When is the individually rational payoff in a repeated game equal to the minmax payoff?*

Olivier Gossner and Johannes Hörner

November 21, 2006

Abstract

We study the relationship between a player's (stage game) minmax payoff and the individually rational payoff in repeated games with imperfect monitoring. We characterize the signal structures under which these two payoffs coincide for any payoff matrix. Under a full rank assumption, we further show that, if the monitoring structure of an infinitely repeated game 'nearly' satisfies this condition, then these two payoffs are approximately equal, independently of the discount factor. This provides conditions under which existing folk theorems exactly characterize the limiting payoff set.

1 Introduction

Folk theorems aim at characterizing the entire set of payoff vectors that can be attained at equilibrium in repeated games. While some of the early literature focused on Nash reversion (see [Fri71]), it was recognized early on that, under perfect monitoring, players could be held to a perhaps less intuitive but often strictly lower payoff, the minmax payoff, defined to be

$$\min_{\alpha_{-i} \in \otimes_{j \neq i} \triangle A_j} \max_{a_i \in A_i} g_i \left(a_i, \alpha_{-i}\right),$$

where $a_i \in A_i$ are player i's pure actions, $\alpha_j \in \triangle A_j$ is player j's mixed action and g_i is player i's payoff function.[1] That is, the minmax payoff is the lowest payoff player i's opponents can hold him to by any choice α_{-i} of *independent* actions, provided that player i correctly foresees α_{-i} and plays a best-response to it.

When the folk theorem was later extended to imperfect public monitoring (see [FLM94]), the minmax payoff thus appeared as a natural bound to focus on. Yet it is known that, unlike in the case of perfect monitoring, there are games with imperfect public monitoring in which some player's individually rational payoff is strictly lower than his minmax payoff. The main purpose of this paper is to characterize the (public or private) monitoring structures for which such phenomena occur.

*We thank Tristan Tomala and Jonathan Weinstein for very useful discussions, as well as audiences at Brown University, the University of Pennsylvania and Yale University

[1] See, among others, [AS76], [FM86], [Rub77]

The minmax payoff is certainly an upper bound on a player's individually rational payoff (the lowest payoff he can be held down to), as his opponents can always choose to play the minimizing mixed strategy profile in every period. Yet in games with imperfect monitoring, private histories offer some scope for correlation among a player's opponents, and his individually rational payoff can sometimes be lower, in the event that player i's correlated minmax payoff is strictly below his minmax payoff (the definition of correlated minmax payoff is obtained by replacing $\otimes_{j \neq i} \triangle A_j$ by $\triangle \otimes_{j \neq i} A_j$ as the domain of the minimization in the definition of the minmax payoff).

There are important known special cases in which the identification of individually rational payoff and minmax payoff is warranted:

- if there are only two players, as correlated minmax and minmax payoffs then coincide;
- if monitoring is perfect, as all players then hold the same information at any point in time, so that the probability distribution over player i's opponents' actions given his information corresponds to independent randomizations by his opponents;
- if monitoring is public, but information is semi-standard ([Leh90]);
- if monitoring is public, but attention is restricted to public strategies, as in this case as well the information relevant to forecasting future play is commonly known.

However, as mentioned, examples of games are known in which a player's equilibrium payoff can be strictly below his minmax payoff. *A fortiori*, the same phenomenon arises in repeated games with private monitoring, an active area of research in game theory (see [Kan02] for a survey). In fact, we provide simple examples to show that this is even possible when:

- monitoring is almost-perfect;
- the punished player perfectly monitors his opponents;
- monitoring is public (see also [FT91], exercise 5.10).

In general, individually rational payoffs depend both on the details of the monitoring structure and the payoff functions, and computing them is a daunting task. The present paper identifies the general condition on the monitoring structure under which the minmax payoff is indeed the players' individually rational payoff. This condition is trivially satisfied in the cases listed above and precisely characterizes those monitoring structures for which, after every history, each player can view his opponents' actions as independent randomizations, conditional on his information. Roughly speaking, what is needed is that, for all mixed action profiles, the distribution over player i's opponents' 'signals' be independent, conditional on player i's 'signal' or a garbling thereof, where a player's 'signal' includes both his own action and the actual signal he observed.

Because we would like to identify the condition on the signal structure that is necessary and sufficient for our result, independently of the payoff matrix, we first derive our characterization in the context of static games with exogenous signals. In this framework, the number of signals and the number of actions can be treated independently. We precisely determine for which signal structures there exists at least one payoff matrix (not necessarily the same for all signal structures) for which, in the game in which players receive a signal before choosing their action, a given player can be held down to a lower payoff than in the same game without signals.[2] Signal and action sets can no longer be treated independently in the repeated game, considered next, and thus the condition becomes only sufficient there.

[2] A complementary problem is to determine for which payoff matrices there exists at least one signal structure that allows one player to be held down below his minmax payoff. The answer is trivial, as it amounts to comparing the minmax and the correlated minmax payoff of the payoff matrix. Our question is motivated by the folk theorems, in which conditions are identified on the signal structure that are sufficient for all games.

Because a growing literature examines the robustness of folk theorems with respect to small perturbations in the monitoring structures, starting from either perfect monitoring (see [Sek97], [EV02], [BO02], [Pic02], [HO06]), or imperfect public monitoring ([MM02], [MM]), we actually also prove a stronger result. We show that, as the distance of the monitoring structure to any monitoring structure satisfying the aforementioned condition converges to zero, so does the distance between the individually rational payoff and the minmax payoff. Further, we show that this convergence is uniform in the discount factor, provided that the monitoring structure satisfies some standard identifiability assumptions that are necessary for the result.

As mentioned, almost-perfect monitoring is *not* sufficient in general for such a result; that is, the metric defining almost-perfect monitoring is weaker than what is necessary for such convergence. Nevertheless, it is an immediate corollary of our result that convergence obtains in the case of almost-perfect monitoring if in addition attention is restricted to the canonical signal structure, in which players signals are (not necessarily correct) action profiles of their opponents. This provides therefore a converse to Theorem 1 of [HO06].

An important question left open is how to actually determine the individually rational payoff when it is below the minmax payoff. Such a characterization is obtained under significantly more restrictive assumptions on the monitoring structure by [GT04]. Similarly, [Leh90] restricts attention to the case in which information is semi-standard. That is, there exists a partition of each player's action set, and, for a given action profile, players publicly observe the element of the partition corresponding to each player's action. Another open question is how to generally construct equilibria that achieve payoffs below the minmax payoff, assuming that the individually rational payoff is indeed below the minmax payoff. It is important to note that this may help support efficient outcomes as equilibria of the repeated game, as it enhances the severity of punishments.

Section two presents examples that motivate our condition. Section three proves that this condition precisely characterizes the signal structures for which, for any payoff matrix, a player can guarantee his minmax payoff. Section four extends the analysis to the case of infinitely repeated games and proves that the individually rational payoff is almost equal to the minmax payoff when the monitoring structure 'almost' satisfies our condition. Section five proves the main theorem of Section four.

2 The duenna game[3]

Two lovers (Player 1 and 2) attempt to coordinate on a place of secret rendezvous. They can either meet on the landscape garden bridge (B) or at the woodcutter's cottage (C). Unfortunately, the incorruptible duenna (Player 3) prevents any communication between them, and wishes to disrupt their meeting. Therefore, the rendezvous only succeeds if both lovers choose the same place and the duenna picks the other place. In all other cases, the duenna exults.

We model this situation as a 'team' zero-sum game, where the common payoff to the lovers is the probability of a successful meeting. Figure 1 displays the probability of a successful meeting,

[3] This game is sometimes referred to as the "three player matching pennies" game (see [MW98]). We find this name slightly confusing, given that the "three person matching pennies" game introduced earlier by [Jor93] is a different, perhaps more natural generalization of matching pennies.

as a function of the players' choice (lovers choose row and column; the duenna chooses the matrix).

	C	B
C	1	0
B	0	0

B

	C	B
C	0	0
B	0	1

C

Figure 1: The Duenna game

In the absence of any correlation device, the team can secure a payoff of 1/4, as the unique equilibrium point calls for all three players to choose both actions with equal probability. Yet if players 1 and 2 could secretly coordinate, they could achieve a probability of 1/2, by choosing the action profiles (B, B) and (C, C) with equal probability.

Now we turn to the repeated game, and assume that monitoring is potentially imperfect. This requires introducing some notation. Let Ω_i denote player i's (finite) set of signal, with generic element ω_i. The distribution of $\omega :\ = (\omega_1, \omega_2, \omega_3) \in \Omega := \otimes_i \Omega_i$ under action profile $a \in A$ is denoted q^a, with marginal distribution on Player i's signal given by q_i^a. A monitoring structure is denoted (Ω, q), where $q = \{q^a : a \in A\}$.

Example 1: (Almost-perfect monitoring) Recall from [MM02] that the monitoring structure (Ω, q) is ε-perfect if there exist signaling functions $f_i : \Omega_i \to A_{-i}$ for all i such that, for all $a \in A, i = 1, 2, 3$:

$$q^a \left(\{\omega : f_i (\omega_i) = a_{-i}\} \right) \geq 1 - \varepsilon.$$

That is, a monitoring structure is ε-perfect if the probability that the action profile suggested by the signal is incorrect does not exceed ε, for all possible action profiles.

Let $\Omega_1 = \{\omega_1^a, \omega_1'^a : a \in A\}$, $\Omega_2 = \{\omega_2^a, \omega_2'^a : a \in A\}$, $\Omega_3 = \{\omega_3^a : a \in A\}$. Consider

$$q^a \left((\omega_1^a, \omega_2^a, \omega_3^a) \right) = q^a \left((\omega_1'^a, \omega_2'^a, \omega_3^a) \right) = \frac{1 - \varepsilon}{2}, \text{ all } a \in A,$$

where $\varepsilon > 0$, and set $f_3 (\omega_3^a) = a$, $f_i (\omega_i^a) = f_i (\omega_i'^a) = a$, all $i = 1, 2$ and $a \in A$. The specification of the remaining probabilities is arbitrary. Observe that monitoring is ε-perfect, since the probability that a player receives either ω_i^a or $\omega_i'^a$ is at least $1 - \varepsilon$ under action profile a.

