The Folk Theorem for Games with Private Almost-Perfect Monitoring*

Johannes Hörner† Wojciech Olszewski‡

January 2005

Abstract

We prove the folk theorem for discounted repeated games under private, almost-perfect monitoring. Our result covers all finite, n-player games satisfying the usual full-dimensionality condition. Mixed strategies are allowed in determining the individually rational payoffs. We assume no cheap-talk communication between players and no public randomization device.

KEYWORDS: Repeated games, private monitoring, folk theorem.

1 Introduction

The central result of the literature on discounted repeated games is probably the folk theorem (Fudenberg and Maskin (1986)): with only two players, or when a full dimensionality condition holds, any feasible payoff vector Pareto-dominating the minmax point of the stage game is achieved by some subgame-perfect equilibrium of the infinitely repeated game provided that the players are sufficiently patient. Under some identifiability conditions, this result has been subsequently generalized by Fudenberg, Levine and Maskin (1994) to the case in which players do not observe the chosen action profile, but only a public signal that is a stochastic function of the action profile. For that purpose it suffices to consider a restricted class of sequential

*We thank seminar audiences at the University of Chicago, the London School of Economics, New York University, Northwestern University, Penn State University, Princeton University, Rutgers University, University College London, Université de Toulouse and WZB, Berlin, the 15th Summer Festival on Game Theory at Stony Brook, the 2nd World Congress of the Game Theory Society in Marseille, and Paul Heidhues for his discussion.

†Department of Managerial Economics and Decision Sciences, Kellogg School of Management and HEC School of Management, Paris. j-horner@kellogg.northwestern.edu.

‡Department of Economics, Northwestern University. wo@northwestern.edu.

equilibria. In *perfect public equilibria* (PPE), players' continuation strategy only depends on the public history, that is, on the history of public signals. The analysis of PPE is tractable because after any history the continuation strategies correspond to an equilibrium in the original game, so that the set of PPE payoffs can be characterized by techniques borrowed from dynamic programming (see Abreu, Pearce and Stacchetti (1990)).[1]

Thus, common knowledge of relevant aspects of players' histories plays an essential role in the proofs of the folk theorem so far. This sort of common knowledge is missing in games with private monitoring. In such games, each player only observes a private signal that is a stochastic function of the action profile. If, for each action profile, the signals of all players are perfectly correlated, then the monitoring is public, and if moreover the signals are perfectly correlated with the action profile, the monitoring is perfect. Yet, in general, signals are neither perfect nor public, so that players share no public information to coordinate continuation play. This paper shows that the folk theorem is robust. It remains valid under the standard full-dimensionality assumption, provided only that the private signals are sufficiently close to perfect. In particular, signals are not restricted to be almost-public or conditionally independent.

More specifically, take any finite n-player game whose set of feasible, individually rational payoffs has non-empty interior V^*, where the individually rational payoffs are determined by considering (independent) mixed strategies. Consider the canonical signal space, in which a player's set of signals is the set of action profiles of its opponents. More general signal spaces are discussed in Section 5. Monitoring is ε-*perfect* if, for any player i, under any action profile a, player i obtains signal $\sigma_i = a_{-i}$ with probability at least $1-\varepsilon$. The parameter ε is the *noise level*. When $\varepsilon = 0$, monitoring is perfect. Payoffs are discounted at common factor $\delta \in (0,1)$. No public randomization or communication device is assumed. Given discount factor δ, denote by $E(\delta,\varepsilon)$ the set of average payoff vectors in the repeated game that are sequential equilibrium payoffs for all ε-perfect monitoring structures. This paper shows that:

$$\forall_{v\in V^*} \quad \exists_{\bar{\delta}<1,\bar{\varepsilon}>0} \quad \forall_{(\delta,\varepsilon)\in\left(\bar{\delta},1\right)\times[0,\bar{\varepsilon})} \quad v \in E(\delta,\varepsilon).$$

Observe that the result does not posit any particular order of limits, as it holds for a joint neighborhood of discount factors and noise levels. In addition, the result states that the payoff vector v is exactly achieved, not only approximated.

There are several related contributions. Lehrer (1990) obtains efficient equilibria while considering time-average payoffs, while Fudenberg and Levine (1991) require approximate optimization. The equilibrium strategies proposed in these papers are no longer equilibrium strategies once discounting and exact optimization are introduced.

Compte (1998), Kandori and Matsushima (1998), Aoyagi (2002) and Fudenberg and Levine (2002) prove versions of the folk theorem while allowing players to communicate. While a realistic

[1]In general, however, the set of sequential equilibrium payoffs is strictly larger than the set of PPE payoffs. See Kandori and Obara (2004) for details.

assumption in many applications, communication reintroduces an element of public information that is somewhat at odds with the motivation of private monitoring as a robustness test outlined above. Mailath and Morris (2002) prove a folk theorem for almost-perfect monitoring, assuming in addition that monitoring is also *almost-public*.

Sekiguchi (1997) achieves the efficient outcome and Bhaskar and Obara (2002) establishes the folk theorem, under almost-perfect monitoring, for the special case of the two-player prisoner's dilemma. They isolate a set of (continuation) strategies closed under best-response: for any relevant belief a player may have about his opponent's continuation strategy (within that set), *some* strategy within the set is a best-response. Using a different approach, Ely and Välimäki (2002) and Piccione (2002) prove the folk theorem under almost-perfect monitoring for the two-player prisoner's dilemma. They isolate a set of (continuation) strategies satisfying a stronger property: for any belief a player may have about his opponent's continuation strategy (within that set), *any* strategy within that set is a best-response. This approach has been further used by Matsushima (2004) to extend the two-player prisoner dilemma's folk theorem from the case of almost-perfect monitoring to the case of conditionally independent, but not necessarily almost perfect, monitoring. Finally, Yamamoto (2004) shows, by modifying the construction of Ely and Välimäki (2002) and Matsushima (2004), that the efficient outcome can be achieved in a class of N-player games, similar in structure to the prisoner's dilemma, under almost-perfect as well as conditionally independent monitoring.

While the first, *belief-based*, approach is more general than the second *belief-free* approach, it appears less tractable and has not been generalized so far to other stage games. The belief-free approach has been studied more generally by Ely, Hörner and Olszewski (2004), which characterizes the set of payoffs that can be achieved using sequential equilibria satisfying this property. For many stage games, this set of payoffs is larger than the convex hull of static Nash equilibrium payoffs, but for "almost all" stage games, it fails to yield the folk theorem even under almost-perfect monitoring.

Although the equilibria studied in this paper are not belief-free, they retain some essential features of belief-free equilibria. To get some insight into the construction, consider the case of two players. In each consecutive block of T periods, players use one of two strategies of the T-finitely repeated stage game. The length T is chosen so that the average payoff over the horizon T of each of the four resulting strategy profiles surrounds the average payoff vector v to be achieved overall: if a player uses one of his two strategies, his opponent is guaranteed to receive more than v_i, independently of which of the two strategies he uses himself; if he uses the other strategy, his opponent gets less than v_i, no matter which strategy he uses, among all strategies of the T-finitely repeated game. Therefore, within each block, a player is not indifferent over his opponent's choice of strategy. By choosing appropriately the probability with which a player uses one or the other strategy at the beginning of each non-initial block (*the transition probabilities*), as a function of his *recent* history and of his *recent* strategy (that is, of his private history and strategy within the previous block), we ensure that players are indifferent, at the beginning of each block, across their two strategies, and weakly prefer those two strategies to all others. Further, by choosing appropriately the probability with which a player uses one or the other strategy in the initial block, we ensure that the payoff vector v is achieved.

This guarantees that beliefs are irrelevant *at the beginning* of each block, and more generally, that only beliefs *about the recent history* matters. Belief-free equilibria obtain for the choice $T = 1$.[2] One of the insights of the general construction is that the special features of the prisoner's dilemma payoff matrix ensuring that the folk theorem obtains with $T = 1$ obtain for any stage game, and any payoff to be achieved for that stage game, provided one chooses $T \geq 1$ appropriately.[3]

Sequential rationality poses several difficulties when $T > 1$. After recent histories that are consistent with both players having only observed correct signals, a player's belief about his opponent's recent history has a tractable structure: when the noise level is sufficiently small, he assigns probability almost one to his opponent having observed the same recent history, regardless of the fine details of the monitoring structure. This is not the case, however, for all other recent histories (*erroneous histories*), as in such events, his posterior may dramatically vary with small differences in the relative likelihood of incorrect signals. As a player's intertemporal incentives depend on his opponent's recent history, specifying best-responses after such histories is much less tractable. Worse, a player's belief, and thus best-response, after such a recent history may in principle depend on his belief about his rival's recent strategy, and this belief typically depends on his whole private history, not only his recent one. This would imply that for some types of recent histories, best-response could not be specified independently of the entire history, destroying the recursive structure of our construction.

This problem is circumvented as follows. For one of his two strategies (the one yielding lower payoffs to his rival), it is possible to specify a player's transition probabilities, at the end of the block, so as to guarantee that his opponent is indifferent over *all* strategies within the block, not only over the two strategy he actually chooses from. This implies that, while computing his best-response after any recent history, a player may condition on his opponent playing the other strategy (independently of the beliefs he actually has about the recent strategy used by his opponent) to determine, given his recent history, the probability distribution over his opponent's recent histories. As this best-response depends on what this other strategy specifies, for each recent history, as well as on the corresponding transition probabilities at the end of the block, it follows that this strategy, and the transition probabilities that go along, must be determined jointly, which is achieved here by applying Kakutani's fixed point theorem.

This guarantees that optimal play after recent histories is indeed a function of that recent history only. It leaves the play not explicitly specified after erroneous histories (in particular, we do not know the payoffs contingent on such histories), but it does not pose any major problem. Roughly because, by choosing transition probabilities that yield lower payoffs contingent on

[2] More precisely, this is the case for belief-free equilibria using a constant regime (see Ely, Hörner and Olszewski (2004)).

[3] Thus, T-period blocks do not serve the purpose of statistical discrimination between actions, as in Radner (1986) or Matsushima (2004), but the purpose of enlarging the set of payoffs generated by belief-free equilibria.

erroneous histories, players can be given incentives not to "trigger" such histories, and since such histories appear with small probability, total payoffs are not affected much by the play on erroneous histories.

The case with more than two players creates additional challenges, related to coordination and dimensionality issues. The construction in that case is more intricate, but we postpone discussion of it to the relevant section. While the construction for $n > 2$ also works for $n = 2$, it is more natural to introduce this construction by first considering the case of two players.

Section 2 introduces the notation and states the results. Section 3 presents the construction for two players, first under perfect monitoring, and then under imperfect private monitoring. Section 4 presents the construction for $n > 2$ players, following the same two steps as for $n = 2$. Finally, Section 5 relaxes the restriction on the signal set and offers concluding comments.

2 Notation and result

Consider the following finite n-person game. Each player $i = 1, ..., n$ has a (finite) action set A_i and a (finite) set of signals Σ_i. Without loss of generality, assume that A_i contains at least two elements, for all i. Throughout Sections 2 to 4, we maintain the assumption that $\Sigma_i = A_{-i}$, where $A_{-i} := A_1 \times \cdots \times A_{i-1} \times A_{i+1} \times \cdots \times A_n$. This assumption is convenient to measure the distance of a particular monitoring structure from perfect monitoring.

For each action profile $a \in A := A_1 \times \cdots \times A_n$, $m(\cdot \mid a)$ specifies a probability distribution over $\Sigma := \Sigma_1 \times \cdots \times \Sigma_n$. The collection of probability distributions over signal profiles $\{m(\cdot \mid a) : a \in A\}$ defines the *monitoring structure*. For each action profile $a \in A$, $m_i(\cdot \mid a)$ denotes the marginal distribution of $m(\cdot \mid a)$ over Σ_i. Thus, $m_i(\sigma_i \mid a)$ is the probability that player i receives signal $\sigma_i \in \Sigma_i$ under action profile $a \in A$.

We focus attention on the case in which the monitoring structure is close to perfect monitoring. Following Ely and Välimäki (2002), we formalize this notion as follows: for $\varepsilon \geq 0$, the monitoring structure $\{m(\cdot \mid a) : a \in A\}$ is ε-perfect if for each player and each action profile $a \in A$,

$$m_i(\sigma_i = a_{-i} \mid a) \geq 1 - \varepsilon.$$

That is, under any action profile, the probability that a player observes an erroneous signal does not exceed ε. The perfect monitoring structure is a special case that obtains for $\varepsilon = 0$. Observe that this definition is stated in terms of marginal distributions only. Therefore, while this definition is consistent with almost-public or conditionally independent signals, it does not impose any such restriction. We do not impose any full-support restriction either.

Mixed actions are unobservable. For any finite set W, let $\triangle W$ denote the set of probability distributions over W. With some abuse of notation, we use $\triangle A := \triangle A_1 \times \cdots \times \triangle A_n$ to denote the set of (independent) mixed action profiles. Similarly, $\triangle A_{-i} := \triangle A_1 \times \cdots \times \triangle A_{i-1} \times \triangle A_{i+1} \times \cdots \times \triangle A_n$. No public randomization device is assumed.

