1 < -\left(1-\delta^{n+c}\right) M + \delta^{n+c} \left(v_1^{jk} + \tilde{\varepsilon}_1^{jk} \right), \tag{3A}$$

$$v_1^{jk} - \bar{\varepsilon} > (1-\delta^c) M + \delta^c \left(1-\delta^{n-c}\right) \left(v_1^{jk} - 2\bar{\varepsilon} \right) + \delta^n \left(v_1^{jk} - \bar{\varepsilon} \right). \tag{4A}$$

Given $\bar{\varepsilon}$, fixing δ^n, inequality (4A) is satisfied as $\delta \to 1$, provided that the value of δ^n is large enough. Similarly, given $\bar{\varepsilon}$, fixing δ^n, inequality (2A) is satisfied as $\delta \to 1$ for $\tilde{\varepsilon}_1^{jk} = -\bar{\varepsilon}$, and (3A) is satisfied for $\tilde{\varepsilon}_1^{jk} = 3\bar{\varepsilon}/4$, provided that the value of δ^n is large enough and $\varepsilon_1 < \bar{\varepsilon}/2$ (recall that $\rho = 2(1-\delta)\delta^{-T} M \to 0$ for fixed δ^{-T}). Observe that the left-hand side of (4A) is the lowest possible payoff for Player 1, evaluated in the first period of a communication round concluding either a punishment phase or a regular phase, if he communicates his true row j and Player 2 communicates his true column k, while the right-hand side is the most he can expect by communicating another row $j' \neq j$ and Player 2 communicates his true column k. Similarly, the left-hand side of (2A) and (3A) is Player 1's payoff, evaluated in the first period of a communication round concluding either a punishment phase or a regular phase, if he communicates his true row j and Player 2 communicates his true column k (and the upcoming regular phase is $R^{jk}(\varepsilon_1, \varepsilon_2)$), while the right-hand side of (2A) (resp. 3A) is the highest (resp. lowest) payoff he can expect if he communicates $\left(L, n_1^U\right)$ for some n_1^U. Therefore, if $\varepsilon_1 < \bar{\varepsilon}/2$, by the intermediate value theorem, we can find $\tilde{\varepsilon}_1^{jk} \in (-\bar{\varepsilon}, 3\bar{\varepsilon}/4)$ so that the payoff from communicating the true row exceeds the payoff from communicating $\left(L, n_1^U\right)$ for all n_1^U, which in turn exceeds the payoff from communicating another row $j' \neq j$, provided Player 2 communicates the true column. If $\varepsilon_1 \geq \bar{\varepsilon}/2$, we can set $\tilde{\varepsilon}_1^{jk} = 0$: in that case as well, the same ordering obtains provided that the value of δ^n is large enough as $\delta \to 1$. The values $\tilde{\varepsilon}_2^{jk}$ are defined similarly.

Consider now the two inequalities:

$$-(1-\delta^n) M + \delta^n \left(v_1^{jk} + \bar{\varepsilon} \right) > (1-\delta^n) M + \delta^n \left(v_1^{jk} + \rho c \right), \tag{5A}$$

$$-(1-\delta^n) M + \delta^n v_1^{jk} > (1-\delta^n) M + \delta^n \left(v_1^{jk} - \bar{\varepsilon} \right) \tag{6A}$$

Conditional on Player 2 communicating $\left(L, n_2^U\right)$ for some n_2^U: the left-hand side of (5A) is the lowest possible payoff for Player 1, evaluated in the first period of a communication round concluding either a punishment phase or a regular phase, if he communicates his true row j, while the right-hand side is the highest payoff he can get if he communicates $\left(L, n_1^U\right)$ for some n_1^U; similarly, the left-hand side of (6A) is the lowest possible payoff for Player 1, evaluated in the first period of a communication round concluding either a punishment phase or a regular phase, if he communicates $\left(L, n_1^U\right)$ for some n_1^U, while the right-hand side is the highest payoff he can get if he communicates another row $j' \neq j$. Observe that both inequalities hold, given $\bar{\varepsilon}$, letting $\delta \to 1$, provided δ^n is large enough.

Finally, observe that the choice of ρ trivially ensures that, conditional on having started communicating $\left(L, n_1^U\right)$ for some n_1^U, Player 1 has strict incentives to play U in all remaining periods of the communication round, no matter where this round takes place.