Yet players 1 and 2 can secure $(1 - \varepsilon)/2 \to_{\varepsilon \to 0} 1/2$ (as the discount factor tends to one). Indeed, they can play B if ω_i^a is observed at the previous stage, and C if $\omega_i'^a$ is observed at the previous stage, independently of a. Therefore, even under almost-perfect monitoring, the payoff of player 3 in this equilibrium is bounded away from his minmax payoff.

This shows that the set of equilibrium payoffs under almost-perfect monitoring can be significantly different from the set of equilibrium payoffs under perfect monitoring. In this example, the set of signals is richer under imperfect private monitoring than under perfect monitoring. Therefore, one may argue that the comparison of the minmax levels is not the most relevant one here. Indeed, the natural “limit” monitoring structure (as $\varepsilon \to 0$) in the example is a monitoring structure that is perfect but that also allows for a private correlation device for Players 1 and 2. Indeed, it is generally true that the minmax level is a continuous function of the signal distribution for fixed sets of signals (and discount factor). But this severely restricts the set of monitoring structures that can be considered close to the one usually considered in the literature on perfect monitoring.

Further, restricting the set of signals does not rule out correlation possibilities such as in Example 1, as they arise even under the *canonical* signal structure, the one for which $\Omega_i = A_{-i}$, for all i. It is a consequence of our main result that, if the monitoring structure is almost-perfect *and* the signal structure is canonical, then individually rational payoff and minmax payoff coincide.

Our second example shows that almost perfect monitoring and canonical signals is not enough to rule out situations in which the individually rational payoff of the repeated game is significantly lower than the min max of the one-shot game.

Example 2: (Perfect monitoring by player 3, canonical signal structure) Each player's set of signals is equal to his opponents' set of actions: $\Omega_i = A_{-i}$, for all i. Player 3's information is perfect:

$$q_3^a(a_{-3}) = 1, \ \forall a \in A.$$

Player 1 perfectly observes player 2's action, and similarly player 2 perfectly observes player 1's action. Their signal about player 3's action is independent of the action profile, but perfectly correlated across them. In particular:

$$\begin{aligned} q_1^a((a_2, C)) &= q_1^a((a_2, B)) = 1/2, \\ q_2^a((a_1, C)) &= q_2^a((a_1, B)) = 1/2. \end{aligned}$$

Consider the following strategies for player 1 and 2: play both actions with equal probability in the first period, and in later period, play C if the last signal about player 3 is C, and play B otherwise. This achieves a payoff of 1/2.

Example 2 shows that it is not enough to require that player 3 have perfect information about his opponents' actions, and/or that the signal structure be canonical, to rule out cases in which a player's individually rational payoff is lower than his minmax payoff. In Example 2, player 1 and 2's signal are completely uninformative about player 3's action, but it is straightforward to construct more complicated variations (with canonical signal structures) for which their signal is arbitrarily informative, and yet such that, for a given level of informativeness, the minmax payoff is bounded away from the individually rational payoff. One may argue that the problem here is that player 3's information is not nearly rich enough, as it does not include its opponents' signal about its own action. This, however, requires a larger set of signals, and takes us back to our initial example.

The issue is not solved either by requiring that the player's signals be almost public, a stronger requirement introduced and studied in [MM]. Indeed, even under public monitoring, it is known that the individually rational payoff may be lower than the minmax payoff. As mentioned, one example can be found in [FT91], exercise 5.10. A simpler one is provided by Example 3.

Example 3: (Public monitoring) Players' signal sets are identical: $\Omega_i = \{\omega_1, \omega_2\}$. Monitoring is public, that is, signals are identical for all three players. For all $a \in A$:

$$q^a(\omega_1) = 1 \text{ if } (a_1, a_2) = (B, B) \text{ or } (a_1, a_2) = (C, C), \text{ and } q^a(\omega_2) = 1 \text{ otherwise.}$$

Thus, signal ω_1 obtains if and only if players 1 and 2 choose the same action, and signal ω_2 obtains otherwise. Observe that players 1 and 2 can infer from the signal whether $(a_1, a_2) = (C, C)$ or (B, B) but player 3 cannot. The strategies of player 1 and 2 are as follows: play both actions with equal probability in odd periods, including the first one; play C (respectively, B) in even periods if and only the action profile in the previous period was (C, C) (respectively, (B, B)), and play both actions with equal probability otherwise. Such a strategy guarantees a payoff approaching 3/8 as the discount factor approaches one.

Finally, we provide an example in which the monitoring is imperfect, signals are not (conditionally) independent, yet players 1 and 2 cannot take advantage of it, independently of the payoff function for player 3.

Example 4: (A monitoring structure for which the individually rational payoff equals the minmax payoff) Each of the three players has two signals: $\Omega_i = \{\omega_i, \omega_i'\}$. Signal ω_3 has probability 1/2. The distribution of player 1 and 2's signals, conditional on player 3's signal, is given in Figure 2. Since one and only one of the signal profiles of player 1 and 2 has probability zero, conditional on either signal of player 3, it is immediate that these conditional distributions are not product distributions. Yet we claim that, in any game that may be played, along with this signal structure, player 3 can guarantee his minmax payoff.

	ω_2	ω_2'
ω_1	2/9	0
ω_1'	5/9	2/9

ω_3

	ω_2	ω_2'
ω_1	2/9	5/9
ω_1'	0	2/9

ω_3'

Figure 2: Conditional distributions

Why? Observe that player 3 can always decide to 'garble' his information, and base his decision on the garbled information, as summarized by two fictitious signals, $\tilde{\omega}_3$ and $\tilde{\omega}_3'$. In particular, upon receiving signal ω_3, he can use a random device that selects $\tilde{\omega}_3$ with probability 1/5, and selects $\tilde{\omega}_3'$ with probability 4/5; similarly, upon receiving signal ω_3', he can use a device that selects $\tilde{\omega}_3'$ with probability 1/5 and $\tilde{\omega}_3$ with probability 4/5. He then takes a decision that is optimal for his signal given player 1 and 2's strategies (i.e., maximizes his conditional expected payoff). It is easy to check that, *after* such a garbling, the distribution of player 1 and 2's signals, conditional on signal $\tilde{\omega}_3$, is the product distribution assigning probability 1/3 to signal ω_1 and

2/3 to signal ω_2. Likewise, the distribution of player 1 and 2's signals, conditional on signal $\tilde{\omega}_3'$, is the product distribution assigning probability 2/3 to signal ω_1 and 1/3 to signal ω_2. Observe finally that, if the distribution of player 1 and 2's signal is a product distribution, conditional on each possible signal of player 3, player 3 can guarantee his minmax payoff for whatever payoff function, since no matter how player 1 and 2 play, the resulting distribution over their action profiles remains a product distribution.

This last example shows that there is a close connection between the individually rational payoff and the existence of a garbling of player 3's signal that satisfies conditional independence (more precise definitions are introduced in Section 3). To study this connection in its simplest form, we abstract in the following section from the dynamic nature of repeated games and pose our problem in the context of static games with exogenous signals. This allows us to treat the size of the signal space and the size of the action space as independent, a property that no longer holds in the repeated game.

3 Games with a Signal Structure

In this section, we examine finite one-shot games in which, prior to play, each player receives a signal $\omega_i \in \Omega_i$ that follows some exogenous (joint) distribution. Each player can then condition his play on his signal, if he wishes to. We are interested in characterizing the signal structures under which some fixed player's equilibrium payoff may be lower than his minmax payoff. Therefore, we consider payoff matrices such that all players but player i have the same payoff, which is the exact opposite of player i's payoff. Such a game is called a *team zero-sum game.*[4] Given any set B, $\triangle B$ denotes the set of probability distributions over B, and when B is a subset of a vector space $\operatorname{co} B$ denotes the convex hull of B. Given a collection of sets $\{B_i\}$, $\otimes_i B_i$ denotes the Cartesian product of these sets.

Consider for example the duenna game. For which signal structures (i.e., distributions) does there exist (Perfect Bayesian) equilibria for which the payoff exceeds 1/4?

Let us first consider the simplest case, in which player 3 receives no signal.

Lemma 1 *If Ω_3 is a singleton set, there exists an equilibrium in which players 1 and 2's payoff strictly exceeds 1/4 if and only if the signals ω_1 and ω_2 are not independently distributed.*

Proof. The necessity part is obvious. To prove sufficiency, observe first that we can restrict attention to the case in which players 1 and 2 receive two signals only. Indeed, if player 1 and 2's signals are not independently distributed, we can always pick some $(\omega_1, \omega_2) \in \Omega_1 \times \Omega_2$ so that $\Delta := \Pr\{(\omega_1, \omega_2)\} \Pr\{(\omega_1', \omega_2')\} - \Pr\{(\omega_1, \omega_2')\} \Pr\{(\omega_1', \omega_2)\} \neq 0$, where $\omega_j' := \Omega_j \setminus \{\omega_j\}$: i.e., players can always treat all signals but one as a unique signal.[5] Reversing the meaning of signals if necessary, we may assume that $\Delta > 0$.

Further, we claim that we can assume that $\Pr\{(\omega_1, \omega_2)\} = \Pr\{(\omega_1', \omega_2')\} > \Pr\{(\omega_1, \omega_2')\} = \Pr\{(\omega_1', \omega_2)\}$. To see this, observe that players can 'garble' their signal by using randomizations

[4] Team zero-sum games are introduced and studied in [vSK97]

[5] To do so, observe that if signals are not independently distributed, there exists ω_1, ω_2 such that $\Pr\{(\omega_1, \omega_2)\} \neq \Pr\{\omega_1\} \Pr\{(\omega_2)\}$. Identify then all signals $\omega_i' \neq \omega_i$ and consider the resulting information structure.