Player i's realized payoff in the stage game, $u_i : A_i \times \Sigma_i \to \mathbb{R}$, is a function of his action and signal alone, so that his expected payoff $g_i : A \to \mathbb{R}$ is given by

$$g_i(a) = \sum_{\sigma_i \in \Sigma_i} m_i(\sigma_i \mid a)\, u_i(a_i, \sigma_i).$$

The domain of g_i is extended to mixtures $\alpha \in \triangle A$ in the usual manner:

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

where $\alpha(a)$ denotes the probability assigned to action profile a by the mixture $\alpha \in \triangle A$. Observe that repeated games with public monitoring are special cases of this formulation. If signals are perfectly correlated with each other, we obtain a game with imperfect public monitoring, while under the perfect monitoring structure, we obtain a standard game with perfect monitoring.

Players share a common discount factor $\delta < 1$. All repeated game payoffs, both infinite and finite, 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 $1-\delta$, sometimes referred to as the *average*, or *normalized payoffs. Total, or unnormalized payoffs* are payoffs that are discounted, but not normalized.

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

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

Choose $\alpha_{-i}^* \in \triangle A_{-i}$ so that

$$v_i^* = \max_{a_i \in A_i} g_i\left(a_i, \alpha_{-i}^*\right).$$

The action α_{-i}^* is the (not necessarily unique) minmax action against player i, and v_i^* is the smallest payoff that the other players can keep player i below in the static game.[4]

Let:

$$U := \{(v_1, \ldots, v_n) \mid \exists a \in A, \forall i,\ g_i(a) = v_i\},$$
$$V := \text{Convex Hull of } U,$$

and

$$V^* := \text{ Interior of } \{(v_1, \ldots, v_n) \in V \mid \forall i,\ v_i > v_i^*\}.$$

The set V consists of the feasible payoffs, and V^* is the set of payoffs in the interior of V that strictly Pareto-dominate the minmax point $v^* := (v_1^*, \ldots, v_n^*)$. We assume throughout that V^* is non-empty. Given discount factor δ, recall that $E(\delta, \varepsilon)$ is the set of average payoff vectors in

[4] Under some imperfect monitoring structures, it may be possible to keep player i's payoff even lower if $n \geq 3$, as signals may allow players $-i$ to correlate their actions without being observed by player i.

the repeated game that are sequential equilibrium payoffs for all ε-perfect monitoring structures. We can now state our main result.

Theorem 1:[5] *(The Folk Theorem) For any $(v_1, \ldots, v_n) \in V^*$, if players discount the future sufficiently little and the noise level is sufficiently small, there exists a sequential equilibrium of the infinitely repeated game where, for all i, player i's average payoff is v_i. That is,*

$$\forall_{v \in V^*} \quad \exists_{\bar{\delta}<1, \bar{\varepsilon}>0} \quad \forall_{(\delta,\varepsilon) \in (\bar{\delta},1) \times [0,\bar{\varepsilon})} \quad v \in E(\delta, \varepsilon).$$

The proof uses the following notations. A t-length (private) history for player i is an element of $H_i^t := (A_i \times \Sigma_i)^t$. A pair of t-length histories is denoted h^t. Such a pair is also referred to as a history. As (private) histories are always indexed by the relevant player, no confusion should arise. Each player's initial history is the null history, denoted $\emptyset$. Let H^t denote the set of all t-length histories, H_i^t the set of i's (private) t-length histories, $H = \cup_t H^t$ the set of all histories, and $H_i = \cup_t H_i^t$ the set of all (private) histories for i. A repeated-game (behavior) strategy for player i is a mapping $s_i : H_i \to \triangle A_i$. The mixed action prescribed by strategy s_i, given private history h_i^t is denoted $s_i[h_i^t]$, while the probability assigned to action any a_i by $s_i[h_i^t]$ is denoted $s_i[h_i^t](a_i)$. The set of all strategies of player i in the infinitely repeated game is denoted S_i, and a strategy profile is denoted $s \in S := S_1 \times \cdots \times S_n$. For any history $h_i^t \in H_i$, let $s_i | h_i^t$ denote the continuation strategy derived from s_i after history h_i^t, and $s_i | H_i'$ the restriction of s_i to the set of histories $H_i' \subset H_i$.

For $T \geq 1$, we shall also consider the game repeated T times (henceforth simply referred to as the finitely repeated game). The set of all t-length (private) histories of player i in the T-finitely repeated game is denoted by H_i^t, the set of all histories by $H_i^T = \cup_{t \leq T} H_i^t$ and the set of (behavior) strategies in the finitely repeated game by S_i^T. For $t \leq T$, we use the same notation for continuation strategies as in the case of the infinitely repeated game.

Three types of repeated game payoffs are considered. Given strategy profile $s \in S$, player i's payoff is denoted $U_i(s)$ in the infinitely repeated game. Given strategy profile $s \in S^T := S_1^T \times \cdots \times S_n^T$, player i's payoff is denoted $U_i^T(s)$ in the finitely repeated game. Finally, we shall consider the finitely repeated game augmented by a transfer $\pi_i : H_{i+1}^T \to \mathbb{R}$ at the end of the last period (identifying 1 and $n+1$). Given $\pi := (\pi_1, \ldots, \pi_n)$ and some history h_{i+1}^T, player i's payoff in this *auxiliary scenario* is defined as $U_i^A(h_{i+1}^T, \pi_i) := U_i^T(h_{i+1}^T) + (1-\delta)\delta^T \pi_i(h_{i+1}^T)$, and its definition extended to strategies $s \in S^T$ in the usual fashion. Let $U(s)$, $U^T(s)$ and $U^A(s, \pi)$ denote the corresponding payoff vectors. Continuation payoffs given some private history h_i^t are denoted $U_i(s \mid h_i^t)$, $U_i^T(s \mid h_i^t)$ and $U_i^A(s, \pi \mid h_i^t)$.

Given some strategy profile $s_{-i} \in S_{-i}^T := S_1^T \times \cdots \times S_{i-1}^T \times S_{i+1}^T \times \cdots \times S_n^T$ and transfer π_i, let $B_i(s_{-i}, \pi_i)$ denote the set of auxiliary scenario best-responses of player i. Finally, given a set of histories $H_i^E \subset H_i^T$, a strategy $s_{-i} \in S_{-i}^T$, a strategy $\bar{s}_i \in S_i^T$ and transfer π_i, let $B_i(s_{-i}, \pi_i \mid \bar{s}_i)$

[5] Theorem 1 does not rule out equilibrium payoffs outside V^* (see footnote 4), although we believe that player i's minmax payoff in the repeated game tends to his stage game minmax payoff v_i^* as $\varepsilon \to 0$. Also, we do not know whether the full dimensionality condition can be dropped with only two players.

denote the set of strategies that maximize player i's auxiliary-scenario payoff against s_{-i}, π_i among all strategies $s_i \in S_i^T$ such that $s_i \mid H_i^E = \overline{s}_i \mid H_i^E$.

By $B(v, \lambda)$, we mean the ball around payoff vector v of radius λ; by $\operatorname{co} W$, the convex hull of a set W, and by $\#W$, the cardinality of the finite set W.

3 Two-player Games

As discussed in the introduction, the construction retains some of the features of belief-free equilibria (see Ely, Hörner and Olszewski (2004)). Belief-free equilibria in two-player games are sequential equilibria for which, for every history $h^t = \left(h_i^t, h_{-i}^t\right)$, player i's continuation strategy $s_i|h_i^t$ is a best-response to his opponent's continuation strategy $s_{-i}|h_{-i}^t$; that is, given any private history h_i^t, player i would be willing to play according to $s_i|h_i^t$ even if he secretly learnt player $-i$'s private history h_{-i}^t. Belief-free equilibria assume away the problem of statistical inference, and can be analyzed by means of recursive techniques building on the concepts of self-generation due to Abreu, Pearce and Stacchetti (1990). In a belief-free equilibrium $s = (s_1, s_2)$, for each player i, there exists a sequence of subsets $\{\mathcal{A}_i^t\}_{t=0}^{\infty}$ of A_i (independent of histories, but possibly depending on calendar time) such that any strategy of player i that adheres to this sequence from period t on, that is, for which

$$\forall_{r \geq t}, \forall_{h_i^r} \quad s_i\left(h_i^r\right) \in \mathcal{A}_i^r,$$

is an optimal continuation strategy, independently of player $-i$'s history h_{-i}^t. Belief-free equilibria typically support a large set of payoffs, but in general this set is not large enough to establish the folk theorem for almost-perfect monitoring. A notable exception is the prisoner's dilemma. In particular, in the case of the prisoner's dilemma, any feasible and individually rational payoff vector that is Pareto-dominated by the efficient payoff vector can be obtained as a belief-free equilibrium payoff vector provided the noise level is small enough. Moreover, the corresponding sequence of subsets can be chosen to be constant: $\forall i$, $\forall t$, $\mathcal{A}_i^t = \mathcal{A}_i = A_i$. To see this, let $A_i := \{C, D\}$, and normalize the payoff vectors of (C, C) and (D, D) to $(1, 1)$ and $(0, 0)$ respectively. Consider perfect monitoring. Observe that, for every $v \in (0, 1)^2$, there exists actions α_{-i}^G and $\alpha_{-i}^B \in \triangle\mathcal{A}_{-i}$ such that:

$$\min_{\mathcal{A}_i} g_i\left(a_i, \alpha_{-i}^G\right) > v_i > \max_{A_i} g_i\left(a_i, \alpha_{-i}^B\right).$$

Indeed, pick $\alpha_{-i}^G = C$ and $\alpha_{-i}^B = D$. This means that, in any given period, player $-i$ has an action, within his set of optimal actions $\mathcal{A}_{-i}$, that forces his opponent's payoff below v_i no matter what he does, and another action, within his set of optimal actions $\mathcal{A}_{-i}$, that guarantees his opponent a payoff above v_i no matter which action his opponent chooses from his set $\mathcal{A}_i$. Therefore, player $-i$'s current choice of action can be used to give a high or a low flow payoff to player i. Further, by conditioning his future choice of action within $\mathcal{A}_{-i}$ on his current action and on the signal he observes in the current period, player $-i$ can make player i indifferent between his various actions within $\mathcal{A}_i$. The overall payoff v_i is then achieved by an initial (private) randomization between α_{-i}^G and α_{-i}^B.

This argument does not work for all other games. For example, if v_i is close to player i's minmax level, α_{-i}^B must be close to the minmax action α_{-i}^*, implying that $\mathcal{A}_{-i}$ must include the support of the minmax action. Yet there may very well be, within that support, actions that yield player $-i$ low payoffs (even payoffs below his own minmax level) no matter what player i does. In fact, the prisoner's dilemma is the only known example for which the construction works for all $v \in V^*$.

The starting point of our analysis is that this convenient payoff structure can be recovered, for any stage game and any payoff $v \in V^*$, if one considers the normal form of the finitely repeated game, for some finite number of repetitions that depends on the stage game and v. That is, given v, we exhibit T, and for each player, a set of strategies $\mathcal{S}_i \subset S_i^T$, with two distinguished elements s_{-i}^G, $s_{-i}^B \in \mathcal{S}_{-i}$ such that

$$\min_{\mathcal{S}_i} U_i^T\left(s_i, s_{-i}^G\right) > v_i > \max_{S_i^T} U_i^T\left(s_i, s_{-i}^B\right).$$

Viewing the infinitely repeated game as the infinite repetition of the finitely repeated game, it is then possible, abstracting from the issues related to sequential rationality within the finitely repeated game, to replicate the construction of belief-free equilibria. In any given block of T periods, player $-i$ has a strategy, within his set of optimal actions $\mathcal{S}_{-i}$, that forces his opponent's average payoff in the block below v_i no matter which strategy he chooses from S_i^T, and another strategy, within his set of optimal actions $\mathcal{S}_{-i}$, that guarantees his opponent an average payoff in the block above v_i no matter which strategy his opponent chooses from his set $\mathcal{S}_i$. Therefore, player $-i$'s choice of strategy within the current block can be used to give a high or a low flow payoff to player i . Further, by conditioning his future choice of strategy within $\mathcal{S}_{-i}$ on his current strategy and on the history he observes within the current block, player $-i$ can make player i indifferent between his various strategies within $\mathcal{S}_i$. The payoff v_i is then achieved by an initial (private) randomization between s_{-i}^G and s_{-i}^B.

Thus, the time horizon of the infinitely repeated game is divided into T-period blocks, and for each player i there exists a subset $\mathcal{S}_i$ of S_i^T, such that any strategy which from the start of block n, picks within each future block an element of $\mathcal{S}_i$, is an optimal continuation strategy at the beginning of the block, independently of player $-i$'s history. That is, let $s_i^n \mid h_i^{nT}$ denote the restriction of $s_i \mid h_i^{nT}$ to the $(n+1)$-st block. The strategy s_{-i} is such that, if s_i is any strategy with the property that

$$\forall_{m \geq n}, \forall_{h_i^{mT}} \quad s_i^m \mid h_i^{mT} \in \mathcal{S}_i,$$

for any history h_i^{mT} following history h_i^{nT}, then $s_i \mid h_i^{nT}$ is an optimal continuation strategy at the beginning of the block, independently of h_{-i}^{nT}.