Similar considerations hold for Player 2. To summarize, we have shown that we can guarantee that both players prefer to communicate their true information, in any communication round, than to communicate $\left(L, n_i^U\right)$ for all n_i^U; that, conditional on communicating $\left(L, n_i^U\right)$ for some n_i^U, they have strict incentives to choose U in any remaining period of the communication round; and that they prefer to communicate $\left(L, n_i^U\right)$ for any n_i^U than to communicate an incorrect row or column; all this, provided that δ^n (and δ^T) is fixed but large enough, by taking $\delta \to 1$, given $\bar{\varepsilon}$.

Turning down to actions, we must consider:

$$\left(1-\delta^{n+1}\right) M+\delta^{n+1}\left(1-\delta^{T-n}\right)\left(v_i^{jk}-2\bar{\varepsilon}\right)+\delta^{T+1}\left(v_i^{jk}-\bar{\varepsilon}\right)<-\left(1-\delta^n\right) M+\delta^n\left(v_i^{jk}-\bar{\varepsilon}\right), \tag{7A}$$

$$\left(1-\delta^{n+1}\right) M+\delta^{n+1}\left(1-\delta^{T-n}\right)\left(v_i^{jk}-2\bar{\varepsilon}\right)+\delta^{T+1}\left(v_i^{jk}-\bar{\varepsilon}\right)<-\left(1-\delta^T\right) M+\delta^T\left(v_i^{jk}-\frac{\bar{\varepsilon}}{2}\right), \tag{8A}$$

$$\left(1-\delta^T\right) M+\delta^T\left(v_i^{jk}-\bar{\varepsilon}/2\right)<-\left(1-\delta^T\right) M+\delta^T\left(v_i^{jk}-\bar{\varepsilon}/4\right). \tag{9A}$$

Observe that all three inequalities hold, for both $i=1,2$, given $\bar{\varepsilon}$, for δ^T and n fixed, as $\delta \to 1$. This guarantees that, given $\bar{\varepsilon}$, we can choose n, T, δ to satisfy all the inequalities above. As for the interpretation, (7A) guarantees that Player i does not want to deviate during any regular phase; (8A) that Player i does not want to deviate during the punishment phase P_{-i} and (9A) guarantees that we can pick $\varepsilon_2^{k;k}(\cdot)$ and $\varepsilon_2^{k;k'}(\cdot)$ within a range of values not exceeding $\bar{\varepsilon}/4$ in case k. Indeed, the left-hand side of (7A) and (8A) is the highest payoff Player i can hope for by deviating at any time (outside communication rounds), while the right-hand side of (7A) (resp. (8A)) is the lowest payoff he can expect by sticking to the equilibrium strategies in a regular phase (resp. in a punishment phase). Note that $\left(1-\delta^T\right) M$ is the highest payoff he can get during the punishment phase P_{-i} over all actions consistent with his equilibrium strategy, while $-\left(1-\delta^T\right) M$ is the lowest such payoff; inequality (9A) guarantees therefore that there exists functions $\varepsilon_1^{k;k'}$ and $\varepsilon_1^{k;k}$ whose ranges do not exceed $\bar{\varepsilon}/4$ such that Player 1 is playing a best-reply, given $\varepsilon_1^{k;k'}(\cdot)$, whether or not the true column is k.

It remains to be shown that the public randomization device can be dispensed with. Observe that the public randomization device is only used in one place, namely to get the exact desired payoffs during a regular phase. However, since $v \in V^*$, full-dimensionality hold, and we can use the exact same technique as in Hörner and Olszewski (2006) by having each player randomize at the beginning of the regular phase over two subsets of actions, so that the outcome of this initial randomization pins down one of four possible sequences of actions during this round, and have the values of ε_i being adjusted at the end of the regular phase so as to guarantee that both players are indifferent over all actions in this initial period; the randomization then allows to convexify the payoffs. The details are omitted.

Appendix 2: Individual Rationality in the Leading Example

To prove the result, we need to introduce some notations. For $\alpha_2 \in \triangle A_2$ (viewed henceforth as an element in the unit interval), let

$$U(\alpha_2) = \left\{x = \left(x^1, x^2\right) : x^j = u_1^j(\alpha_1, \alpha_2), \text{ some } \alpha_1 \in \triangle A_1\right\}.$$

where $u_1^j(\alpha_1, \alpha_2)$ is Player 1's expected payoff in the stage game when players randomize their actions according to (α_1, α_2) and the true row, or state, is j. Thus, $U(\alpha_2)$ represents the set of expected payoff vectors that Player 1 can obtain in the two states, given that Player 2 randomizes his action according to α_2 as we vary Player 1's mixed action α_1 (but independent of the row j). Note α_1 and α_2 do not depend on j. Let $F \subset \mathbb{R}^2$ be a compact set. For all $x \notin F$, let $\Pi_F(x)$ denote the set of points in F closest to x. If x and y are two distinct points of $\mathbb{R}^2$, H_{xy} is the line through y perpendicular to the line xy. Blackwell (1968) shows that if F is a closed convex set, F is approachable (for Player 2) if and only if it is a B-set: F is a B-set if for all $x \notin F$, there exists a mixed action α_2 for Player 2 and a point y in $\Pi_F(x)$ such that the hyperplane H_{xy} separates x from $U(\alpha_2)$.