$\{p_i, q_i\} \subseteq [0,1]^2$, $i = 1, 2$. That is, we define "signals" $\{\tilde{\omega}_i, \tilde{\omega}'_i\}$ as follows: conditional on observing ω_i, player i assigns probability p_i to the signal $\tilde{\omega}_i$ ($1 - p_i$ to the signal $\tilde{\omega}'_i$); similarly, conditional on observing ω'_i, he assigns probability q_i to the signal $\tilde{\omega}'_i$ ($1 - q_i$ to the signal $\tilde{\omega}_i$). We pick q_i, $i = 1, 2$, so that $\Pr\{(\tilde{\omega}_1, \tilde{\omega}_2)\} = \Pr\{(\tilde{\omega}'_1, \tilde{\omega}'_2)\}$ and $\Pr\{(\tilde{\omega}'_1, \tilde{\omega}_2)\} = \Pr\{(\tilde{\omega}_1, \tilde{\omega}'_2)\}$, i.e.

$$\begin{aligned} q_1 &= \frac{\frac{1}{2} - (1-p_1)\Pr\{(\omega_1,\omega_2)\} + \Pr\{(\omega_1,\omega'_2)\}}{\Pr\{(\omega'_1,\omega_2)\} + \Pr\{(\omega'_1,\omega'_2)\}}, \\ q_2 &= \frac{\frac{1}{2} - (1-p_2)\Pr\{(\omega_1,\omega_2)\} + \Pr\{(\omega'_1,\omega_2)\}}{\Pr\{(\omega_1,\omega'_2)\} + \Pr\{(\omega'_1,\omega'_2)\}}. \end{aligned}$$

Observe that we can choose p_1 and p_2 in $(0, 1/2)$ so that q_1 and q_2 are both in $[0, 1]$.[6] In addition,

$$\Pr\{(\tilde{\omega}_1,\tilde{\omega}_2)\} - \Pr\{(\tilde{\omega}_1,\tilde{\omega}'_2)\} = \frac{(1-2p_1)(1-2p_2)\Delta}{2\left(\Pr\{(\omega_1,\omega'_2)\} + \Pr\{(\omega'_1,\omega'_2)\}\right)\left(\Pr\{(\omega'_1,\omega_2)\} + \Pr\{(\omega'_1,\omega'_2)\}\right)},$$

which is strictly positive. Therefore, the common value on the diagonal strictly exceeds the common value off the diagonal; in other words, the common value on the diagonal exceeds $\frac{1}{4}$.

We are now ready to show that there are strategies which strictly improve upon the minmax payoff. Player $i = 1, 2$ plays T if he receives signal $\tilde{\omega}_i$ and B otherwise. Since $\Pr\{(\tilde{\omega}_1, \tilde{\omega}_2)\} = \Pr\{(\tilde{\omega}'_1, \tilde{\omega}'_2)\}$, player 3 is indifferent between both his actions, and his payoff is then equal to

$$\Pr\{(\tilde{\omega}_1, \tilde{\omega}_2)\} > 1/4,$$

which was to be shown. ■

Remark 1 *Given the previous lemma, one may conjecture that, when Ω_3 is a singleton and if the correlated* min max *payoff of the one-shot game is strictly below the* min max *payoff, then the individually rational payoff for player 3 in the repeated game is less than the* min max *payoff if and only if the same holds for the duenna game. This conjecture is incorrect.*

What does this suggest for the general case in which player 3 also receives signals? If it is the case that, conditional on each signal he can receive, player 1 and player 2's signals are independently distributed, player 1 and 2 cannot improve upon player 3's minmax payoff. Therefore, conditional independent distributions are part of the distributions for which the minmax payoff equals the individually rational payoff. This is not the only case, however: if player 1 and 2's signals are unconditionally distributed, player 3 can also secure his minmax payoff. After all, he can always choose to disregard his signal. Observe that this can occur even when there is no signal of player 3 for which player 1 and 2's signal would be conditionally independent. Example 4 further shows that we must include a larger class of distributions: namely, player 3 can secure his standard minmax if the distribution of his signals is *sufficient* for another distribution of signals for which conditional independence holds. This class includes the two previous cases as special cases but is much richer (for instance, it can be shown that the game of Example 4 remains in this class for *any* prior such that the probability of ω_3 is between 1/5 and 4/5).

[6] Observe that $\Delta \neq 0$ implies that $\Pr\{(\omega_1,\omega_2)\} + \Pr\{(\omega_1,\omega'_2)\}$ and $\Pr\{(\omega'_1,\omega_2)\} + \Pr\{(\omega'_1,\omega'_2)\}$ are both in $(0, 1)$.

In the remainder of this section, we prove that this class of distributions exactly characterizes the distributions for which the individually rational payoff equals the minmax payoff. That is, for any signal structure that does not have the property that some garbling of it satisfies conditional independence (from the point of view of some player i), there exists a finite game in which the individually rational payoff of player i is strictly lower than his minmax payoff. To this end, one can verify that the duenna game is no longer sufficient.

Before stating the result, further notation and formal definitions must be introduced. Given a team zero-sum game G (without signal structure) between player i (fixed throughout) and all other ('team') players, we denote by $\Gamma(q, G)$ the (zero-sum) game obtained from G by adjoining the signal structure q (where the set of signals is given by the domain of q). The highest payoff that player i can secure is denoted $v_i(G)$ in the game G, and $V_i(q, G)$ in the game $\Gamma(q, G)$. The set of independent distributions $M = \otimes_{j \neq i} \triangle\Omega_j$ is viewed as a subset of $\triangle\Omega_{-i}$. As usual, the product of two distributions p and q is denoted $p \otimes q$.

Definition 1 *A family $p = (p_{\omega_i})_{\omega_i}$ of elements of $\triangle M$ is a* garbling of $q \in \triangle\Omega$ on M *for player i if, letting $q \otimes p \in \triangle(\Omega \times M)$ denote the distribution induced by q and p, there is a version of $p \otimes q(\omega_{-i}|m)$ such that:*

$$p \otimes q(\omega_{-i}|m) \in M \qquad a.s.$$

The distribution $q \in \triangle\Omega$ admits a garbling on M for player i *if there exists a garbling of $q \in \triangle\Omega$ on M.*

In words, upon receiving a signal ω_i, player i draws a random $m \in M$ according to p_{ω_i}, remembers m and forgets ω_i. When $(p_{\omega_i})_{\omega_i}$ is a garbling of q on M, the new belief of player i over Ω_{-i} is then $p \otimes q(\omega_{-i}|m) \in M$.

Given $q \in \triangle\Omega$, let $\boldsymbol{\omega}$ be a random variable with law q. We view $q(\omega_{-i}|\boldsymbol{\omega}_i)$ as a random variable with values in $\triangle\Omega_{-i}$ and denote its distribution by $\mu_q^i \in \triangle(\triangle\Omega_{-i})$ (note that μ_q^i depends only on q, not on the particular choice of $\boldsymbol{\omega}$).

Lemma 2 *Let $\beta \in \triangle M$. There exists a garbling $p = (p_{\omega_i})_{\omega_i}$ of $q \in \triangle\Omega$ on M for player i such that $p \otimes q$ has marginal β on M if and only if for every bounded convex function ψ on $\triangle\Omega_{-i}$,*

$$\mathbf{E}_\beta \psi^\beta \leq \mathbf{E}_\mu \psi^\beta.$$

Proof. This is an extension of the theorem of [Bla51] to the infinite-dimensional case. See [CFM64] and [Str65]. ■

We may now state and prove the main result of this section, establishing that the signal structures for which the individually rational payoff equals the minmax payoff for all payoff matrices are precisely those admitting a garbling satisfying conditional independence.

Proposition 3 *The distribution $q \in \triangle\Omega$ admits a garbling on M for player i if and only if, for every finite game G,*

$$v_i(G) = V_i(q, G).$$

Proof. We start with the easier "only if" part. Assume that $p = (p_{\omega_i})_{\omega_i}$ is a garbling of q on M for player i, and let G be a finite game with actions sets $(A_j)_j$ and payoff function g_i for player i. Consider strategies $\sigma_j \colon S_j \to \triangle A_j$ for players $j \neq i$ in the game $\Gamma(q, G)$, and for every $m \in M$, let $x(m) \in \otimes_{j \neq i} A_j$ be the image of m by σ_{-j} given by $x(m)(a_{-i}) = \sum_{\omega_{-i}} m(\omega_{-i}) \otimes_j \sigma_j(\omega_j)(a_j)$. For every $x \in \otimes_{j \neq i} \triangle A_j$, let $a_i(x)$ be a best response to x for player i. Consider the strategy σ_i for player i in $\Gamma(q, G)$ defined by $\sigma_i(\omega_i) = \int_M a_i(x(m)) dp_{\omega_i}(m)$. The expected payoff for player i in $\Gamma(q, G)$ when σ is the profile of strategies followed is

$$E_d g(\sigma_{-i}(\omega_{-i}), \sigma_i(\omega_i)) = \int_M g(a_i(x(m)), x(m)) d(p \otimes q)(m) \geq V_i(G).$$

For the "if" part, endow $A := \{\beta, \operatorname{supp} \beta \subseteq M\}$, subset of the set of distributions over $D = \triangle S_{-i}$, with the weak* topology. Let B be the set of continuous convex functions on D bounded by 0 and -1, endowed with the topology of the uniform norm. Both A and B are compact convex sets.