If $T > 1$, such a *block equilibrium* typically fails to be belief-free. There are histories within a block such that a player's set of optimal actions depends on the history. However, it depends only on the *recent* history (the finite, terminal segment of the player's private history, corresponding to the actions taken and signals observed within the current block).

Because such an equilibrium is not belief-free, sequential rationality within each block raises several difficulties, as mentioned in the introduction. These difficulties affect the specific way $\mathcal{S}_i$, s_i^G and s_i^B are defined. It is convenient to first define $\mathcal{S}_i$, s_i^G and s_i^B under perfect monitoring. Yet the reasons behind some peculiarities of those definitions will only become clear in the following subsection.

3.1 Perfect Monitoring

Fix a stage game, and let v be any payoff vector in V^*. To construct a subgame-perfect equilibrium achieving v, provided players are sufficiently patient, it is necessary to first define some payoff vectors and action profiles.

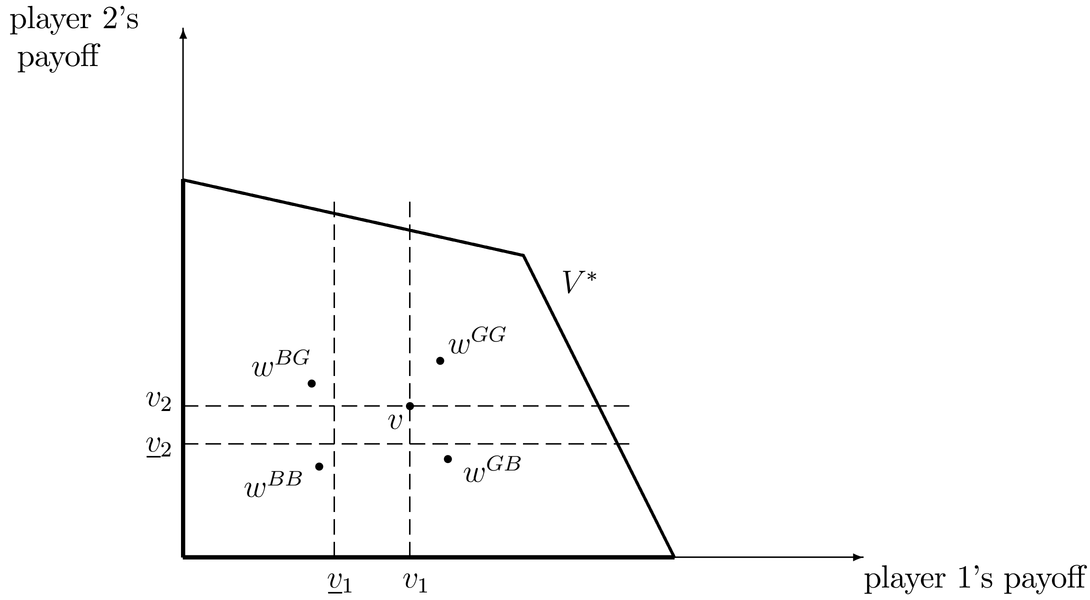

Figure 1

Take four payoff vectors $w^{GG}, w^{GB}, w^{BG}, w^{BB}$ in V^* such that: (i) $w_1^{GG} > v_1$ and $w_2^{GG} > v_2$; (ii) $w_1^{GB} > v_1$ and $w_2^{GB} < v_2$; (iii) $w_1^{BG} < v_1$ and $w_2^{BG} > v_2$; (iv) $w_1^{BB} < v_1$ and $w_2^{BB} < v_2$. That is, these four payoff vectors surround the payoff vector v. More precisely, there exists $\underline{v}_i < v_i$ such that the interior of $\mathrm{co}(\{w^{GG}, w^{GB}, w^{BG}, w^{BB}\})$ contains the rectangle

$$[\underline{v}_1, v_1] \times [\underline{v}_2, v_2].$$

See Figure 1. Assume that there exists pure action profiles $a^{GG}, a^{GB}, a^{BG}, a^{BB}$ achieving those payoffs. That is,

$$g_i(a^{GG}) = w_i^{GG},\ g_i(a^{GB}) = w_i^{GB},\qquad (1)$$
$$g_i(a^{BG}) = w_i^{BG},\ g_i(a^{BB}) = w_i^{BB}.$$

While such action profiles typically do not exist, there always exist an m and four finite sequence of action profiles $\left\{\left(a_1^{XY}, ..., a_m^{XY}\right) \in A^m : X, Y \in \{B, G\}\right\}$, such that w^{XY}, defined as the average discounted payoff vector over the sequence $\left\{a_1^{XY}, ..., a_m^{XY}\right\}$, $X, Y \in \{B, G\}$, satisfies properties (i)-(iv) for sufficiently high discount factors. If $m > 1$, the construction that follows must be accordingly modified, by replacing each single period within a block but the first by a finite sequence of length m, and each occurrence of the action profile a^{XY} by the finite sequence of action profiles $\left(a_1^{XY}, ..., a_m^{XY}\right)$. The details are omitted. We will show that each payoff in the set $[\underline{v}_1, v_1] \times [\underline{v}_2, v_2]$ is achieved by some block equilibrium. The length of the blocks is specified later. Throughout the subsection, the time index t refers to the number of periods relative to the block under consideration, not to the absolute number of periods in the infinitely repeated game. In particular, h_i^t refers to the recent history only, simply referred to as a history. Since monitoring is perfect and i's signal can be identified with $-i$'s actions, $h_i^t = h_{-i}^t$, up to the ordering of signals and actions.

For each i, partition the set of actions of player i into two non-empty subsets G_i and B_i. The action chosen within the first period of each block plays essentially a role of communication. If, in the first period of a block $(t = 1)$, player i picks an action from the set G_i, we say that he sends message $M_i = G$; otherwise, we say that he sends message $M_i = B$.

We define now $\mathcal{S}_i \subset S_i^T$. The set of strategies $\mathcal{S}_1$, $\mathcal{S}_2$ are *all* strategies $s_1 \in S_1^T$, $s_2 \in S_2^T$, satisfying, for $i = 1, 2$

$({*}_2)$ If $h_i^t = \left(a, a^{GG}, a^{GG}, \ldots, a^{GG}\right)$ for some $a \in G_1 \times G_2$, then $s_i\left(h^t\right) = a_i^{GG}$, $i = 1, 2$. If $h_1^t = \left(a, a^{GB}, a^{GB}, \ldots, a^{GB}\right)$ for some $a \in B_1 \times G_2$, then $s_1\left(h_1^t\right) = a_1^{GB}$; if $h_2^t = \left(a, a^{BG}, a^{BG}, \ldots, a^{BG}\right)$ for some $a \in G_1 \times B_2$, then $s_2\left(h_2^t\right) = a_2^{BG}$.

That is, the only restriction placed on $\mathcal{S}_i$ is that its elements $s_i \in \mathcal{S}_i$ be such that, if player $-i$ sent message $M_{-i} = G$ and both players repeatedly chose the actions determined by M_i and $M_{-i} = G$ since then, the strategy s_i requires that player i choose that action. No restriction is imposed on the specification of s_i in the initial period; after any history along which $M_{-i} = B$; and after any history along which at least one player has deviated (in some period $t \geq 2$) from the actions determined by M_i and $M_{-i} = G$. Observe that player i's message M_i defines player $-i$'s payoff level, high or low (the superscripts of action profiles are in the reverse order of the message profiles).

Because only those histories that are off the equilibrium path under perfect monitoring raise serious problems under imperfect private monitoring, it is useful to consider those strategies in $\mathcal{S}_i$ for which there are as few histories off the equilibrium path as possible. To this purpose, for each player i and each history h_i^t, define

$$\mathcal{A}_i\left(h_i^t\right) = \left\{a_i \in A_i : \exists_{s_i \in \mathcal{S}_i} \quad s_i\left[h_i^t\right](a_i) > 0\ \right\}.$$

The set $\mathcal{A}_i\left(h_i^t\right)$ is the set of all actions *prescribed by* $\mathcal{S}_i$, that is all the actions that are chosen, with positive probability, after history h_i^t, by some strategy $s_i \in \mathcal{S}_i$. Observe that $\mathcal{A}_i\left(h_i^t\right)$ is either A_i or a singleton. For instance, if $i = 1$, then this singleton is either $\left\{a_1^{GG}\right\}$ or $\left\{a_1^{GB}\right\}$. With some abuse of notation, we say that $\mathcal{S}_i$ *prescribes action* a_i whenever $\mathcal{A}_i\left(h_i^t\right) = \{a_i\}$.

Given some (small) $\rho > 0$, denote by $\mathcal{S}_i^\rho$ the set of strategies $s_i \in \mathcal{S}_i$ that assign probability at least ρ to action a_i after every history h_i^t and any action $a_i \in \mathcal{A}_i\left(h_i^t\right)$, that is,

$$\mathcal{S}_i^\rho = \left\{s_i \in \mathcal{S}_i : \forall_{h_i^t}\ \forall_{a_i \in \mathcal{A}_i\left(h_i^t\right)} \quad s_i\left[h_i^t\right](a_i) \geq \rho\right\}.$$

For any history, a strategy in $\mathcal{S}_i^\rho$ assigns positive probability to any action unless this is precisely ruled out by condition $(^*{}_2)$. In particular, the set of histories off the equilibrium path is independent of the particular choice of strategy profile $s \in \mathcal{S}_1^\rho \times \mathcal{S}_2^\rho$.

Next, we construct s_{-i}^G, $s_{-i}^B \in \mathcal{S}_{-i}^\rho$ with the desired properties. Consider for instance $-i = 1$. First, define strategy $s_1^g \in \mathcal{S}_1$ as follows. At the beginning of a block, strategy s_1^g sends message $M_1 = G$. That is, choose $s_1^g[\emptyset] \in \triangle G_1$. For any history $h_1^t = (a^0, a^1, a^2, \ldots a^t)$ such that $a^0 \in G_1 \times G_2$, consider the first action $a^\tau \neq a^{GG}$, $\tau > 1$, if any. If $a_1^\tau = a_1^{GG}$, $a_2^\tau \neq a_2^{GG}$, then $s_1^g\left[h_1^t\right] = \alpha_1^*$, the action minmaxing player 2. If $a_1^\tau \neq a_1^{GG}$, $a_2^\tau = a_2^{GG}$, then choose $s_1^g\left[h_1^t\right]$ among the best-responses to α_2^*, the action minmaxing him. Otherwise, $s_1^g\left[h_1^t\right] = a_1^{GG}$. The specification for the case $a^0 \in B_1 \times G_2$ (respectively, $G_1 \times B_2$ and $B_1 \times B_2$) is identical, replacing a_i^{GG} by a_i^{GB} (respectively, a_i^{BG} and a_i^{BB}) everywhere.

That is, strategy s_1^g sends message G and specifies the action determined by the pair of messages until the first unilateral deviation. If player 2 has deviated, strategy s_1^g minmaxes player 2. Otherwise, strategy s_1^g is a best-response to minmaxing.

Strategy $s_1^b \in \mathcal{S}_1$ sends message $M_1 = B$. That is, choose $s_1^b[\emptyset] \in \triangle B_1$. For any other history, $s_1^b\left[h_i^t\right] = s_1^g\left[h_i^t\right]$. Thus, strategy s_1^g and s_1^b differ only in the message they specify in the initial period.

If player 1 plays s_1^b, then player 2 obtains a (per-period) payoff higher than w_2^{GB} or w_2^{BB} in at most two periods: period 1 and the period in which he unilaterally deviates, if any. If player 1 plays s_1^g and player 2 plays a strategy $s_2 \in \mathcal{S}_2$, then player 2 can obtain a payoff lower than w_2^{GG} or w_2^{BG} at most in period 1. Therefore, by choosing T large enough, and δ close enough to one, we can ensure that the average payoff from each of the four strategy profiles is arbitrarily close to the corresponding payoff vector w^{XY}, $X, Y \in \{G, B\}$.

Perturb very slightly strategies s_i^g and s_i^b so that, after any history h_i^t, player i plays each action $a_i \in \mathcal{A}_i\left(h_i^t\right)$ with probability at least ρ, so as to obtain a pair of strategies s_i^G, $s_i^B \in \mathcal{S}_i^\rho$ such that the average payoff of player i from any strategy $s_i \in \mathcal{S}_i$ against s_{-i}^G is at least v_i, while the average payoff of player i from any strategy $s_i \in S_i^T$ against s_{-i} does not exceed $\underline{v}_i$.

Finally, in order to specify transition probabilities, it is necessary to define two further strategies. While these strategies are not actually used by the players in equilibrium, they are benchmarks relative to which the actual play (as inferred from the private history) is compared. Let r_i^G be a strategy $s_i \in \mathcal{S}_i$ such that, for every history $h_i^t \in H_i^T$, the strategy $s_i|h_i^t$ yields the lowest payoff against s_{-i}^G among all strategies $s_i \in \mathcal{S}_i$. Similarly, let r_i^B be a strategy $s_i \in S_i^T$ such that, for every history $h_i^t \in H_i^T$, the strategy $s_i|h_i^t$ yields the highest payoff against s_{-i}^B among all strategies $s_i \in S_i^T$. Without loss of generality, we may take r_i^G and r_i^B to be pure.