Assume from now on that $F = \{x \in \mathbb{R}^j : x^j \leq v_1^j \text{ for all } j = 1, 2\}$ and let $(v_1^1, v_1^2) =: v$. Consider first the case $G < L$. See Figure 1A. If v lies (weakly) below the segments D_1 (see Figure 1A), then it is always possible to find $x \notin F$ such that $\Pi_F(x) = v$ and H_{xv} is parallel to D_1. Since for all α_2, $U(\alpha_2) \cap D_1$ is non-empty, [in fact, D_1 and D_2 represent the set of extreme points of $U(\alpha_2)$] there is no α_2 such that the line H_{xv} separates x from $U(\alpha_2)$. Hence F is not a B-set. Similarly, if v lies (weakly) below one of the segments D_2 or it has a negative coordinate, F is not approachable as it is always possible to find a point $x \notin F$ such that $\Pi_F(x) = v$ and H_{xv} does not separate D_1 or D_2 from x. Suppose now that v lies above both segments D_1 and D_2 and has strictly positive coordinates. Then all points outside F can be separated from $U(0)$, $U(1)$ or $U(1/2)$ (see Figure 1A). It is straightforward to check that the segments D_1 and D_2 are precisely given by the equations:

$$v_1^2 = \frac{1+G}{1+L}\left(1 - v_1^1\right) \text{ and } 1 - v_1^2 = \frac{1+L}{1+G} v_1^1,$$

for the appropriate ranges of values v_1^1 and v_1^2, giving the desired result.