Since q admits no garbling on M for player i, by Lemma 2,

$$\forall \beta \in A, \exists \psi^\beta \in B, \mathbf{E}_\beta \psi^\beta > \mathbf{E}_\mu \psi^\beta. \tag{1}$$

The map $g \colon A \times B \to \mathbb{R}$ given by $g(\beta, \psi) = \psi^\beta - \mathbf{E}_\beta \psi^\beta$ is continuous, so, by the min max theorem, the two-player zero-sum game in which player I's action set is A, player II's action set is B and the payoff to I is given by g has a value v, and 1 implies $v < 0$. Hence there exists $\psi \in B$ such that

$$\forall \beta \in A, \mathbf{E}_\beta \psi > \mathbf{E}_\mu \psi - v. \tag{2}$$

Letting $\psi' = \psi - \mathbf{E}_\mu \psi + v/2$ we have $\mathbf{E}_\mu \psi' = v/2 < 0$ and $\psi'(m) > -v/2 > 0$ for all $m \in M$.

For every $m \in M$, there exists a linear map ϕ_m on D such that $\phi_m(m) > 0$ and $\phi_m \leq \psi$. Let O_m be a neighborhood of m such that $\psi_m > 0$ on O_m. Since M is a compact subspace of D and $\cup_m O_m = M$, there exists a finite subfamily $M_0 \subset M$ such that $\psi'' := \max_{m \in M_0} \phi_m > 0$ on M. Furthermore $\mathbf{E}_\mu \psi'' \leq \mathbf{E}_\mu \psi < 0$.

Now consider the game G in which each $j \neq i$ has strategy set Ω_j, player i has strategy set M_0, and i's payoff function g_i is given by $g_i(m_0, \omega_{-i}) = \phi_{m_0}(\omega_{-i})$. For each m in M and $m_0 \in M_0$, $\mathbf{E}_m g_i(m_0, \omega_{-i}) = \phi_{m_0}(m) > 0$, so that $\min_{m \in M} \max_{m_0 \in M_0} g_i(m_0, m) > 0$. On the other hand, in the game in which signals are distributed according to μ and the payoff function is g_i, if players $j \neq i$ play their signals in G, player i's best response yields a payoff of $\mathbf{E}_\mu \max_{m_0 \in M_0} \phi(d) \leq \mathbf{E}_\mu \psi(d) < 0$. ■

This indicates that the existence of a garbling that satisfies conditional independence is the appropriate condition for our purposes. However, in the repeated game, the problem is slightly more complicated. First, the distribution is itself a function of (past) actions. Therefore, it must be strengthened to the existence of a garbling providing conditional independence, for each possible action profile, whether pure or mixed.[7] Second, because it is important to include cases in which conditional independence need not hold exactly (as is typically the case if we consider, for instance, a monitoring structure that is almost, but not perfect), we must allow for small departures from conditional independence, which complicates the analysis, especially since we aim for a bound that is uniform in the discount factor.

[7] Observe that conditional independence of signals for each pure action profile does not imply conditional independence for all mixed action profiles.

4 Repeated Games with Imperfect Monitoring

A stage game G specifies the set of players $i = 1, ..., n$, each player's (finite) action set A_i and, for each player i, a payoff function $g_i \colon A := A_1 \times \cdots \times A_n \to \mathbb{R}$. We restrict attention to games G for which $|g_i(a)| < 1$, all $i = 1, \ldots, n$, $a \in A$ (the specific choice of the upper bound is obviously irrelevant).

Players can use mixed actions $\alpha_i \in \triangle A_i$. The domain of g_i is extended to mixed action profiles $\alpha \in \triangle A$ in the usual manner:

$$g_i(\alpha) = \sum_{a \in A} \alpha(a) g_i(a).$$

Mixed actions are unobservable. No public randomization device is assumed.

We consider the infinitely repeated game, denoted G^∞. In each period, player i observes a private signal ω_i from some finite set Ω_i, whose distribution depends on the action profile being played. Therefore, player i's information in period t consists of both his action a_i and his private signal ω_i. Let $s_i = (a_i, \omega_i)$ denote this information, or *signal* for short, and define $S_i := A_i \times \Omega_i$. The monitoring structure determines a distribution over private signals for each action profile. For our purpose, it is more convenient to define it directly as a distribution over $S := S_1 \times \cdots \times S_n$. Given action profile $a \in A$, $q^a(s)$ denotes the distribution over signal profile $s = (s_1, \ldots, s_n)$. We extend the domain of this distribution to mixed action profiles $\alpha \in \triangle A$, and write $q^\alpha(s)$. Let q_i^α denote the marginal distribution of q^α over player i's signals s_i, and given $s_i \in S_i$ and $\alpha \in \triangle A$ such that $q_i^\alpha(s_i) > 0$, let $q_{-i}^\alpha(\cdot \mid s_i)$ denote the marginal distribution over his opponents' signals, conditional on player i's signal s_i. From now on, a monitoring structure refers to such a distribution q.

Players share a common discount factor $\delta \in (0,1)$, but as will be clear, its specific value is irrelevant (statements do not require that it be sufficiently large). Repeated game payoffs are discounted, and their domain is extended to mixed strategies in the usual fashion; unless explicitly mentioned otherwise (as will occur), all payoffs are normalized by a factor $1 - \delta$.

For each i, the *minmax payoff* v_i^* of player i (in mixed strategies) is defined as

$$v_i^* := \min_{\alpha_{-i} \in \otimes_{j \neq i} \triangle A_j} \max_{a_i \in A_i} g_i(a_i, \alpha_{-i}).$$

A private history of length t for player i is an element of $H_i^t := S_i^t$. A (behavioral) private strategy for player i is a family $\sigma_i = (\sigma_i^t)_t$, where $\sigma_i^t \colon H_i^{t-1} \to \triangle A_i$. We denote by Σ_i the set of these strategies. Player i's individually rational payoff in the repeated game is the lowest payoff he can be held down to by any choice $\sigma_{-i} = (\sigma_j)_{j \neq i}$ of *independent* choices of strategies in the repeated game, provided that player i correctly foresees σ_{-i} and plays a best-reply to it. Since our purpose is to investigate individually rational payoffs and minmax payoffs, there is no need to introduce a solution concept.

A distribution q on S, or S_{-i}, is a *product distribution* if it is the product of its marginal distributions: for instance a distribution q on S is a product distribution if and only if $q((s_1, \ldots, s_n)) = q_1(s_1) \cdots q_n(s_n)$ for all $s_1, \ldots, s_n$. This class is the key to our characterization.

Theorem 4 *If*

$$q_{-i}^\alpha(\cdot \mid s_i) \in \otimes_{j \neq i} \triangle S_j,\ \forall \alpha \in \otimes_j \triangle A_j,\ \forall s_i \in S_i \text{ such that } q_i^\alpha(s_i) > 0,$$

then player i's individually rational payoff equals his minmax payoff.

This result is proved in the next section, as an immediate consequence of the first step of the proof of Theorem 6. To state the next important but straightforward extension of Theorem 4, one more notation must be introduced. A distribution q on S is in $\mathcal{D}_i^B$ if q admits a garbling on $\otimes_{j \neq i} \triangle S_j$ for player i.

Corollary 5 *If*

$$q^\alpha \in \mathcal{D}_i^B,\ \forall \alpha \in \otimes_j \triangle A_j,$$

then player i's individually rational payoff equals his minmax payoff.

Observe that this corollary generalizes Theorem 4, as if $q_{-i}^\alpha (\cdot \mid s_i)$ is in $\otimes_{j \neq i} \triangle S_j$ for all s_i in S_i, then q^α is in $\mathcal{D}_i^B$. Also, if q_{-i}^α is a product distribution, that is, if signals of players $j \neq i$ are unconditionally independently distributed, then q^α is in $\mathcal{D}_i^B$, as player i can altogether disregard his signal and take a best-reply to his prior belief. This may occur for distributions such that $q_{-i}^\alpha (\cdot \mid s_i) \notin \otimes_{j \neq i} \triangle S_j$, for all s_i in S_i. In fact, $\mathcal{D}_i^B$ is much 'richer' than those special cases. However, while it is straightforward to check whether $\left\{ q_{-i}^\alpha (\cdot \mid s_i) : s_i \in S_i \right\} \subseteq \otimes_{j \neq i} \triangle S_j$ by considering each element $q_{-i}^\alpha (\cdot \mid s_i)$ 'in turn', it is harder to ascertain whether $q^\alpha \in \mathcal{D}_i^B$.

For a given discount factor, it is straightforward to see that the individually rational payoff is continuous in the monitoring structure. In particular, if q 'almost' satisfies the condition in Theorem 4, or the more general condition in Corollary 5, then the individually rational payoff is approximately equal to the minmax payoff. Such a result is unsatisfactory because it does not rule out that, even if q^α is arbitrarily close to $\mathcal{D}_i^B$ (for all mixed action profiles α in $\triangle A$), there may exist a discount factor, sufficiently close to one, for which the individually rational payoff is bounded away from the minmax payoff. Over time, 'small' amounts of secret correlation among player i's opponents may accumulate, allowing them eventually to successfully coordinate their play. Indeed, this is in general possible, as the following example demonstrates.

Example 5: (A monitoring structure that satisfies 'almost' independence) The payoff matrix is given by the duenna game. Player 1 and 2's signal set has two elements, $\Omega_i = \{\omega_i, \omega_i'\}$. Player 3 receives no signal. The distribution of player 1 and 2's signals is independent of the action profile and perfectly correlated. With probability $\epsilon > 0$, the signal profile is (ω_1, ω_2), and it is equal to (ω_1', ω_2') with probability $1 - \epsilon > 0$.

Given ϵ, let H_i^T denote the set of private histories of signals of length T for players $i = 1, 2$, and let $H_i^{\prime,T}$ denote the subset of H_i^T consisting of those histories in which the empirical frequency of signals ω_i exceeds the expectation of this number, $T\epsilon$. Observe that, by the central limit theorem, the probability that a history of length T is in H_i^T tends to 1/2 as $T \to \infty$. Define s_i^T as the strategy consisting in playing C for the first T periods and then playing C forever after if the private history is in $H_i^{\prime,T}$, and playing B forever after if it is not. It follows that the lowest equilibrium payoff to players 1 and 2 from using the strategy profile (s_1, s_2) tends to 1/2 as $\delta \to 1$ tends to 1/2, as we take $T \to \infty$ but $\delta^T \to 0$. On the other hand, for any fixed δ, players 1 and 2 cannot secure more than 1/4 as $\epsilon \to 0$.