Enlarging the box $[\underline{v}_1, v_1] \times [\underline{v}_2, v_2]$ if necessary, we may assume that:

$$U_i^T(r_i^G, s_{-i}^G) = v_i \text{ and } U_i^T(r_i^B, s_{-i}^B) = \underline{v}_i.$$

We conclude this subsection by establishing the two-player folk theorem under perfect monitoring.

Theorem 1 ($n = 2$, perfect monitoring): *Under perfect monitoring, for any $(v_1, v_2) \in V^*$, if players discount the future sufficiently little, there exists a block equilibrium of the infinitely repeated game where, for all i, player i's average payoff is v_i.*

Proof: For each i, construct a strategy for player $-i$ as follows. At the beginning of each block, player $-i$'s continuation strategy only depends on its state $u \in [\underline{v}_i, v_i]$. Thus, the state space is $[\underline{v}_i, v_i]$, the set of possible continuation payoffs of player i. The initial state is v_i.

At the beginning of each block, in state $u \in [\underline{v}_i, v_i]$, player $-i$ performs an initial randomization: for $q \in [0, 1]$ such that $u = qv_i + (1-q)\underline{v}_i$, player $-i$ picks strategy s_{-i}^G with probability q and strategy s_{-i}^B with probability $1 - q$. Thus, player $-i$ uses one or the other strategy throughout the block as a function of the outcome of this initial randomization.

To define transition probabilities, consider first the case in which player $-i$ plays s_{-i}^G within the block. In this case, given the realized history h_{-i}^T in the block, consider all periods t along this history in which player i deviated from the action prescribed by r_i^G. Let θ_t denote the difference between player i's unnormalized continuation payoff from playing r_i^G from period t on and player i's unnormalized continuation payoff from choosing i's action as observed by player $-i$ in period t, followed by reversion to r_i^G from period $t + 1$ on. Let $\theta_t^G := \min\{0, \theta_t\}$, and $\pi_i^G\left(h_{-i}^T\right) := \delta^{-T} \sum_{t=1}^T \delta^{t-1}\theta_t^G$. Pick δ close enough to 1 so that $\left(1 - \delta^T\right) \pi_i^G\left(h_{-i}^T\right) > \underline{v}_i - v_i$ for all histories h_{-i}^T. At the end of the block, player $-i$ transits to the state:

$$v_i + \left(1 - \delta^T\right) \pi_i^G\left(h_{-i}^T\right), \tag{2}$$

which is in $[\underline{v}_i, v_i]$ for any history h_i^T. Observe that, if player i knew that player $-i$ was playing s_{-i}^G, every strategy $s_i \in \mathcal{S}_i$ would be a best-response. That is, player $-i$ "punishes" player i from taking any continuation strategy that improves upon the continuation strategy derived r_i^G by the exact amount that makes him indifferent between deviating or not from the action prescribed by r_i^G. For any deviation consistent with a continuation strategy in $\mathcal{S}_i$, this is necessarily a punishment, by definition of r_i^G. Observe that, as $\theta_t^G = \min\{0, \theta_t\}$, player $-i$'s transition probability is unaffected by deviations that yield a lower continuation payoff than the payoff from the continuation strategy derived from r_i^G, as such deviations are necessarily unprofitable for player i.

Suppose now that player $-i$ plays s_{-i}^B. Similarly, given the realized history h_i^T in the block, consider all periods t along this history in which player i deviated from the action prescribed by r_i^B. Let θ_t^B denote the difference between player i's unnormalized continuation payoff from playing r_i^B from period t on, and player i's unnormalized continuation payoff from choosing i's action as observed by player $-i$ in period t, followed by reversion to r_i^B from period $t+1$ on. By definition of r_i^B, the value θ_t^B is necessarily non-negative. Let $\pi_i^B\left(h_{-i}^T\right) := \delta^{-T}\sum_{t=1}^T \delta^{t-1}\theta_t^B$. Pick δ close enough to 1 so that $\left(1-\delta^T\right)\pi_i^B\left(h_{-i}^T\right) < v_i - \underline{v}_i$ for all histories h_{-i}^T. At the end of the block, player $-i$ then transits to the state:

$$\underline{v}_i + \left(1-\delta^T\right)\pi_i^B\left(h_{-i}^T\right), \tag{3}$$

which is in $[\underline{v}_i, v_i]$ for any history h_i^T. Observe that, if player i knows that player $-i$ was playing s_{-i}^B, *every* strategy $s_i \in S_i^T$ of player i would be a best-response, so that player i's choice of strategy is payoff-irrelevant in this case. This property, which is preserved under imperfect private monitoring, implies that player i may always assume that player $-i$ is using s_{-i}^G, when computing best-responses.

It follows from equations (2)-(3) and the one-stage deviation property that, given the strategy of player $-i$, any strategy s_i of player i that is such that its restriction to any given block is an element to $\mathcal{S}_i$ is a best-response (independently of how s_i selects this element of $\mathcal{S}_i$, at the beginning of the block, possibly as a function of the entire history). The payoff of player i is equal to the weighted average of the payoff of playing r_i^G against s_{-i}^G and the payoff of playing r_i^B against s_{-i}^B with weights q and $1-q$. Both the average payoff within the block and the continuation payoff from playing r_i^G against s_{-i}^G are equal to v_i, and both the average payoff within the block and the continuation payoff from playing r_i^B against s_{-i}^B are equal to $\underline{v}_i$. Thus, at the beginning of a block, player i's payoff when player $-i$'s state is u is $qv_i + (1-q)\underline{v}_i = u$. *Q.E.D.*

3.2 Imperfect Private Monitoring

In this subsection, we extend the two-player folk theorem under perfect monitoring of the previous subsection to the case in which the noise level is small enough. This requires modifying the strategies s_i^G, s_i^B as well as the transfers π_i^G, π_i^B for all $i = 1, 2$.

Observe first that the definitions of $\mathcal{S}_i$, $\mathcal{S}_i^\rho$, of strategies s_i^G, $s_i^B \in \mathcal{S}_i^\rho$ and of transfers π_i^G, π_i^B are well-defined under imperfect private monitoring, as their definitions are stated in terms of private histories, and the domain of private histories is the same under perfect and imperfect private monitoring, under the canonical signal structure.

For some histories h_i^t, the continuation strategy specified by $s_i^G \mid h_i^t$ and the transfers π_i^G applied to the resulting history h_i^T may be significantly different from the specification above.

As already mentioned, under perfect monitoring, the set of histories that are off the equilibrium path given a strategy profile $s \in \mathcal{S}_1^\rho \times \mathcal{S}_2^\rho$ is independent of the particular choice of s. Therefore, given $\mathcal{S}_1^\rho$, $\mathcal{S}_2^\rho$, we can unambiguously define the set of *erroneous histories* $H_i^E \subset H_i^T$ as the set of all recent histories that are off the equilibrium path. More precisely, the recent history h_i^t is *erroneous* if, under perfect monitoring, it is off the equilibrium path for some strategy profile in $\mathcal{S}_1^\rho \times \mathcal{S}_2^\rho$. Otherwise the history is *non-erroneous*. Let $H_i^{E,t}$ denote the set of all erroneous t-length histories, and let $H_i^{N,t} = H_i^t \backslash H_i^{E,t}$ denote the complement of $H_i^{N,t}$. Let

$$H_i^N = \bigcup_{t \leq T} H_i^{N,t},\ H_i^E = \bigcup_{t \leq T} H_i^{E,t}.$$

Erroneous histories may arise with positive probability under imperfect monitoring. Indeed, if the monitoring structure happens to satisfy the full support assumption, then with the exception of histories that follow one's own deviation, every history, erroneous or not, occurs with positive probability. Thus, a history $h_i^t \in H_i^E$ is erroneous since along such a history, if neither player has deviated from his strategy in $\mathcal{S}_1^\rho$, $\mathcal{S}_2^\rho$, at least one player must have observed an incorrect, or *erroneous* signal. This need not be player i: for instance, all signals of player i along the history in h_i^t may be correct, but player $-i$ may have observed in some period, but the last, an incorrect signal, that called upon a continuation strategy that is inconsistent with the continuation strategy he would have followed if he had observed the correct signal.

Since player i's best-response after a recent history $h_i^t \in H_i^t$ depends on $-i$'s continuation strategy $s_{-i} \mid h_{-i}^t$ and transfer $\pi_i\left(h_{-i}^T\right)$ that will result, it depends on his beliefs about h_{-i}^t and on his belief about whether, in the current block, (s_{-i}, π_i) is equal to $\left(s_{-i}^G, \pi_i^G\right)$ or $\left(s_{-i}^B, \pi_i^B\right)$. As in the case of perfect monitoring, s_{-i}^B (and π_i^B) will be jointly defined in a manner ensuring that, conditional on player $-i$ using s_{-i}^B (and π_i^B), player i is indifferent over all strategies in S_i^T, and therefore, over all continuation strategies (within the block) after $h_i^t \in H_i^t$, even if h_i^t is erroneous. Therefore, for every recent history $h_i^t \in H_i^t$, player i may condition on his opponent playing strategy s_{-i}^G (and applying transfer π_i^G), no matter how unlikely this event may be. Thus, as $s_{-i}^G \mid h_{-i}^t$ only depends on the recent history $h_{-i}^t \in H_{-i}^t$, player i's best-response after a recent history $h_i^t \in H_i^t$ also depends only on the recent history. This ensures that optimality only depends on the recent history, and not on the entire history.

If the noise level ε is sufficiently small, in particular small relative to ρ, player i's belief about h_{-i}^t given h_i^t (that is, the probability distribution over the recent history $h_{-i}^t \in H_{-i}^t$, derived from Bayes' rule given h_i^t, conditional on $s_{-i} = s_{-i}^G$) is almost degenerate if h_i^t is non-erroneous. In this case, player i assigns probability almost one to player $-i$'s recent history $h_{-i}^t = h_i^t$.[6] That is, if erroneous signals are unlikely, player i assigns probability almost 1 to his opponent's recent actions being equal to his own recent signals, and to his opponent's recent signals being equal to his own recent actions. Linear algebra can therefore be used to define the transfer π_i^G that

[6] Strictly speaking, as h_i^t is a sequence of ordered pairs of i's actions and i's signals ($-i$'s actions), h_i^t and h_{-i}^t differ in the ordering of those pairs.

guarantees that continuing with strategies from $\mathcal{S}_i$ (in particular, s_i^G and s_i^B defined as under the perfect monitoring) is optimal after such histories (see Lemma 1 and 2). As non-erroneous histories occur with probability almost 1, provided the noise level is sufficiently small, transfers can be specified in a way that ensures that expected payoffs in the auxiliary scenario, at the beginning of the block, are arbitrarily close to the corresponding expected payoffs under perfect monitoring.

After erroneous histories h_i^t, however, player i's belief about h_{-i}^t is not necessarily well-behaved. As player i conditions on at least one incorrect signal having been observed, the relative likelihood of incorrect signals may drastically affect his belief. For instance, he may assign low probability to $h_{-i}^t = h_i^t$. Fortunately, erroneous histories occur with small probability, and therefore, independently of player i's continuation strategy (within the block) after such histories, its specification has a small impact on expected payoffs. Nevertheless, observe that, for $h_i^t \in H_i^E$, the optimal continuation strategy $s_i^G \mid h_i^t$ depends not only on π_i^G, but also on $s_{-i}^G \mid H_{-i}^E$; on the other hand, π_i^G that guarantees that continuing with strategies from $\mathcal{S}_i$ is optimal after non-erroneous histories depends on $s_{-i}^G \mid H_{-i}^E$ and $s_i^G \mid H_i^E$. Therefore, it is necessary to define $\left(s_1^G, \pi_2^G\right)$ and $\left(s_2^G, \pi_1^G\right)$ jointly, which is achieved here by applying Kakutani's fixed-point theorem (see the proof of Theorem 1).

While continuation play after erroneous histories hardly affects expected payoffs, we need to make sure that it does not affect incentives after non-erroneous histories. After all, if player i picks an action outside the support of what his equilibrium strategy prescribes, this results with probability almost 1 in player $-i$ observing an erroneous history, for which we only know that an optimal continuation strategy exists. That continuation strategy could potentially yield high flow payoffs to player i in the remaining periods of the block, so that picking this action could offer a profitable deviation.

This is an issue, however, only if player i gets the opportunity to pick such an action. This is where the definition of $\mathcal{S}_i^\rho$ and of s_{-i}^B comes into play. If player $-i$ uses s_{-i}^B ("punishing" thereby player i), he sends message $M_{-i} = B$ with probability almost one, when ρ is small enough. In addition, $H_{-i}^{N,1} = H_{-i}^1$, as both s_i^G and s_i^B specify totally mixed actions in the first period. More generally, given message $M_{-i} = B$, no history h_{-i}^t along which player $-i$ has not deviated himself is erroneous. Therefore, if player $-i$ uses strategy s_{-i}^B, then with probability almost one, no recent history $h_{-i}^t \in H_{-i}^T$ of player $-i$ is erroneous, independently of player i's strategy $s_i \in S_i^T$. In particular, we can use linear algebra to ensure that player $-i$ finds it optimal to minmax player i in the appropriate contingencies. This guarantees the possibility to punish player i, independently of the specification of play after erroneous histories.