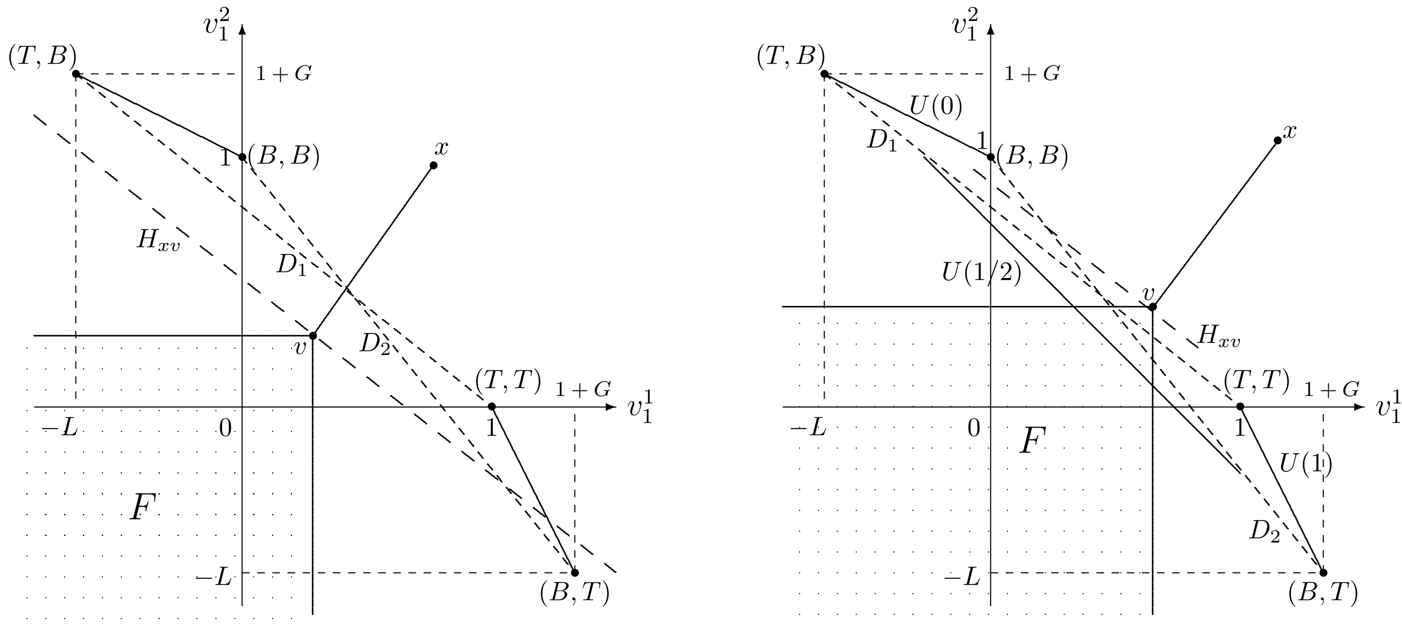


Figure 1A: Case $G < L$

Consider now the case $G > L$. See Figure 2A. The curve C represents the sets of payoffs (v_1^1, v_1^2) such that v_1^2 equals the right hand side of (1). This region corresponds to the set of maximal payoffs that Player 1 can achieve in a one-shot game in which players use mixed (uncorrelated) actions. Namely, a point in C satisfies

$$v_1^2 = \max_{\alpha_1,\alpha_2} u_1^2(\alpha_1, \alpha_2)$$

such that:

$$v_1^1 = u_1^1(\alpha_1, \alpha_2)$$

where $v_1^1 \in [-L, 1+G]$. Hence there is no α_2 such that $U(\alpha_2)$ has a point that lies strictly above C. Moreover, for any α_2, $U(\alpha_2)$ intersects C in one point since the value of α_2 that solves the maximization problem varies between 0 and 1 as v_1^1 varies from $-L$ to $(1+G)$. If v lies (weakly) below the curve C then it is always possible to find $x \notin F$ that lies above C such that $\Pi_F(x) = v$ and H_{xv} lies (weakly) below C. Since for all α_2, $U(\alpha_2) \cap C$ is non-empty, there is no α_2 such that the line H_{xv} separates x from $U(\alpha_2)$. Consider now v above the curve C and such that v has strictly positive coordinates. Then all points outside F can be separated from $U(0)$, $U(1)$ or some intermediate $U(\alpha_2)$ (see Figure 2A).

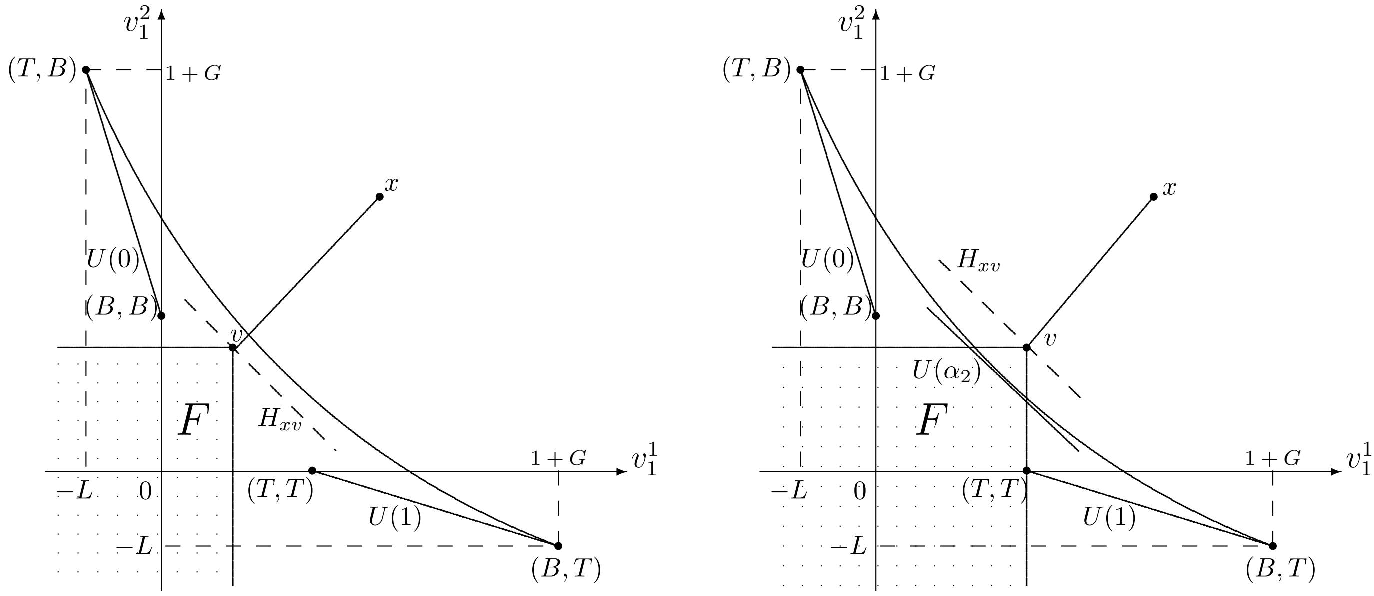


Figure 2A: Case $G > L$