Example 5 shows that, in general, the order of limits is important. While payoff is continuous in the monitoring structure for a fixed discount factor, the limit of this payoff as the discount factor tends to one may be discontinuous in the monitoring structure. To rule out such an example, player i's signal must be sufficiently informative. That is, for player $-i$'s secret correlation to dissipate, player i must be able to statistically discriminate among action profiles of his opponents.

Definition 2 *(Identifiability) The conditional distribution* $q_i := \{q_i^a : a \in A\}$ *satisfies identifiability if, for all* a_{-i} *in* A_{-i} *and* α_i *in* $\triangle A_i$*,*

$$q_i^{a_{-i},\alpha_i} \notin \text{co}\left\{q_i^{a'_{-i},\alpha_i} : a'_{-i} \neq a_{-i}, a'_{-i} \in A_{-i}\right\}.$$

That is, q_i satisfies identifiability if, for any possibly mixed action of player i, the distribution over his signals that is generated by any pure action profile of his opponents cannot be replicated by some convex combination of other action profiles of theirs. Let d denote the total variation distance between probability measures. The conditional distribution q_i satisfies *ρ-identifiability* if, for all a_{-i} in A_{-i} and α_i in $\triangle A_i$, and any distribution q'_i in co $\left\{q_i^{a'_{-i},\alpha_i} : a'_{-i} \neq a_{-i}, a'_{-i} \in A_{-i}\right\}$,

$$d\left(q_i^{a_{-i},\alpha_i}, q'_i\right) > \rho.$$

Thus, the concept of ρ-identifiability measures the distance between the distribution q_i and the nearest conditional distribution that fails to satisfy identifiability.

Finally, we need to introduce a measure of the distance between a given conditional distribution $q_{-i} := \left\{q_{-i}^a : a \in A\right\}$ and the nearest product distribution. The conditional distribution q_{-i} is *ε-dependent* if, for all action profiles α in $\otimes_j \triangle A_j$, there exists q'_{-i} in $\otimes_{j\neq i} \triangle S_j$ such that

$$\mathbf{E}\left[d\left(q_{-i}^{\alpha}\left(\cdot \mid s_i\right), q'_{-i}\right)\right] < \varepsilon,$$

where $\mathbf{E}[\cdot]$ denotes the expectations operator. That is, the conditional distribution q_{-i} is ε-dependent, if those signals for which the conditional distribution of player i is not close to a product distribution are sufficiently unlikely, given any action profile that corresponds to independent randomizations.

Theorem 6 *For any* $\nu > 0$*, if* q_i *satisfies* $\rho-$*identifiability, for some* $\rho > 0$*, then there exists* $\varepsilon > 0$ *such that, if* q_{-i} *is* ε*-dependent, then player* i*'s individually rational payoff is within* ν *of its minmax payoff.*

Theorem 6 strengthens Theorem 4 and provides a continuity result that is uniform in the discount factor (observe that the discount factor does not enter the statement of the theorem, i.e., the value of ε is independent of δ). This theorem is important for the literature on the robustness of equilibria in a neighborhood of perfect, or imperfect public monitoring. Indeed, while almost-perfect monitoring structures need not be ε-dependent for small ε (as expected given Example 1), it is immediate to see that they must be if attention is restricted to canonical signal structures. Therefore, Theorem 4 provides a converse to Theorem 1 in [HO06].

It is straightforward to extend Theorem 6 to distributions that are close to $\mathcal{D}_i^B$ *provided* that the garbled signal satisfies the identifiability condition (that is, the belief of player i, conditional on his garbled signal, should satisfy ρ-identifiability. The generalization is straightforward and omitted.

Finally, a large literature has considered a restricted class of strategies, namely public strategies, in the context of games with public monitoring. In such games, the minmax payoff in public strategies in the repeated game cannot be lower than the static minmax payoff, a result which is not generally true without the restriction on strategies (see the last example of the previous section). A natural question is to what extent the ε-dependence assumption can be weakened for such a class of strategies. To be more specific, assume that strategies must be a function of the history of private signals ω_i alone, rather than of the history of all signals s_i. Observe that this reduces to public strategies in the case of public monitoring, but is well-defined even under private monitoring. Then Theorem 6 remains valid, if we require that only the restriction of q_{-i} to private signals be ε-dependent. This is a significantly weaker restriction, which is indeed trivially satisfied if monitoring is public. The proof is a trivial modification of the proof of Theorem 6.

All statements are either trivial or follow from the proof of Theorem 6. The proof of Theorem 6 is rather long, and divided in two parts. In the first part, we replace the private monitoring structure by another one. Player i's information is unchanged. His opponents' information is now public among them, but it is not simply the information resulting from pooling their individual signals from the original monitoring structure: doing so would enable them to correlate their play in many situations in which they would not be able to do so if their strategy were only based on their own signals. The common information must be 'poorer' than that, but we still need to make sure that any probability distribution over plays that could be generated in the original monitoring structure by some strategy profile of player i's opponents (for some strategy of player i) can still be generated in this alternate monitoring structure. The argument is somewhat delicate and presented in the next subsection.

By considering the alternate monitoring structure, tools from information theory are brought to bear. This is done in the second part of the proof, in which we show that, under ε-epsilon dependence and ρ-identifiability, it takes time to accumulate sufficient public information for player i's opponents to successfully correlate their action profile, relative to the time it takes player i to detect on which of the plays his opponents have coordinated upon. The argument is presented in Subsection 3.2.

5 Proof of Theorem 6

5.1 Reduction to public strategies

A result that will prove useful here is the following.

Lemma 7 *If q is a distribution over some product finite set $S = \Pi_{k \in K} S_k$, then there exists a product distribution $p \in \otimes_j \triangle S_j$ and a 'residual' distribution r such that*

$$q = \lambda p + (1 - \lambda) r,$$

for some $\lambda = \lambda(q)$ in $[0, 1]$. Further, for every $\nu > 0$, there exists $\varepsilon > 0$ such that, if $d(q, q') < \varepsilon$ for some $q' \in \otimes_j \triangle S_j$, then we can choose $\lambda > 1 - \nu$.

Proof. Indeed, if we define λ as the supremum over all such values in the unit interval for which we can write q as a convex combination of distributions p and r, with $p \in \otimes_j \triangle S_j$, it by the maximum theorem it follows that (i) this maximum is achieved, (ii) it is continuous in q. In fact, since q belongs to a compact metric space, this continuity is uniform, by the Heine-Cantor theorem. The result follows, since $\lambda = 1$ if $q \in \otimes_j \triangle S_j$. ■

Given this result, we can view signals in the repeated game as being drawn in three stages. Given the action profile (α_{-i}, a_i):

- first, the signal s_i is drawn according to the marginal distribution $q_i^{\alpha_{-i},a_i}$. Given s_i, apply (7) and write

$$q_{-i}^{\alpha_{-i},a_i}(\cdot|s_i) = \lambda p_{-i}^{\alpha_{-i},a_i}(\cdot|s_i) + (1-\lambda)\, r_{-i}^{\alpha_{-i},a_i}(\cdot|s_i).$$

where λ depends on $q_{-i}^{\alpha_{-i},a_i}(\cdot|s_i)$.

- second, a Bernoulli random variable l with $\Pr\{l = 1\} = 1 - \Pr\{l = 0\} = \lambda$ is drawn;

- third, if $l = 0$, the signal profile s_{-i} is drawn according to $r_{-i}^{\alpha_{-i},a_i}(\cdot|s_i)$; if instead $l = 1$, s_{-i} is drawn according to $p_{-i}^{\alpha_{-i},a_i}(\cdot|s_i)$.

We show in this subsection that player i's individually rational payoff is no larger under the original monitoring structure than under an alternate monitoring structure in which player i's opponents can condition their strategy on the history of values of s_i, l, and, whenever $l = 0$, of s_{-i}. This is nontrivial because player i's opponents are not allowed to condition their strategy on their own signals any longer, unless $\lambda = 0$. Yet the conclusion is rather intuitive, for when $\lambda = 1$, the signals of player i's opponents are independently distributed anyway (conditional on s_i). This result will allow us in the next subsection to view the histories used by player i's opponents as common.

Before stating the result, further notation must be introduced.

Histories: Recall that a *history* of length t in the original game is an element of S^t. The set of plays is $H^\infty = S^{\mathbb{N}}$ endowed with the product σ-algebra. We define an *extended history* of length t as an element of $(S \times \{0,1\} \times S')^t$, that is, as a history in the original game augmented by the history of realizations of the Bernoulli variable l. The set of extended plays is $\tilde{H}^\infty = (S \times \{0,1\})^{\mathbb{N}}$ endowed with the product σ-algebra.

A private history of length t for player j (in the original game) is an element of $H_j^t = S_j^t$. A *public history* of length t is an element of $H_p^t := S_p^t$, where $S_p\colon = S_i \times \{0\} \times S_{-i} \cup S_i \times \{1\}$; that is, S_p is the set of public signals s_p, where $s_p = (s_i, 0, s_{-i})$ if $l = 0$ and $s_p = (s_i, 1)$ if $l = 1$.

Strategies: A (behavioral) private strategy for player j (in the original game) is a family $\sigma_j = (\sigma_j^t)_t$, where $\sigma_j^t\colon H_j^{t-1} \to \triangle A_j$. Let Σ_j denote the set of these strategies. A (behavioral) public strategy for player j is a family $\tau_j = (\tau_j^t)_t$, where $\tau_j^t\colon H_p^{t-1} \to \triangle A_j$. Let $\Sigma_{p,j}$ denote the set of these strategies. Finally, a (behavioral) general strategy for player j is a family $\tilde{\sigma}_j = (\tilde{\sigma}_j^t)_t$, where $\sigma_j^t\colon (S_p \times S_j)^{t-1} \to \triangle A_j$. Let $\tilde{\Sigma}_j$ denote the set of these strategies.