It is convenient to consider now the auxiliary scenario, described in Section 2. Recall that $B_i(s_{-i}, \pi_i)$ denotes the set of auxiliary scenario best-responses of player i to the T-period strategy s_{-i} and to the transfer π_i, and that, given strategies s_{-i}, $s_i \in S_i^T$, transfer π_i, $B_i(s_{-i}, \pi_i \mid \overline{s}_i)$ denotes the set of strategies that maximize player i's auxiliary-scenario payoff against s_{-i}, π_i among all strategies $s_i \in S_i^T$ such that $s_i \mid H_i^E = \overline{s}_i \mid H_i^E$.

Lemma 1 *For every strategy $\overline{s} \mid H^E$, there exists $\overline{\varepsilon} > 0$ such that for $\varepsilon < \overline{\varepsilon}$:*
There exists a non-positive transfer $\pi_i^G : H_{-i}^T \to \mathbf{R}_-$ such that

$$\{s_i \in S_i^T : s_i \mid H_i^N = \widetilde{s}_i \mid H_i^N \text{ for some } \widetilde{s}_i \in \mathcal{S}_i \text{ and } s_i \mid H_i^E = \overline{s}_i \mid H_i^E\} \subset B_i(\overline{s}_{-i}^G, \pi_i^G \mid \overline{s}_i), \quad (3)$$

where $\overline{s}_{-i}^G \mid H_{-i}^N = s_{-i}^G \mid H_{-i}^N$ and $\overline{s}_{-i}^G \mid H_{-i}^E = \overline{s}_{-i} \mid H_{-i}^E$.

Proof: Assume without loss of generality that $-i = 1$ and $i = 2$. Given a history h_1^T, let (h_1^t, a_2) denote the truncation of h_1^T to h_1^t and the signal obtained by player 1 in period t. The transfer will have the form:

$$\pi_2^G(h_1^T) = \frac{1}{\delta^T}\left[\sum_{t=1}^{T} \delta^{t-1}\theta(h_1^t, a_2)\right],$$

for some function $\theta(\cdot, \cdot)$ to be specified. Pick $r_2^G \in S_2^T$ to be a strategy that satisfies:
(a) $r_2^G \mid H_2^N = s_2 \mid H_2^N$ for some $s_2 \in \mathcal{S}_2$;
(b) $r_2^G \mid H_2^E = \overline{s}_2 \mid H_2^E$;
(c) $\forall h_2^t \in H_2^N$, $r_2^G | h_2^t$ yields the lowest payoff against $\overline{s}_1^G$, in the T-period repeated game, among all strategies with properties (a) and (b).
Without loss of generality, assume that $r_2^G \mid H_2^N$ is a pure strategy.

We now define $\theta(h_1^t, a_2)$, $\forall h_1^t \in H_1^t$, $\forall a_2 \in \Sigma_1 = A_2$ by backward induction with respect to t. To satisfy (3), it suffices to pick non-positive values for $\theta(h_1^t, a_2)$ such that (given $\theta(h_1^\tau, a_2)$, $\tau > t$), the following constraints, or properties, are satisfied:

1. For every history $h_2^t \in H_2^{N,t}$ such that $\mathcal{S}_2$ prescribes A_2, player 2 is indifferent under the auxiliary scenario between playing all actions $a_2 \in A_2$, each followed by switching to r_2^G from period $t + 1$ on;

2. For every history $h_2^t \in H_2^{N,t}$ such that $\mathcal{S}_2$ prescribes $a_2 = a_2^{GG}$ or a_2^{BG}, the payoff of player 2 under the auxiliary scenario to playing a_2 exceeds the payoff to playing any other action, both followed by switching to r_2^G from period $t + 1$ on.

[Note that the payoff difference across continuation strategies $s_2 \mid h_2^t$ of player 2 is independent of $\theta(h_1^{\widetilde{t}}, a_2)$ for $\widetilde{t} < t$ as those values cannot be affected by actions taken in periods $\widetilde{t} \geq t$. The differences, even the preference ordering over continuation strategies, may of course depend on the values of $\theta(h_1^{\widetilde{t}}, a_2)$ for $\widetilde{t} > t$ and $a_2 \in A_2$, but these values are already determined by backward induction.]

For every $\nu > 0$, observe that there exists ε/ρ small enough such that: for any history $h_2^t \in H_2^{N,t}$, there exists a history $h_1^t \in H_1^{N,t}$ such that, conditional on observing h_2^t, player 2 assigns probability at least $1 - \nu$ to the event that player 1 observed the *corresponding* history

$h_1^t \in H_1^{N,t}$, along which the signals of player 1 along h_1^t coincide with the actions of player 2 in h_2^t and the actions of player 1 along h_1^t coincide with the signals of player 2 in h_2^t. Note that, given two distinct histories $h_2^t, h_2'^t \in H_2^{N,t}$, the corresponding histories h_1^t and $h_1'^t$ are distinct. Given any history $h_2^t \in H_2^{N,t}$ and any action $a_2 \in A_2$, consider as a row vector the probabilities assigned by player 2, conditional on history h_2^t and on action a_2 taken by player 2 in period t, to the different histories $h_1^t \in H_1^t$ and signals $a_2 \in \Sigma_1 = A_2$ observed by player 1 in period t. Construct a matrix D^t by stacking the row vectors for all non-erroneous histories $h_2^t \in H_2^{N,t}$ and actions $a_2 \in A_2$.

By the previous paragraph, the matrix D^t has full row rank, provided ε/ρ is small enough. Therefore, there exist values $\theta(h_1^t, a_2)$ satisfying constraints 1 and 2. Indeed the number of columns (rows) of D^t exceeds the number of linear equality (or inequality) constraints that 1-2 imposes on $\theta(h_1^t, a_2)$ by k, where $k := \#H_2^{N,t}$ is the number of t-length non-erroneous histories h_2^t. [Say, $A_2 = \{a_2^1, ..., a_2^l\}$ consists of l actions and suppose that, given a history $h_2^t \in H_2^{N,t}$, constraint 1 must be satisfied (the argument for constraint 2 is analogous). That is, $l-1$ equations have to be satisfied: player 2 must be indifferent between playing a_2^k and a_2^{k+1} for $k = 1, ..., l-1$. Since there are l actions a_2, there are l rows of D^t corresponding to each $h_2^t \in H_2^{N,t}$, but only $l-1$ constraints.] Further, we can assume that these values $\theta(h_1^t, a_2)$ are all non-positive since properties 1-2 define the values $\theta(h_1^t, a_2)$ up to a constant. *Q.E.D.*

Let from now on the notation $U_i^T(s)$ *exclusively* refers to player i's average payoff in the finitely repeated game *under perfect monitoring*, given strategy profile $s \in S_1^T \times S_2^T$, while $U_i^A(s, \pi_i)$ denotes player i's average payoff given transfer π_i and strategy profile $s \in S_1^T \times S_2^T$ under imperfect private monitoring. The next Lemma shows that, provided the noise level is sufficiently small, the applied transfer π_i^G is arbitrarily close to zero, given s_{-i}^G, if player i chooses the "worst" pure strategy against s_{-i}^G among all strategies in $\mathcal{S}_i$, in the finitely repeated game.

Lemma 2 *In Lemma 1, the positive transfer $\pi_i^G : H_{-i}^T \to \mathbf{R}_-$ can be chosen so that, for every $s_i \in B_i(\overline{s}_{-i}^G, \pi_i^G \mid \overline{s}_i)$, (i)*

$$\lim_{\varepsilon \to 0} U_i^A(s_i, \overline{s}_{-i}^G, \pi_i^G) = \min_{\widetilde{s}_i \in \mathcal{S}_i} U_i^T(\widetilde{s}_i, \overline{s}_{-i}^G); \tag{4}$$

(ii) π_i^G is bounded away from $-\infty$, i.e. there exists $\underline{\pi}$ (independent of $\overline{s}$) such that $\pi_i^G \geq \underline{\pi}$, and (iii) π_i^G depends continuously on $\overline{s}$.

Proof: To guarantee (4), the constants $\theta(\cdot, \cdot)$ from the proof of Lemma 1 must be further specified. We will show by backward induction with respect to t that, in addition to properties 1-2, we can assume that:

(i) $\theta(h_1^t, a_2)$ tends to 0 as $\varepsilon \to 0$ whenever $h_1^t \in H_1^t$ is the history corresponding to some history $h_2^t \in H_2^{N,t}$ and $a_2 = r_2^G(h_2^t)$;

(ii) given a history $h_2^t \in H_2^{N,t}$, the expected transfer that player 2 receives in period T if he uses the continuation strategy $r_2^G \mid h_2^t$ tends to 0 as $\varepsilon \to 0$.

Let further $\theta(h_1^t, a_2) = 0$, $\forall a_2 \in A_2$, whenever h_1^t does not correspond to any $h_2^t \in H_2^{N,t}$. Notice further that there exist (not necessarily non-positive) values for $\theta(h_1^t, a_2)$ with properties 1-2 such that, for every $h_2^t \in H_2^{N,t}$, $\theta(h_1^t, a_2) = 0$ where h_1^t is the history corresponding to h_2^t and $a_2 = r_2^G(h_2^t)$. (Remember that the number of constraints imposed on $\theta(h_1^t, a_2)$ by 1-2 falls below the rank of D^t by k). In addition, property 2 can without loss of generality be replaced with (given $\theta(h_1^\tau, a_2)$, $\tau > t$):

3. For every history $h_2^t \in H_2^{N,t}$ and action $a_2 \in A_2$ such that $\mathcal{S}_2$ prescribes a_2^{GG} (or respectively a_2^{BG}) and playing a_2 yields payoff in the T-period repeated game *higher than or equal to* the payoff to playing the prescribed action, both followed by switching to r_2^G from period $t+1$ on: the payoff of player 2 under the auxiliary scenario from playing the prescribed action is equal to the payoff from playing any such a_2, both followed by switching to r_2^G from period $t+1$ on;
4. For every history $h_2^t \in H_2^{N,t}$ and action $a_2 \in A_2$ such that $\mathcal{S}_2$ prescribes a_2^{GG} (or respectively a_2^{BG}) and playing a_2 yields payoff in the T-period repeated game *lower* than the payoff to playing the prescribed action, both followed by switching to r_2^G from period $t+1$ on: the expectation of the transfer that player 2 receives in period T if he uses the continuation strategy $r_2^G \mid h_2^t$ is equal to the expectation of the transfer if he plays a_2 followed by switching to r_2^G from period $t+1$ on.

Note that properties 3 and 4 imply property 2.

Since $\theta(h_1^t, a_2) = 0$ for h_1^t corresponding to $h_2^t \in H_2^{N,t}$ and $a_2 = r_2^G(h_2^t)$, the expectation of the transfer that player 2 receives in period T if he uses the continuation strategy $r_2^G \mid h_2^t$ tends to 0 as $\varepsilon \to 0$ by the induction hypothesis. This yields (i)-(ii), except that $\theta(h_1^t, a_2)$ may be positive.

We shall now show that, for any $a_2 \in A_2$ and any history h_1^t corresponding to some $h_2^t \in H_2^N$, $\theta(h_1^t, a_2)$ tends to a non-positive value as $\varepsilon \to 0$. This will guarantee that all values $\theta(h_1^t, a_2)$ can be made non-positive, by subtracting a constant from all of them, and this will not affect the required properties since the constant can be tending 0 for $\varepsilon \to 0$.

If $\mathcal{S}_2$ prescribes A_2 at history $h_2^t \in H_2^{N,t}$, then any continuation strategy that uses $a_2 \neq r_2^G(h_2^t)$ in period t and switches to r_2^G from period $t+1$ on, satisfies conditions (a) and (b) from the definition of r_2^G; therefore, by condition (c) of the same definition, such a continuation strategy yields at least as high a payoff against $\overline{s}_1^G$ in the T-period repeated game as r_2^G. Since the expected transfer to player 2 who uses the continuation strategy $r_2^G \mid h_2^t$ tends to 0 for $\varepsilon \to 0$, the expected transfer to player 2 who uses the other continuation strategy has to tend to a non-positive number; otherwise property 1 would be violated. This in turn implies that the value $\theta(h_1^t, a_2)$, where h_1^t is the history corresponding to $h_2^t \in H_2^{N,t}$, tends to a non-positive number, because the difference between $\theta(h_1^t, a_2)$ and the expected transfer to player 2 who plays a_2 and switches to r_2^G from period $t+1$ on converges to 0 as $\varepsilon \to 0$. The same argument, except referring to properties 3 and 4 instead of 1, applies to the histories $h_2^t \in H_2^{N,t}$ such that $\mathcal{S}_2$ prescribes a_2^{GG} or a_2^{BG}.