Note that both $\Sigma_{p,j}$ and Σ_j can naturally be identified as subsets of $\tilde{\Sigma}_j$, but $\Sigma_{p,j}$ and Σ_j cannot be ordered by set inclusion.

A (pure) strategy for player i is a family $\sigma_i = (\sigma_i^t)_t$, where $\sigma_i^t\colon S_i^{t-1} \to A_i$.

Any profile of general strategies σ_{-i} for player i's opponents, together with a strategy σ_i for player i, induces a probability distribution P_{σ_{-i},σ_i} on $\tilde{H}^\infty$.

Proposition 8 *For any private strategy profile σ_{-i}, there exists a public strategy profile τ_{-i} such that, for every pure strategy σ_i, $h_i^t \in S_i^t$, $s^{t+1} \in S$:*

$$P_{\tau_{-i},\sigma_i}(h_i^t) = P_{\sigma_{-i},\sigma_i}(h_i^t) \tag{3}$$
$$P_{\tau_{-i},\sigma_i}(s^{t+1}|h_i^t) = P_{\sigma_{-i},\sigma_i}(s^{t+1}|h_i^t) \qquad \text{if} \qquad P_{\sigma_{-i},\sigma_i}(h_i^t) > 0. \tag{4}$$

Proof. We first define a *public strategy up to stage* t for player j as a family $\tau_{t,j} = (\tau_{t,j}^{t'})_{t'}$ where

$$\begin{cases} \tau_{t,j}^{t'} : S_p^{t'-1} \to \triangle A_j, & \text{if } t' \leq t; \\ \tau_{t,j}^{t'} : S_p^{t-1} \times S_j^{t'-t+1} \to \triangle A_j, & \text{otherwise.} \end{cases}$$

The proof of the Proposition relies on the following lemma. This lemma exhibits a sequence of strategy profiles for player i's opponents, up to stage t, based on σ_{-i}, that do only depend on the first t public signals, and not on the realizations of the first t private signals (conditional on these public signals). This sequence of strategies is constructed by iterated applications of Kuhn's theorem.

Lemma 9 *For any private strategies σ_{-i}, there exist strategies $(\tau_{t,-i})_t = (\tau_{t,j})_{j\neq i,t}$ where $\tau_{t,j}$ is a public strategy up to stage t for player j and*

$$\tau_{0,j} = \sigma_j \tag{5}$$
$$\tau_{t,j}^t = \tau_{t',j}^t \quad \text{for } t' \geq t \tag{6}$$
$$P_{\tau_{t+1,-i},\sigma_i}(s^{t+1},\ldots,s^{t+n}|h_p^t) = P_{\tau_{t,-i},\sigma_i}(s^{t+1},\ldots,s^{t+n}|h_p^t) \tag{7}$$

for every $\sigma_i, n, (s^{t+1},\ldots,s^{t+n}) \in S^n$, and $h_p^t \in H_p^t$.

Proof. Define $\tau_{t,-i}$ by induction on t. First let $\tau_{0,-i} = \sigma_{-i}$ so that (5) is met. Assume $\tau_{t,-i}$ has been defined. Let $\tau_{t+1,-i}^{t'} = \tau_{t,-i}^{t'}$ for $t' \leq t$ so that (6) is satisfied.

For each history h_p^t and for every $s_j \in S_j$, let $\tilde{\tau}_{t+1,j}[h_p^t, s_j]$ be the private (continuation) strategy defined by $\tilde{\tau}_{t+1,j}[h_p^t, s_j](s_j^1,\ldots,s_j^n) = \tau_{t,j}(h_p^t, s_j, s_j^1,\ldots,s_j^n)$.

The probability $q^{\tau_{t,-i}(h_p^{t-1}),a_i^t}(s_j|s_p)$ over S_j defines a mixture of private strategies $\tilde{\tau}_{t+1,j}^{t'}[h_p^t, s_j]$, where a_i^t is Player i's action in period t as specified by s_p. By Kuhn's theorem, there exists a private strategy $\tau_{t+1,j}[h_p^t]$ which is equivalent to this mixture. Finally set $\tau_{t+1,j}^{t+n+1}(h_p^t, s_j^1,\ldots,s_j^n) = \tau_{t+1,j}^{t+n+1}[h_p^t](s_j^1,\ldots,s_j^n)$. Condition (7) is met by equivalence of the mixed and the behavioral strategy and because all $(s_j)_j$ are independent conditional on s_p and h_p^t. ■

Back to the proof of Proposition 8, define τ_{-i} by $\tau_j^t(h_p^{t-1}) = \tau_{t,j}^t(h_p^{t-1})$, where $(\tau_{t,-i})_t$ is given by Lemma 9.

¿From (6), for every t', $P_{\tau_{t'+1,-i},\sigma_i}$ and $P_{\tau_{t',-i},\sigma_i}$ induce the same probability on H_p^t, and from (7), $P_{\tau_{t+1,-i},\sigma_i}(h_i^{t'}|h_p^t) = P_{\tau_{t,-i},\sigma_i}(h_i^{t'}|h_p^t)$ for $t' \geq t$. Thus $P_{\tau_{-i},\sigma_i}(h_i^t) = P_{\tau_{t,-i},\sigma_i}(h_i^t) = P_{\tau_{t-1,-i},\sigma_i}(h_i^t) = \ldots = P_{\sigma_{-i},\sigma_i}(h_i^t)$, which gives (3).

Also, $P_{\tau_{-i},\sigma_i}(s^{t+1}|h_i^t) = P_{\tau_{t+1,-i},\sigma_i}(s^{t+1}|h_i^t) = \frac{P_{\tau_{t+1,-i},\sigma_i}(s^{t+1},h_i^t)}{P_{\tau_{t+1,-i},\sigma_i}(h_i^t)} = \frac{P_{\sigma_{-i},\sigma_i}(s^{t+1},h_i^t)}{P_{\sigma_{-i},\sigma_i}(h_i^t)} = P_{\sigma_{-i},\sigma_i}(s^{t+1}|h_i^t)$ whenever $P_{\sigma_{-i},\sigma_i}(h_i^t) > 0$ which gives (4). ■

An immediate consequence of Proposition (8) is that the individually rational payoff of player i is necessarily at least as high under the alternate monitoring structure in which player i's opponents use so-called public strategies, as under the original monitoring structure (whether we consider the finitely or infinitely-repeated game, and independently of the discount factor). Note that this already establishes Theorem 2 (and therefore Corollary 3). Indeed, under the assumption of Theorem 2, we have $\Pr\{l=1\}=1$, so that public strategies only depend on s_i, which is known by player i. That is, conditional on his history of private signals, player i can view the choices of continuation strategies of his opponents as independent. Matters are more complicated when the monitoring structure is only ε-dependent, as $\Pr\{l=1\}<1$. Nevertheless, the event $\{l=0\}$ is unfrequent for ε small enough.

5.2 Measuring secret correlation and its dissipation

The aim of this section is to prove that, relying on public strategies, and under the conditions of Theorem 6, little secret correlation can be generated by team members in the course of the repeated game and if this correlation is used there is enough dissipation of this correlation with time. This implies that the individually rational payoff of player i is uniformly close to his minmax payoff in mixed strategies.

In order to measure the amount of secret correlation generated and dissipated by the team in the course of the repeated game, we rely on the entropy measure of randomness and information.

We start with some reminders on information theory, then derive a bound on the minmax payoffs based on entropy, and finally utilize this bound to prove the main result.

5.2.1 Reminder on information theory

The entropy of a finite random variable $\mathbf{x}$ with law P is by definition:

$$H(\mathbf{x})=-\mathbf{E}[\log P(\mathbf{x})]=-\sum_x P(x)\log P(x)$$

where log denotes the logarithm with base 2, and $0\log 0=0$. Note that $H(\mathbf{x})\geq 0$ and that $H(\mathbf{x})$ depends only on the law P of $\mathbf{x}$. The entropy of $\mathbf{x}$ is thus the entropy $H(P)$ of its distribution P, with $H(P)=-\sum_x P(x)\log P(x)$.

If $(\mathbf{x},\mathbf{y})$ is a couple of finite random variables with joint law P, the conditional entropy of $\mathbf{x}$ given $\{\mathbf{y}=y\}$ is the entropy of the conditional distribution $P(\mathbf{x}|y)$:

$$H(\mathbf{x}\mid y)=-\mathbf{E}[\log P(\mathbf{x}\mid y)].$$

The conditional entropy of $\mathbf{x}$ given $\mathbf{y}$ is the expected value of the previous:

$$H(\mathbf{x}\mid\mathbf{y})=\sum_y H(\mathbf{x}\mid y)P(y).$$

We have the following relation of additivity of entropies:

$$H(\mathbf{x},\mathbf{y})=H(\mathbf{y})+H(\mathbf{x}\mid\mathbf{y}).$$

Finally define the mutual information between x and y as

$$I(\mathbf{x};\mathbf{y})=H(\mathbf{x})-H(\mathbf{x}|\mathbf{y})=H(\mathbf{y})-H(\mathbf{y}|\mathbf{x}).$$

5.2.2 An entropy bound on minmax payoffs

Let $\sigma = (\sigma_{-i}, \sigma_i)$ be a strategy profile, where σ_{-i} is a profile of public strategies. Suppose that after stage t, the history for player i is $h_i^t = (s_i^1, \ldots, s_i^t)$. Let $h_p^t = (s_p^1, \ldots, s_p^t)$ be the public history after stage t. The mixed action profile played by the team at stage $t+1$ is $\sigma_{-i}^{t+1}(h_p^t) = \otimes_{j \neq i} \sigma_j(h_p^t)$. Player i holds a *belief* on this mixed action, namely he believes that players $-i$ play $\sigma_{-i}^{t+1}(h_p^t)$ with probability $\mathbf{P}_\sigma(h_p^t | h_i^t)$. The distribution of the action profile $\mathbf{a}_{-i}^{t+1}$ of players $-i$ at stage $t+1$ given the information h_i^t of player i is $\sum_{h^t} \mathbf{P}_\sigma(h_p^t | h_i^t) \sigma_{-i}^{t+1}(h_p^t)$, element of the set $\triangle A_{-i}$ of correlated distributions on A_{-i}.

Let $X = \otimes_{j \neq i} \triangle A_j$ be the set of independent probability distributions on A_{-i}. A *correlation system* is a probability distribution on X and we let $C = \triangle X$ be the set of correlation systems.

X is identified to a closed subset of $\triangle A_{-i}$ and thus C is compact with respect to the weak* topology.

We study the evolution of the uncertainty of player i concerning the public history along the play of the repeated game.

Given a correlation system c and $a_i \in A_i$, let $(\mathbf{x}, \mathbf{s}_p)$ be a random variable with values in $X \times S_p$ such that the law of $\mathbf{x}$ is c and the law of $(\mathbf{s}_p, s_i)$ given $\{\mathbf{x} = x\}$ is $q^{x, a_i}(\cdot)$. The entropy variation is:

$$\Delta H(c, a_i) = H(\mathbf{s}_p \mid \mathbf{x}) - H(\mathbf{s}_i)$$

The entropy variation is the difference between the entropy gained by the team and the entropy lost. The entropy gain is the conditional uncertainty contained in $\mathbf{s}_p$ given $\mathbf{x}$; the entropy loss is the entropy of $\mathbf{s}_i$ which is observed by player i. If $\mathbf{x}$ is finite, from the additivity formula:

$$H(\mathbf{x}, \mathbf{s}_p) = H(\mathbf{x}) + H(\mathbf{s}_p \mid \mathbf{x}) = H(\mathbf{s}_i) + H(\mathbf{x}, \mathbf{s}_p \mid \mathbf{s}_i)$$

and therefore,

$$\Delta H(c, a_i) = H(\mathbf{x}, \mathbf{s}_p \mid \mathbf{s}_i) - H(\mathbf{x})$$

The entropy variation is thus the new entropy of the information possessed by the team and not by i minus the initial entropy.

Now we define, given a correlation system c, the payoff obtained when player i plays a best reply to the expected distribution on A_{-i}. Given a correlation system c, the distribution of the action profile for the team is $x_c \in \triangle A_{-i}$ such that for each $a_{-i} \in A_{-i}$, $x_c(a_{-i}) = \int_X x_{-i}(a_{-i}) dc(x)$. The best response payoff for player i against c is $\pi(c) = \max_{a_i \in A_i} g_i(x_c, a_i)$, and let $B_i(c) = \operatorname{argmax} g_i(x_c, \cdot) \subseteq A_i$.

Consider the set of feasible vectors $(\Delta H(c, a_i), \pi(c))$ where $a_i \in B_i(c)$ in the (entropy variation, payoff) plane:

$$V = \{(\Delta H(c, a_i), \pi(c)) \mid c \in C, a_i \in B_i(c)\}$$

Lemma 10 *V is compact.*

Proof. Since $\mathbf{s}$ is independent of $\mathbf{x}$ conditionally on $\mathbf{a}$, the additivity formula gives $H(\mathbf{a}, \mathbf{s} | \mathbf{x}) = H(\mathbf{a} | \mathbf{x}) + H(\mathbf{s} | \mathbf{a})$ and the entropy variation is:

$$\Delta H(c, a_i) = H(\mathbf{a} | \mathbf{x}) + H(\mathbf{s} | \mathbf{a}) - H(\mathbf{s})$$

¿From the definitions of entropy and conditional entropy, recalling that the law of $\mathbf{a}$ given $\{\mathbf{x} = x\}$ is x:

$$\Delta H(c) = \int_X H(x)dc(x) + \sum_a x_c(a)H(q(\cdot|a)) - H(\sum_a x_c(a)q(\cdot|a))$$

which is clearly a continuous function of c. ΔH and π are thus continuous on the compact set C so V is compact. ■

Define w as the lowest payoff associated to a convex combination of correlations systems under the constraint that the average entropy variation is non-negative:

$$w = \inf\{x_2 \in \mathbb{R} \mid (x_1, x_2) \in \operatorname{co} V, x_1 \geq 0\}$$

For every correlation system c such that $\mathbf{x}$ is a.s. constant, $\Delta H(c) \geq 0$ thus V intersects the half-plane $\{x_1 \geq 0\}$. Since V is compact, so is its convex hull and the supremum is indeed a maximum. Techniques for computing the set V are developed by Gossner et *al.*, [GLT06]. The set V need not be convex (see an example in Goldberg [Gol03]) and the supremum in the definition of w above might not be achieved by a point in V but by a convex combination involving two points of V with positive weights.

It is convenient to express the number w through the boundary of $\operatorname{co} V$. Define for each real number h:

$$u(h) = \min\{\pi(c) \mid (c \in C, \Delta H(c) \geq h\} = \min\{x_2 \mid (x_1, x_2) \in V, x_1 \geq h\}$$

Since V is compact, $u(h)$ is well defined. Let $\operatorname{cav} u$ be the least concave function pointwise greater than u. Then:

$$w = \operatorname{cav} u(0)$$

Indeed, u is upper-semi-continuous, non-increasing and the hypograph of u is the comprehensive set $V^* = V - \mathbb{R}^2_+$ associated to V. This implies that $\operatorname{cav} u$ is also non-increasing, u.s.c. and its hypograph is $\operatorname{co} V^*$.

Here we prove that for every strategy of the team, if player i plays stage-best replies, the average vector of (payoffs, entropy variation) generated belongs to V. This later implies that no strategy for the team can guarantee a better payoff than w.

Definition 3 *Let σ_{-i} be a public strategy for the team, and define inductively $\sigma_{i,\sigma_{-i}}$ as the strategy of player i that plays stage-best replies to σ_{-i}:*

At stage 1, $\sigma_{i,\sigma_{-i}} \in \operatorname{argmax}_{a_i} g_i(\sigma_{-i}(\emptyset), a_i)$ where $\emptyset$ is the null history that starts the game. Assume that $\sigma_{i,\sigma_{-i}}$ is defined on histories of length less that t. For every history h_i^t of player i, let $x^{t+1}(h_i^t) \in \triangle A_{-i}$ be the distribution of the action profile of the team at stage $t+1$ given h_i^t (arbitrary if h_i^t has zero probability) and select $\sigma_{i,\sigma_{-i}}(h_i^t)$ in $\operatorname{argmax}_{a_i} g_i(x^{t+1}(h_i^t), a_i)$.

It is convenient to consider the zero-sum game in which player i's stage-game payoff given action profile α is $g_i(\alpha)$, as in the original game, while all the other player's common payoff is $-g_i(\alpha)$. We let $\gamma_n(\sigma)$ denote player i's payoff in the finitely repeated game with n repetitions, given strategy profile σ (in the finitely repeated game), and the value of this finitely repeated game is denoted v_n. Similarly, let $\gamma_\delta(\sigma)$ denote player i's payoff in the infinitely repeated game with discount factor δ, given strategy profile σ, and the value of this infinitely repeated game is denoted v_δ.

Lemma 11 *For every σ_{-i}, $\sigma_{i,\sigma_{-i}}$ defends w in every n-stage game, i.e. for every σ_{-i}, n:*

$$\gamma_n(\sigma_{-i}, \sigma_{i,\sigma_{-i}}) \geq w$$

Hence for each n, $v_n \geq w$.

Proof. The proof follows the same lines as some previous papers using entropy methods (see e.g., [NO99],[NO00], [GV02] and [GT06]). Fix a profile of public strategies σ_{-i} for the team and let $\sigma_i = \sigma_{i,\sigma_{-i}}$. Let $\mathbf{s}_p^t, \mathbf{s}_i^t$ be the random signals to the team and to player i under P_{σ_{-i},σ_i}, $\mathbf{h}_p^t = (\mathbf{s}_p^1, \ldots, \mathbf{s}_p^t)$ and $\mathbf{h}_i^t = (\mathbf{s}_i^1, \ldots, \mathbf{s}_i^t)$ be the public history and the history of player i after stage t. Let $\mathbf{x}^t = \sigma_{-i}^t(\mathbf{h}_i^{t-1})$ and $c^t(\mathbf{h}_i^{t-1})$ be the distribution of $\mathbf{x}_m^t$ conditional on $\mathbf{h}_i^{t-1}$ i.e. $c^t(\mathbf{h}_i^{t-1})$ is the correlation system at stage t after history $\mathbf{h}_i^{t-1}$. Under $\sigma = (\sigma_{-i}, \sigma_i)$, the expected payoff to player i at stage t given $\mathbf{h}_i^t$ is $\max_{a_i} g_i(\mathbf{E}_\sigma[\mathbf{x}^t|\mathbf{h}_i^{t-1}], a_i) = \pi(\mathbf{c}^t)$ from the definition of σ_i and thus $\gamma_n(\sigma) = \mathbf{E}_\sigma[\frac{1}{n}\sum_{m=1}^n \pi(\mathbf{c}_m)]$.