The values $\theta(h_1^t, a_2)$ can be picked continuous functions of $\overline{s}_2 \mid H_2^{E,t}$. To see this, apply again backward induction with respect to t. Note first that the system of linear equations: 1, 3, 4, $\theta(h_1^t, a_2) = 0$ for every h_1^t corresponding to some $h_2^t \in H_2^{N,t}$ and $a_2 = r_2^G(h_2^t)$, as well as $\theta(h_1^t, a_2) = 0$ for every h_1^t not corresponding to any $h_2^t \in H_2^{N,t}$, uniquely determines $\theta(h_1^t, a_2)$ $\forall h_1^t \in H_1^t$, $\forall a_2 \in \Sigma_1 = A_2$. Obviously, this system of equations depends continuously on $\overline{s}$. Thus, the values $\theta(h_1^t, a_2)$ depend continuously on $\overline{s}$ as well. Obviously, $\theta(h_1^t, a_2)$ still depends continuously on $\overline{s}$ if we subtract from all of them

$$\max_{h_1^t, a_2} \theta(h_1^t, a_2),$$

which is a constant that satisfies the desired properties specified above.

Finally, we can choose the transfers $\theta(h_1^t, a_2)$ to be bounded away from $-\infty$. To see this, use backward induction with respect to t. Consider first the perfect monitoring case. Then it can be assumed that $\theta(h_1^t, a_2)$ with properties 1, 3, and 4 is at least as large as:

$$-B := -T[\max_a u_2(a) - \min_a u_2(a)] - \Sigma_{s>t}[\max_{h_1^s, a_2} \theta(h_1^s, a_2) - \min_{h_1^s, a_2} \theta(h_1^s, a_2)].$$

Thus, for $\varepsilon > 0$ small enough, we can choose the values $\theta(h_1^t, a_2)$ satisfying properties 1, 3, and 4 that all exceed $-2B$. By compactness of the set of strategies $\overline{s}$, it can be therefore assumed that $\theta(\cdot)$ is bounded away from $-\infty$ by some $\underline{\theta}$ independent of $\overline{s}$. This implies that $\pi_2(h_1^T) > \underline{\pi}$, $\forall_{h_1^T \in H_1^T}$. *Q.E.D.*

As the counterparts of Lemma 1 and 2 for the case $\overline{s}_{-i} = \overline{s}_{-i}^B$ are straightforward, we gather them in a unique lemma.

Lemma 3 *For every strategy $\overline{s} \mid H^E$, there exists $\overline{\varepsilon} > 0$ such that for $\varepsilon < \overline{\varepsilon}$:*
There exist a non-negative transfer $\pi_i^B : H_{-i}^T \to \mathbf{R}_+$ such that

$$S_i^T = B_i(\overline{s}_{-i}^B, \pi_i^B), \tag{5}$$

where $\overline{s}_{-i}^B \mid H_{-i}^N = s_{-i}^B \mid H_{-i}^N$ and $\overline{s}_{-i}^B \mid H_{-i}^E = \overline{s}_{-i} \mid H_{-i}^E$, and for every $s_i \in B_i(\overline{s}_{-i}^B, \pi_i^B)$,

$$\lim_{\varepsilon \to 0} U_i^A(s_i, \overline{s}_{-i}^B, \pi_i^B) = \max_{\widetilde{s}_i \in S_i^T} U_i^T(\widetilde{s}_i, \overline{s}_{-i}^B). \tag{6}$$

Proof: The argument is similar to Lemma 1 and 2, but simpler. We consider a pure strategy r_2^B of player 2 such that, for every history $h_2^t \in H_2^t$ (both erroneous and non-erroneous), the continuation strategy $r_2^B|h_2^t$ yields the highest payoff against $\overline{s}_1^B$ in the T-period repeated game. We again define $\theta(h_1^t, a_2)$, $\forall h_1^t \in H_1^t$, $\forall a_2 \in \Sigma_1 = A_2$ by backward induction with respect to t. Suppose all values $\theta(h_1^\tau, a_2)$ for $\tau > t$ have already been defined, and we are given $h_1^t \in H_1^t$ and $a_2 \in \Sigma_1 = A_2$. Let $\theta(h_1^t, a_2)$ be the difference in player 2's payoff in the T-period repeated game (given $\theta(h_1^\tau, a_2)$, $\tau > t$) between playing $r_2^B|h_2^t$, where h_2^t corresponds to h_1^t, and playing a_2 followed by switching to r_2^B from period $t+1$ on. By definition, the transfer:

$$\pi_2^B(h_1^T) = \frac{1}{\delta^T}\left[\sum_{t=1}^{T} \delta^{t-1}\theta(h_1^t, a_2)\right]$$

satisfies (5), and (6) is satisfied because $\theta(h_1^t, a_2) = 0$ whenever $a_2 = r_2^B(h_2^t)$. *Q.E.D.*

Lemmas 1-3 imply that, for any specification of strategies on H^E, players can be given incentives to prefer, among all strategies coinciding with the specified strategy on H^E, those that correspond to some strategy from $\mathcal{S}$ on H^N, by choosing appropriately the transition probabilities (in particular, they are indifferent among all such strategies). Those probabilities typically depend on the specified strategies on H^E. Yet the optimal strategies on H^E also depend on the transition probabilities. We will thus define transition probabilities and strategies on H^E simultaneously, by applying a fixed point theorem and use this to establish Theorem 1 for the case of two players.

Proof of Theorem 1 (2 players): Define $\overline{s}_1 \mid H_1^E$ and $\overline{s}_2 \mid H_2^E$ as the first two coordinates of a fixed point of a correspondence F from the set of all strategies $s_1 \mid H_1^E$, $s_2 \mid H_2^E$ and non-positive transfers π_1, π_2 into itself. Note that the set of all strategies $s_{-i} \mid H_{-i}^E$ can be identified with a convex subset of a finite-dimensional Euclidean space; similarly (non-positive) transfers π_i can be identified with a point of a finite-dimensional cube assuming they are bounded away from $-\infty$ by $\underline{\pi}$.

Consider the correspondence F defined by

$$F\left(\overline{s}_1 \mid H_1^E, \overline{s}_2 \mid H_2^E, \pi_1, \pi_2\right) = \left\{\left(s_1' \mid H_1^E, s_2' \mid H_2^E, \pi_1', \pi_2'\right)\right\}$$

as the set of (restricted) strategies and transfers such that: $s_i' \mid H_i^E$ is the restriction to H_i^E of a strategy of player i that is a best-response (in the auxiliary scenario) to player $-i$'s strategy that coincides with s_{-i}^G on H_{-i}^N and with $\overline{s}_{-i}$ on H_{-i}^E and to transfers π_i. The transfer π_i' is defined as the (non-positive) transfer π_i^G whose existence is established in Lemmas 1-2, for $\overline{s} \mid H^E = (\overline{s}_1 \mid H_1^E, \overline{s}_2 \mid H_2^E)$.

The set $F\left(\overline{s}_1 \mid H_1^E, \overline{s}_2 \mid H_2^E, \pi_1, \pi_2\right)$ is non-empty and convex, as the set of agent $-i$'s best-responses $s_{-i}' \mid H_{-i}^E$ is non-empty and convex and π_i^G is single-valued. The best-response correspondence is obviously upper hemi-continuous. Since π_i' is independent of π_i and, by Lemma 2, continuous with respect to $\overline{s} \mid H^E$, F is upper hemi-continuous.

Let $\left(\overline{s}_1 \mid H_1^E, \overline{s}_2 \mid H_2^E, \pi_1^G, \pi_2^G\right) \in F\left(\overline{s}_1 \mid H_1^E, \overline{s}_2 \mid H_2^E, \pi_1^G, \pi_2^G\right)$. By construction, playing any strategy s_i such that $s_i \mid H_i^N = \widetilde{s}_i \mid H_i^N$ for some $\widetilde{s}_i \in \mathcal{S}_i$ and $s_i \mid H_i^E = \overline{s}_i \mid H_i^E$ is a best-response against both $\overline{s}_{-i}^G$, π_i^G and $\overline{s}_{-i}^B$, π_i^B. It yields the payoffs close to v_i and $\underline{v}_i$, respectively, if ε is sufficiently close to 0; slightly perturbing the box $\prod_{i=1}^2[\underline{v}_i, v_i]$ if necessary, we can assume that the payoffs are exactly equal to v_i and $\underline{v}_i$.

We show that the payoff set $\prod_{i=1}^2[\underline{v}_i, v_i]$ can be achieved under almost perfect private monitoring. Divide the horizon of the infinitely repeated game into T-period blocks with the required properties. In particular, assume that the discount factor $\overline{\delta}$ is close enough to 1 so that $v_i + (1-\delta^T)\pi_i^G\left(h_i^T\right) > \underline{v}_i$ and $\underline{v}_i + (1-\overline{\delta}^T)\pi_i^B\left(h_i^T\right) < v_i$ for all $h_i^T \in H_i^T$.

For each i, construct a strategy for player $-i$ as follows. At the beginning of each block, player $-i$'s continuation strategy only depends on its state $u \in [\underline{v}_i, v_i]$. Thus, the state space is $[\underline{v}_i, v_i]$, the set of possible continuation payoffs of player i. The initial state is v_i.

At the beginning of each block, in state $u \in [\underline{v}_i, v_i]$, player $-i$ performs an initial randomization: for $q \in [0, 1]$ such that $u = qv_i + (1-q)\underline{v}_i$, he picks strategy $\overline{s}^G_{-i}$ with probability q and strategy $\overline{s}^B_{-i}$ with probability $1 - q$. Thus, player $-i$ uses one or the other strategy throughout the block as a function of the outcome of this initial randomization.

To define transition probabilities, consider first the case in which player $-i$ plays $\overline{s}^G_{-i}$ within the block. Given the realized history h^T_i in the block, player $-i$ transits to the state $v_i + (1 - \delta^T)\pi^G_i\left(h^T_i\right) \in [\underline{v}_i, v_i]$ at the end of the block; if player $-i$ plays $\overline{s}^B_{-i}$ within the block, then, at the end of the block, he transits to the state $\underline{v}_i + (1 - \delta^T)\pi^B_i\left(h^T_i\right) \in [\underline{v}_i, v_i]$.

It follows from the one-stage deviation property that, given the strategy of player $-i$, any strategy for player i such that, in every block, $s_i \mid H^N_i = \widetilde{s}_i \mid H^N_i$ for some $\widetilde{s}_i \in \mathcal{S}_i$ and $s_i \mid H^E_i = \overline{s}_i \mid H^E_i$ is a best-response. The payoff of player i is equal to the weighted average of the payoff to playing the best response against $\overline{s}^G_{-i}$ and the payoff to playing the best response against $\overline{s}^B_{-i}$ with weights q and $1 - q$. Thus, the payoff of player i in block-state u is $qv_i + (1-q)\underline{v}_i = u$. *Q.E.D.*

4 More players

When the stage game involves more than two players, two additional difficulties must be addressed. First, a player's opponents must coordinate their play, to punish effectively the player if necessary, but also to achieve the exact payoff that ensures the player's willingness to randomize over his continuation strategies at the beginning of the block. The construction must ensure that player i maintains the belief that his opponents coordinate their continuation strategies almost perfectly, whenever this coordination is essential to determine player i's best-response.

The second difficulty is an issue of dimensionality. With two players, it is possible to generate a neighborhood of the payoff vector to be achieved by considering only two strategies for each player. In particular, it is possible to guarantee that, whenever a player uses one particular strategy among the two, along with the corresponding transfer, his opponent is indifferent over all possible continuation strategies (within the block). This is an essential element of the construction, as it ensures that recent histories are a sufficient statistic to compute best-responses.

With three players or more, the straightforward generalization of the two-player construction consists in defining a set of strategies $\left\{\overline{s}_i^{M_1,\ldots,M_{i-1},M_{i+1},\ldots,M_n} : M_j \in \{G, B\},\, j \neq i\right\}$, for each player i, indexed by the payoff level, high or low, assigned to player $j \neq i$ by his opponents, assuming they coordinate.[7] That is, each player chooses from 2^{n-1} strategies. To replicate the two-player

[7] We do not know whether it is possible to prove the result, for all stage games, by considering a more restricted

construction, it would then be necessary to specify strategies and transfers such that, for every player i, and for all but one strategy profiles of player i's opponents, player i is indifferent over all his continuation strategies. Thus, player i must be indifferent over all continuation strategies (within the block), for some strategy profiles of his opponents that assign him a high payoff, as there are more than one such strategy profile. This is a demanding requirement for most stage games, as player i must then be willing to pick actions that yield low stage-game payoffs independently of his opponents' actions (a notable exception, here again, is the n-prisoner's dilemma, for which the difficulty does not arise).

Coordination is roughly[8] achieved as follows. Each player i is responsible for the payoff of his successor, player $i+1$ (identifying player $n+1$ and 1), and uses the history he observed and the strategy he played in the previous block to decide whether player $i+1$'s continuation payoff should be high or low. He then announces his decision through a choice of action (a message) in the initial period of the current block. The profile of such messages determines the strategy profile that should be played within the block. This, however, only ensures that players believe just after the initial period that coordination will obtain with high probability, and does not preclude that, after some recent histories, player i believe that this coordination has actually failed. To make sure that, whenever such coordination is payoff-relevant, player i maintains a high degree of confidence in his opponent successfully coordinating and keeping doing so, we introduce a second round of messages (through choices of actions) at the end of the block. If coordination among $i+1$'s opponents fails, player i learns (with high probability) about the cause of this failure at the end of the block, and he adjusts his transfer to make player $i+1$'s choice of action within the block payoff-irrelevant. The failure of coordination may relate to the strategy profile to be used by player $i+1$'s opponents within the block, but also to the date and identity of a unilateral deviator from the action profile to be chosen with high probability in each period of the block. To make sure that player i learns about the cause of the failure players need several periods of messages, but the duration of this communication may be taken arbitrarily short, relative to the actual play phase.

That is, each player may safely assume that his opponents coordinate (with high probability), even if his recent history suggests otherwise, as if they do not, he will be made indifferent across all his actions, once miscoordination is revealed at the end of the block.

This does not solve the second problem. While a player may condition on the event that his opponents coordinate, he still has beliefs about what strategy profile they are coordinating on. After non-erroneous histories (defined in a similar way to the previous section), he assigns probability almost 1 to the particular strategy profile consistent with his recent history, including the initial messages. After erroneous histories, however, his belief about this strategy profile may depend on his entire history, rather than on his recent history alone. To circumvent this difficulty, we add another initial round of communication, in which each player $i+1$ must "repeat" what

set of strategy profiles.

[8] The actual construction described in the next two subsections is slightly more complex.

strategy profile he believes his opponents intend to play. If he gets this wrong (more precisely, if the repeated strategy profile does not coincide to the intended strategy profile according to the signals of player i), transfers are adjusted so as to ensure that player $i+1$ is indifferent over all continuation strategies within the remaining periods of the block. But this transfer also guarantees that a player repeats what he truly believes his opponents intend to play.

This guarantees that, after any recent history, a player may not only assume that his opponents coordinate, but also that they coordinate on the strategy profile he repeated within that recent history, as he is indifferent over all continuation strategies (within the block) otherwise. In this way, we ensure that, after any history, the best-response only depends on the recent history. This device, that is useless with two players, comes at a cost when signal spaces more general than canonical are considered, as described in Section 5. Indeed, it requires that, when player $i+1$ repeats what he thinks his opponents intend to play, player $i+1$ assigns probability almost 1 to player i observing a "correct" signal (conditional on which player i assigns probability almost 1 to player $i+1$ communicating what he actually did), *conditional on any signal player $i+1$ may observe himself in that period*. If all other players use totally mixed actions whenever it is player $i+1$' turn to send this message, this necessarily obtains under the canonical signal space, provided the noise level is small enough. This fails, however, for some richer signal spaces.

4.1 Perfect Monitoring

Let v be the payoff vector from the interior of V^* that we wish to achieve. Take 2^n payoff vectors w^M, where $M = (M_1, ..., M_n)$, and $M_i \in \{G, B\}$ such that: (a) $w_i^M > v_i$ if $M_i = G$; (b) $w_i^M < v_i$ if $M_i = B$. There exist $\underline{v}_i < v_i$ such that the interior of $\text{co}(\{w^M : M \in \{G, B\}^n\})$ contains the box

$$[\underline{v}_1, v_1] \times \cdots \times [\underline{v}_n, v_n].$$

As in the two-player case, assume that there exist pure action profiles a^M such that

$$u_i(a^M) = w_i^M.$$

Otherwise replace action profiles by finite sequences of action profiles, as described in Section 3. We follow the notational conventional of Section 3. We show that the payoff set $\prod_{i=1}^n [\underline{v}_i, v_i]$ can be achieved in block strategies.

The play in a block will be divided into *five* phases.

Phase 1 (period 1): every player i simultaneously sends message G or B. For each i, we partition A_i into two non-empty sets G_i and B_i. If, in the first period of a block $(t = 1)$, player i chooses an action from the first set, we say that he sends message $M_i = G$; otherwise, we say that he sends message $M_i = B$. In all strategies of player i defined below, it is understood that he sends message M_i by uniformly randomizing over all actions in the corresponding element of the partition, G_i or B_i.

Phase 2 (periods $t = 2, ..., n(n-1)+1$): players consecutively "report" the message vector that corresponds to all signals they have observed in period 1 (they do not report their own message). That is, player 1 first *reports these signals*, while all other players uniformly randomize over all their actions; then player 2 reports his signals, while all other players uniformly randomize over all their actions, and so forth. More formally, this is achieved by using the partitions described in phase 1. To report the signals player i observed in the initial period, that imply message profile $M_{-i} = (M_1, ..., M_{i-1}, M_{i+1}, ..., M_n)$, $M_j \in \{G, B\}$ all $j \neq i$, player i randomizes uniformly over all actions in A_i in all periods of Phase 2, *except* in the $n-1$ periods $(i-1)(n-1)+2, \ldots, i(n-1)+1$, in which he consecutively randomizes over all actions in the element of the partition that corresponds to the message $M_j \in \{G, B\}$. For instance, in period $(i-1)(n-2)+2$, player $i > 1$ sends message $M_i = G$ (respectively B) if the signal he obtained about player 1 in the initial period is an element of G_1 (respectively B_1), and so forth.

Phase 3 (periods $t = n(n-1)+2, ..., 2n(n-1)+1$): players "repeat" consecutively their predecessors' reports. That is, first player 2 repeats what he observed player 1 report in Phase 2, while all other players uniformly randomize over all the actions in their action sets; next player 3 repeats what he observed player 2 report in Phase 2, and so forth. By convention, player n is the predecessor of player 1. The details are the straightforward analogues of those in Phase 2, and are therefore omitted.

Phase 4 (periods $t = 2n(n-1)+1, ..., T-k$): see below for the specification.

Phase 5 (periods $T-k+1, T-k+2, ..., T$): in period $T-k+1$, every player i sends first the message $M_i = G$ or B. Then player i reveals his signals about the message profile M sent in Phase 1, whether all players (according to his signals) were playing $a^{\widetilde{M}}$ in every period of Phase 4, where $\widetilde{M} = (M_n, M_1, M_2, \ldots, M_{n-1})$. [$\widetilde{M}$ has been used (instead of M) as the superscript of the action profile, because player $i+1$'s payoff depends on player i's message and the $i+1$-st coordinate of the superscript reflects whether player $i+1$'s payoff is high or low.] If he reports a deviation from $a^{\widetilde{M}}$, then player i announces if the first such deviation was unilateral, in which case he also reports: (i) the identity of the player who first deviated and (ii) the period in which this first deviation occurred. All these announcements take place in periods $T-k+2, ..., T$, where k is chosen such that all such reports can be completed. Without loss of generality, we can take k of order $\log T$. All elements are trivial to report, except the period in which the first unilateral deviation from $a^{\widetilde{M}}$ occurred, if any. This period can be announced in no more than $1 + \log T$ periods by dichotomous signaling: the first message signals whether the deviation occurred in the first half or the second half of Phase 4; once this half is determined, the second message further signals whether the deviation occurred in the first half or the second half within that half, and so forth.

The play in Phase 4 is determined by the messages sent in Phase 1. As before, we first define $\mathcal{S}_{i+1} \subset S_{i+1}^T$, $i = 0, 1, ..., n-1$, as the sets of all (behavior) strategies s_{i+1} in the T-period repeated game satisfying the following condition:

($^*{}_n$) In Phase 3, s_{i+1} repeats player i's report from Phase 2, as inferred from player $i+1$'s private history. Suppose M is the message profile sent in Phase 1, according to player $i+1$'s private history. If $M_i = G$, then player $i+1$ plays $a_{i+1}^{\widetilde{M}}$ in all periods of Phase 4 provided that, according to player $i+1$'s private history, players have played $a^{\widetilde{M}}$ in all periods $\bar{t}$, $2n(n-1)+1 \leq \bar{t} < t$.

Note that no restriction is imposed on the actions chosen in Phases 1, 2 and 5. To avoid clutter, the condition ($^*{}_n$) above is informal in two respects. The first sentence is not meant to imply that strategy s_{i+1} specifies uniform randomization over the actions in the appropriate element in the partition, as described above, but simply that s_{i+1} assigns probability 0 to any action in the other element of the partition. That is, the condition only restricts the support of the actions specified in Phase 3. Second, "as inferred from player $i+1$'s private history" and "according to player $i+1$'s private history" refer to the fact that s_{i+1} is defined relative to $i+1$'s private history only, so that the condition $\mathcal{S}_{i+1}$ remains well-defined under imperfect monitoring.

As in the case of two players, we say that $\mathcal{S}_i$ *prescribes* some set of actions $\mathcal{A}_i\left(h_i^t\right)$, given some private history h_i^t, if $\mathcal{A}_i\left(h_i^t\right)$ is the set of all actions a_i such that, for some strategy $s_i \in \mathcal{S}_i$, s_i assigns positive probability to action a_i conditional on history h_i^t. That is, given h_i^t,

$$\mathcal{A}_i\left(h_i^t\right) = \left\{a_i \in A_i : \exists_{s_i \in \mathcal{S}_i} \quad s_i\left[h_i^t\right](a_i) > 0\ \right\}.$$

In Phases 1, 2, and 5, this prescribed set coincides with A_i; in Phase 4, it is either A_i or the singleton $a_i^{\widetilde{M}}$ (in which case we say that $\mathcal{S}_i$ *prescribes* $a_i^{\widetilde{M}}$); in Phase 3, this prescribed set always coincides with an element of the partition of the set of player i's actions, as described in Phase 1.

Given some (small) $\rho > 0$, define $\mathcal{S}_i^\rho$ as the set of strategies $s_i \in \mathcal{S}_i$ that, after every history h_i^t, and for all actions $a_i \in \mathcal{A}_i\left(h_i^t\right)$ assign probability at least ρ to action a_i, i.e.

$$\mathcal{S}_i^\rho = \left\{s_i \in \mathcal{S}_i : \forall_{h_i^t}\ \forall_{a_i \in \mathcal{A}_i\left(h_i^t\right)} \quad s_i\left[h_i^t\right](a_i) \geq \rho\right\}.$$

We shall show now that if T is large enough, then there are strategies s_i^g, $s_i^b \in \mathcal{S}_i$ such that player $i+1$'s average payoff from any strategy $s_{i+1} \in \mathcal{S}_{i+1}$ against s_i^g and $s_j \in \{s_j^g, s_j^b\}$ for $j \neq i, i+1$ is higher than v_{i+1}; and player $i+1$'s average payoff from any strategy $s_{i+1} \in S_i^T$ (including strategies from the complement of $\mathcal{S}_{i+1}$) against s_i^b and $s_j \in \{s_j^g, s_j^b\}$ for $j \neq i, i+1$ is lower than $\underline{v}_{i+1}$.

The strategies s_i^g and s_i^b only differ in the actions taken in Phase 1 and the first period of Phase 5; s_i^g sends message G and s_i^b sends message B. In Phase 2 (respectively, Phase 3), both strategies specify that player i reports what he observed in Phase 1 (respectively, repeats what he observed in Phase 2), as described above; in Phase 4, given the message profile he observed in Phase 1, both strategies specify $a_i^{\widetilde{M}}$, where $\widetilde{M} = (M_n, M_1, M_2, \ldots, M_{n-1})$, until the first deviation from $a^{\widetilde{M}}$, if this first deviation is unilateral. [In particular, both strategies specify $a_i^{\widetilde{M}}$ in case of a simultaneous deviation.] If a player $j \neq i$ unilaterally deviates from $a_j^{\widetilde{M}}$, then both strategies specify i's action in action profile α_j^*, minmaxing player j. If player i unilaterally deviates from $a_i^{\widetilde{M}}$, then he plays in all remaining periods the best-response to the minmaxing profile α_i^* (if there is more than one such best-response, pick one of them). Finally in Phase 5, players first repeat their messages from Phase 1, and they next honestly communicate what they are supposed to communicate in Phase 5.

To determine average payoffs of player $i+1$ against s_i^g and s_i^b, observe that, for sufficiently large T, these average payoffs are approximately equal to the average payoffs in Phase 4. If player i plays s_i^b, then player $i+1$ can obtain a (per-period) payoff above $\underline{v}_{i+1}$ in at most one period of Phase 4 (the period in which he unilaterally deviates). If player i plays s_i^g and player $i+1$ plays a strategy $s_{i+1} \in \mathcal{S}_{i+1}$, then player $i+1$ cannot obtain (in Phase 4) a payoff below v_{i+1}.

Perturb slightly strategies s_i^g and s_i^b so that, after any history h_i^t, player i plays each action $a_i \in \mathcal{A}_i(h_i^t)$ with probability at least ρ; if $\rho > 0$ is small enough, we obtain strategies s_i^G, $s_i^B \in \mathcal{S}_i^\rho$ such that the average payoff of player $i+1$ to playing any strategy $s_{i+1} \in \mathcal{S}_{i+1}$ against s_i^G and $s_j \in \{s_j^G, s_j^B\}$ for $j \neq i, i+1$ is higher than v_{i+1}, and the average payoff of player $i+1$ to playing any strategy s_{i+1} (including strategies from the complement of $\mathcal{S}_{i+1}$) against s_i^B and $s_j \in \{s_j^G, s_j^B\}$ for $j \neq i, i+1$ is lower than $\underline{v}_{i+1}$.