$H^t = H(\mathbf{h}_p^t \mid \mathbf{h}_i^t)$ is the expected entropy of the secret information to the team after stage t. From the additivity of entropies:

$$\begin{aligned} H(\mathbf{s}_p^1, \ldots, \mathbf{s}_p^t|\mathbf{h}_i^t) &= H(\mathbf{s}_i^t|\mathbf{h}_i^{t-1}) + H^t \\ &= H^{t-1} + H(\mathbf{s}_p^t|\mathbf{h}_p^{t-1}) \end{aligned}$$

Thus,

$$\begin{aligned} H^t - H^{t-1} &= H(\mathbf{s}_p^t|\mathbf{h}_p^{t-1}) - H(\mathbf{s}_i^t|\mathbf{h}_i^{t-1}) \\ &= H(\mathbf{s}_p^t|\mathbf{x}^t, \mathbf{h}_i^{t-1}) - H(\mathbf{s}_i^t|\mathbf{h}_i^t) \\ &= \mathbf{E}_{\sigma,\tau}\Delta H(c^t(\mathbf{h}_i^{t-1}), \mathbf{a}_i^t) \end{aligned}$$

Then:

$$\sum_{m=1}^t \mathbf{E}_{\sigma,\tau}\Delta H(c^m(\mathbf{h}_i^{m-1}), \mathbf{a}_i^m) = H_n \geq 0$$

Therefore the vector $(\frac{1}{t}\sum_{m=1}^t \mathbf{E}_{\sigma,\tau}\Delta H(c^m(\mathbf{h}_i^{m-1}), \mathbf{a}_i^m), \gamma_n(\sigma,\tau))$ is in $\operatorname{co} V \cap \{x_1 \geq 0\}$. ■

Corollary 12 *For every σ^{-i} the strategy $\sigma^i_{\sigma^{-i}}$ defends w in every discounted game, i.e. for every σ^{-i}, λ:*

$$\gamma_\delta(\sigma^{-i}, \sigma^i_{\sigma^{-i}}) \geq w$$

Hence for every δ, $v_\delta \geq w$.

5.2.3 The case of ε-dependence and ρ-identifiability

Here we prove that under ε-dependence, little secret correlation can be generated per stage for the team. We also show that, under the ρ-identifiability assumption, secret correlation, if utilized, dissipates. Relying on the previous bound on minmax payoffs, we conclude the proof of the main theorem.

We compare w with the maximum payoff the team could obtain in a modified game where $s_p = s_i$, but the team could use an entropy of ε_h at each stage. If $s_p = s_i$, $\Delta H(c, a_i) = -I(\mathbf{x}, \mathbf{s}_i)$, and the function u' playing the role of u is:

$$u'(h) = \min\{\pi(c)|(c, a_i) \;\; s.t. \;\; I(\mathbf{x}; s_i) \leq h\}$$

Lemma 13 *For every $\varepsilon_h > 0$, there exists $\varepsilon > 0$ such that if the monitoring is ϵ-dependent, for every (c, a_i): $\Delta H(c, a_i) \leq \varepsilon_h - I(\mathbf{x}; \mathbf{s}_i)$. In particular, $u(h) \geq u'(h + \varepsilon_h)$ for all h.*

Proof. First note that $\Delta H(c, a_i) + I(\mathbf{x}; s_i) = H(\mathbf{s}_p|\mathbf{x}, \mathbf{s}_i) = \int_x \sum_{s_i} q_i^{x,a_i}(s_i) H(q^{x,a_i}(\mathbf{s}_p|s_i))dc(x)$. For every x, s_i,

$$H(q^{x,a_i}(\mathbf{s}_p|s_i)) = H(\lambda(q_i^{x,a_i}(s_i))) + (1 - \lambda(q_i^{x,a_i}(\cdot|s_i)))H(r_{-i}(\cdot|s_i))$$

There exists $S_i' \subseteq S_i$ such that $q_i^{x,a_i}(S_i') \geq 1-\varepsilon$ and that for $s_i \in S_i'$, $d(q_i^{x,a_i}(\cdot|s_i), \otimes_{j \neq i} \triangle S_j) \leq \varepsilon$ and in particular $\lambda(q_i^{x,a_i}(\cdot|s_i)) \geq \eta(\varepsilon)$ by Lemma 7. For $s_i \in S_i'$, $H(\lambda(q_i^{x,a_i}(s_i))) \leq \max\{H(\eta(\varepsilon)), 1\}$ and for all $s_i \in S_i'$, $H(r_{-i}(\cdot|s_i)) \leq \log_2(S_{-i} \times A_{-i})$. Hence

$$\Delta H(c, a_i) + I(\mathbf{x}; s_i) \leq \max(\eta(\varepsilon), 1) + \log_2(S_{-i} \times A_{-i})$$

which proves the first part of the lemma. The second part now follows from the definitions of u and u'. ■

Let us finally define the function u'' that corresponds to u' when $s_i = a_{-i}$, namely when i has perfect observation of the actions of the opponents.

$$u''(h) = \min\{\pi(c), (c, a_i) \ \ s.t. \ \ I(\mathbf{x}; \mathbf{a}_{-i}) \leq h\}$$

Lemma 14 *There exists a continuous function α such that $\alpha(0) = 0$ and $u'(h) \geq u''(\alpha(h))$ for all h.*

Proof. Let $\alpha(h) = \max\{I(\mathbf{x}; \mathbf{a}_{-i})|(c, a_i) \ \ s.t. \ \ I(x; s_i) \leq h\}$. It follows from Carathéodory's theorem that the we can restrict to c of support of size at most 3 in the definition of α, and the sup is actually a max. Assume now c has finite support, $I(\mathbf{x}; \mathbf{s}_i) = 0$ implies that $\mathbf{s}_i$ and $\mathbf{m}_{-i}$ are independent, therefore that c is a mass unit and that $I(\mathbf{x}; \mathbf{a}_{-i}) = 0$. Hence $\alpha(0) = 0$. The map α is continuous by the maximum principle. That $u'(h) \geq u''(\alpha(h))$ for all h follows from the definitions of u', u'' and α. ■

We now complete the proof of our main result.

Proof. [Proof of Theorem 6] From Lemma 11 and corollary 12, $v_n \geq w$ and $v_\lambda \geq w$, where by Lemmata 13 and 14, $w \geq \text{cav}\,(u'' \circ \alpha)(\varepsilon_h)$. The result follows since $\text{cav}\,(u'' \circ \alpha)$ is continuous and $\varepsilon_h \to 0$ as $\varepsilon \to 0$. ■

References

[AS76] R. J. Aumann and L. S. Shapley. Long-term competition – a game theoretic analysis. preprint, 1976.

[Bla51] D. Blackwell. Comparison of experiments. In *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, pages 93–102. University of California Press, 1951.

[BO02] V. Bhaskar and I. Obara. Belief-based equilibria in the repeated prisoners' dilemma with private monitoring. *Journal of Economic Theory*, 102:40–70, 2002.

[CFM64] P. Cartier, J.M.G. Fell, and P.-A. Meyer. Comparaison des mesures portées par un ensemble convexe compact. *Bulletin de la Société Mathématique de France*, 92:435–445, 1964.

[EV02] J.C. Ely and J. Välimäki. A robust folk theorem for the prisoner's dilemma. *Journal of Economic Theory*, 102:84–106, 2002.

[FLM94] D. Fudenberg, D. K. Levine, and E. Maskin. The folk theorem with imperfect public information. *Econometrica*, 62:997–1039, 1994.

[FM86] D. Fudenberg and E. Maskin. The folk theorem in repeated games with discounting or with incomplette information. *Econometrica*, 54:533–554, 1986.

[Fri71] J. Friedman. A noncooperative equilibrium for supergames. *Review of Economic Studies*, 38:1–12, 1971.

[FT91] D. Fudenberg and J. Tirole. *Game Theory*. Cambridge, Massachusetts. The MIT Press, 1991.

[GLT06] O. Gossner, R. Laraki, and T. Tomala. On the optimal use of coordination. *Mathematical Programming B.*, to appear, 2006.

[Gol03] Y. Goldberg. On the minmax of repeated games with imperfect monitoring: A computational example. Discussion Paper Series 345, Center the Study of Rationality, Hebrew University, Jerusalem, 2003.

[GT04] O. Gossner and T. Tomala. Secret correlation in repeated games with signals. Document de travail 04-03, CERAS, Paris, 2004.

[GT06] O. Gossner and T. Tomala. Empirical distributions of beliefs under imperfect observation. *Mathematics of Operations Research*, 31(1):13–30, 2006.

[GV02] O. Gossner and N. Vieille. How to play with a biased coin? *Games and Economic Behavior*, 41:206–226, 2002.

[HO06] J. Hörner and W. Olszewski. The folk theorem with private almost-perfect monitoring. *Econometrica*, 2006.

[Jor93] J.S. Jordan. Three problems in learning mixed strategy equilibria. *Games and Economic Behavior*, 5:368–386, 1993.

[Kan02] M. Kandori. Introduction to repeated games with private monitoring. *Journal of Economic Theory*, 102:1–16, 2002.

[Leh90] E. Lehrer. Nash equilibria of n player repeated games with semi-standard information. *International Journal of Game Theory*, 19:191–217, 1990.

[MM] G. Mailath and S. Morris. Coordination failure in repeated games with almost public monitoring. *Theoretical Economics*.

[MM02] G. Mailath and S. Morris. Repeated games with almost-public monitoring. *Journal of Economic Theory*, 102:189–229, 2002.

[MW98] D. Moreno and J. Wooders. An experimental study of communication and coordination in noncooperative games. *Games and Economic Behavior*, 24:47–76, 1998.

[NO99] A. Neyman and D. Okada. Strategic entropy and complexity in repeated games. *Games and Economic Behavior*, 29:191–223, 1999.

[NO00] A. Neyman and D. Okada. Repeated games with bounded entropy. *Games and Economic Behavior*, 30:228–247, 2000.

[Pic02] M. Piccione. The repeated prisoner's dilemma with imperfect private monitoring. *Journal of Economic Theory*, 102:70–84, 2002.

[Rub77] A. Rubinstein. Equilibrium in supergames. Center for Research in Mathematical Economics and Game Theory, Research Memorandum 25, 1977.

[Sek97] T. Sekiguchi. Efficiency in repeated prisoner's dilemma with private monitoring. *Journal of Economic Theory*, 76:345–361, 1997.

[Str65] V. Strassen. The existence of probability measures with given marginals. *Annals of Mathematical Statistics*, 36:423–439, 1965.

[vSK97] B. von Stengel and D. Koller. Team max min equilibria. *Games and Economics Behavior*, 21:309–321, 1997.