We denote, for $s_j \in \{s_j^G, s_j^B\}$ for $j \neq i, i+1$, by $r_{i+1}^G(s_j, j \neq i, i+1)$ the strategy $s_{i+1} \in \mathcal{S}_{i+1}$ such that, for every history h_{i+1}^t, the strategy $s_{i+1}|h_{i+1}^t$ yields the lowest payoff against s_i^G and s_j for $j \neq i, i+1$ among all strategies $s_{i+1} \in \mathcal{S}_{i+1}$. Similarly, we denote by $r_{i+1}^B(s_j, j \neq i, i+1)$ a strategy $s_{i+1} \in S_{i+1}^T$ such that for every history h_{i+1}^t the strategy $s_{i+1}|h_{i+1}^t$ yields the highest payoff against s_i^B and s_j for $j \neq i, i+1$ among all strategies $s_{i+1} \in S_{i+1}^T$.

We now generalize the proof of the folk theorem under perfect monitoring given in section 3.1 to $n \geq 2$.

Proof of Theorem 1 (perfect monitoring): For each i, construct a strategy for player i as follows. At the beginning of each block, player i's continuation strategy only depends on its state $u \in [\underline{v}_{i+1}, v_{i+1}]$. Thus, the state space is $[\underline{v}_{i+1}, v_{i+1}]$, the set of possible continuation payoffs of player $i+1$. The initial state is v_{i+1}.

At the beginning of each block, in state $u \in [\underline{v}_{i+1}, v_{i+1}]$, player i performs an initial randomization: for $q \in [0,1]$ such that $u = qv_{i+1} + (1-q)\underline{v}_{i+1}$, he picks strategy s_i^G with probability q and strategy s_i^B with probability $1-q$. Thus, player i uses one or the other strategy throughout the block as a function of the randomization. For later purposes, we refer to the outcome of this randomization as the *intention* of player i. Thus, player i *intends to play* s_i^B if s_i^B is the outcome of this randomization.

To define transition probabilities, suppose first that player i intends to play s_i^B. Then he records the periods in which player $i+1$ departs from $r_{i+1}^B(s_j, j \neq i, i+1)$, where $s_j = s_j^{M_j}$, $M_j \in \{G, B\}$ stands for the message that has been sent in the first period of Phase 5, for $j \neq i+1$. Let θ_t^B denote the difference between player $i+1$'s unnormalized continuation payoff from $r^B_{i+1}(s_j, j \neq i, i+1)$ from period t on, and player $i+1$'s unnormalized continuation payoff from playing the action chosen by player $i+1$, followed by switching to $r^B_{i+1}(s_j, j \neq i, i+1)$ from period $t+1$ on. By definition of $r^B_{i+1}(s_j, j \neq i, i+1)$, $\theta^B_t \geq 0$ for every action of player $i+1$. At the end of the block, player i then transits to the state:

$$\underline{v}_{i+1} + \frac{1-\delta}{\delta^T}\left(x^B_{i+1}(s_j, j \neq i, i+1) + \sum_{t=1}^{T} \delta^{t-1}\theta^B_t\right), \tag{7}$$

where $x^B_{i+1}(s_j, j \neq i, i+1)$ is the difference between getting $\underline{v}_{i+1}$ in every of T periods and player $(i+1)$'s payoff from playing $r^B_{i+1}(s_j, j \neq i, i+1)$, which is less than v_{i+1} if δ is large enough. Observe that, if player $i+1$ knew that player i was playing s^B_i, every strategy s_{i+1} of player i would be a best-response.

Suppose now that player i intends to play s^G_i. We consider three cases. In the first two cases, the strategy of player i is defined similarly to the case in which player i intends to play s^B_i.

Case i: The message vector sent by player $i+1$ in Phase 3 differs from the message vector reported by player i in Phase 2.

Case ii: The message vector sent by player $i+1$ in Phase 3 coincides with the message vector reported by player i in Phase 2, but the message vector reported by player i in Phase 2 differs from the message vector sent in Phase 5.

Case iii: The message vector reported by player i in Phase 2 coincides with the message vector sent in Phase 5, and the message vector sent by player $i+1$ in Phase 3 coincides with the message vector reported by player i in Phase 2.

In Cases i and ii, player i picks first a "target" transition state $v_{i+1} - \zeta \in (\underline{v}_{i+1}, v_{i+1})$, and then he picks the "actual" transition states as when he intends to play s^B_i, in order to make player $i+1$ indifferent across all strategies $s_{i+1} \in S^T_{i+1}$. As the payoff of player $i+1$ from any strategy in S^t_{i+1} is $\underline{v}_{i+1}$ when player i intends to play s^B_i, now player $i+1$'s payoff from any strategy is $v_{i+1} - \zeta$.

In Case iii, player i records the periods in which player $i+1$ departs from $r^G_{i+1}(s_j, j \neq i, i+1)$, where again $s_j = s^{M_j}_j$, $M_j \in \{G, B\}$ stands for the message that has been sent in Phase 5, for $j \neq i+1$. Let θ_t denote the difference between the unnormalized continuation payoff to playing the action chosen by player $i+1$, followed by switching to $r^G_{i+1}(s_j, j \neq i, i+1)$ from period $t+1$ on, and the unnormalized continuation payoff of playing $r^G_{i+1}(s_j, j \neq i, i+1)$ from period t on. Let $\theta^G_t = \max\{0, \theta_t\}$. At the end of the block, player i transits to the state:

$$v_{i+1} - \frac{1-\delta}{\delta^T}\left(x^G_{i+1}(s_j, j \neq i, i+1) + \sum_{t=1}^{T} \delta^{t-1}\theta^G_t\right), \tag{8}$$

where $x^G_{i+1}(s_j, j \neq i, i+1)$ is the difference between player $(i+1)$'s payoff from playing $r^G_{i+1}(s_j, j \neq i, i+1)$ and getting v_{i+1} in every of T periods, which falls below v_{i+1} if δ is large enough. Observe that, if player $i+1$ knew that player i was playing s_i^G, every strategy $s_{i+1} \in \mathcal{S}_{i+1}$ would be a best-response.

By construction and the one-stage deviation property, given the strategy of player $(i+1)$'s opponents, any strategy for player $i+1$ which belongs to $\mathcal{S}_{i+1}$ in every block is a best-response. The payoff of player $i+1$ in block-state u is equal to the weighted average $qv_{i+1}+(1-q)\underline{v}_{i+1}=u$. *Q.E.D.*

Note that neither Phase 2 and 3 nor Phase 5 played any role in our construction of block-strategy equilibria under perfect monitoring. Under imperfect private monitoring, the information transmitted in Phase 5 (except the first period) will allow player i to pick the transition probabilities that make player $i+1$ indifferent across all strategies when he intends to play s_i^B. Note the difference with the two-player case. To make player $i+1$ indifferent across all actions conditional on some history h_{i+1}^t, player i must know the (possibly mixed) actions of players other than $i+1$ in period t as well as their continuation strategies; more precisely, player $i+1$ must believe that player i will know when he determines the transition probabilities. It will be (approximately) achieved (for $\varepsilon \to 0$) through the information transmitted in Phase 5.

Making player $i+1$ indifferent across all strategies when player i intends to play s_i^B will allow player $i+1$ with an erroneous history to play as if he knew that player i's intention is to play s_i^G. Then Phases $2-3$ and the first period of Phase 5 will further allow player $i+1$ with an erroneous history to play as if he knew the intentions of all other players. The details will only become clear in the following subsection.

4.2 Imperfect Private Monitoring

We call a history h_i^t *erroneous* if, under perfect monitoring, it is a history off the path for every strategy profile from $\mathcal{S}$; otherwise the history is called *non-erroneous*. As in the case of two players, let $H_i^{N,t}$ denote the set of all non-erroneous t-length histories, and let $H_i^{E,t}=H_i^t \backslash H_i^{N,t}$ denote the complement of $H_i^{N,t}$. Let

$$H_i^N=\bigcup_{t\leq T} H_i^{N,t},\ H_i^E=\bigcup_{t\leq T} H_i^{E,t}.$$

We will show that, if player i's intention is to play s_i^B (more precisely, player i intends to play s_i^B on H_i^N and a given strategy on H_i^E), then he can pick the transition probabilities so that for any intentions s_j^M, where $j \neq i, i+1$ and $M \in \{G,B\}$, player $i+1$ is indifferent across all strategies $s_{i+1} \in S_{i+1}^T$ in the T-period repeated game. Simultaneously, assuming now that player i's intention is to play s_i^G (it again means that player i intends to play s_i^G on H_i^N and a given strategy on H_i^E), player i can pick the transition probabilities so that after histories from H_{i+1}^N, player $i+1$ is indifferent across all strategies from $\mathcal{S}_{i+1}$ (for any intentions of other players) and he weakly prefers any of them to any other strategy; moreover, he can pick the transition probabilities so that player $i+1$ is indifferent across all strategies in Phase 4 conditional on the following two events:

1. the message profile *reported by player* i in Phase 2 does not coincide with *the intention profile* revealed by players other than $i+1$ in the first period of Phase 5,

2. the message profile *reported by player* i in Phase 2 does coincide with *the intention profile* revealed by players other than $i+1$ in the first period of Phase 5, but the message profile *sent by player i+1* in Phase 3 does not coincide with the message profile *reported by player* i in Phase 2.

Notice that our construction will guarantee that player $(i+1)$'s set of best replies on erroneous histories contains the strategies that would be best-responses if he knew that player i intended to play s_i^G and the intentions of all players $j \neq i, i+1$ coincided with the message profile sent by himself in Phase 3.

As in the case of two players, we consider the auxiliary scenario, where players play the T-period repeated game and then each of them obtains a transfer that is a function of player i's private history. Recall that $B_{i+1}(s_{-(i+1)}, \pi_{i+1})$ denotes the set of auxiliary scenario best-responses of player $i+1$ to the T-period strategy profile $s_{-(i+1)}$ of his opponents and the transfer function π_{i+1}; given a T-period strategy $s_{-(i+1)}$, a transfer function π_{i+1} and a strategy $\overline{s}_{i+1} \in S_{i+1}^T$, let $B_{i+1}(s_{-(i+1)}, \pi_{i+1} \mid \overline{s}_{i+1})$ denote the set of strategies that maximize player $(i+1)$'s auxiliary-scenario payoff against $s_{-(i+1)}$, π_{i+1} among all strategies $s_{i+1} \in S_{i+1}^T$ such that $s_{i+1} \mid H_{i+1}^E = \overline{s}_{i+1} \mid H_{i+1}^E$. Recall finally that $U_{i+1}^A(s_{i+1}, s_{-(i+1)}, \pi_{i+1})$ denotes the average payoff of player $i+1$ from s_{i+1} against $s_{-(i+1)}$, π_{i+1}, while $U_{i+1}^T(s)$ stands for the payoff in T-period repeated game under perfect monitoring given strategy profile s. The following Lemma is the counterpart, for $n \geq 2$, of Lemmata 1, 2 and 3.

Lemma 4 *For every strategy* $\overline{s} \mid H^E$*, there exists* $\overline{\varepsilon} > 0$ *such that for* $\varepsilon < \overline{\varepsilon}$*:*

(a) *There exist non-negative transfers* $\pi_{i+1}^B : H_i^T \to \mathbf{R}_+$ *such that for every* $M = (M_1, ..., M_n) \in \{G, B\}^n$ *with* $M_i = B$

$$S_{i+1}^T = B_{i+1}(\overline{s}_{-(i+1)}^M, \pi_{i+1}^B), \tag{9}$$

where $\overline{s}_j^M \mid H_j^N = s_j^{M_j} \mid H_j^N$ *and* $\overline{s}_j^M \mid H_j^E = \overline{s}_j \mid H_j^E$*, and for every* $s_{i+1} \in B_{i+1}(\overline{s}_{-(i+1)}^M, \pi_{i+1}^B)$

$$\lim_{\varepsilon \to 0, T \to \infty} U_{i+1}^A(s_{i+1}, \overline{s}_{-(i+1)}^M, \pi_{i+1}^B) = \lim_{T \to \infty} \max_{\widetilde{s}_{i+1} \in S_{i+1}^T} U_{i+1}^T(\widetilde{s}_{i+1}, \overline{s}_{-(i+1)}^M). \tag{10}$$

(b) *There exist non-positive transfers* $\pi_{i+1}^G : H_i^T \to \mathbf{R}_-$ *such that such that for every* $M = (M_1, ..., M_n) \in \{G, B\}^n$ *with* $M_i = G$

$$\begin{aligned} \{s_{i+1} \in S_{i+1}^T : s_{i+1} \mid H_{i+1}^N = \widetilde{s}_{i+1} \mid H_{i+1}^N \text{ for some } \widetilde{s}_{i+1} \in \mathcal{S}_{i+1} \text{ and} \\ s_{i+1} \mid H_{i+1}^E = \overline{s}_{i+1} \mid H_{i+1}^E\} \subset B_{i+1}(\overline{s}_{-(i+1)}^M, \pi_{i+1}^G \mid \overline{s}_{i+1}), \end{aligned} \tag{11}$$

where $\overline{s}_j^M \mid H_j^N = s_j^{M_j} \mid H_j^N$ *and* $\overline{s}_j^M \mid H_j^E = \overline{s}_j \mid H_j^E$*, and for every* $s_{i+1} \in B_{i+1}(\overline{s}_{-(i+1)}^M, \pi_{i+1}^G \mid \overline{s}_{i+1})$

$$\lim_{\varepsilon \to 0, T \to \infty} U_{i+1}^A(s_{i+1}, \overline{s}_{-(i+1)}^M, \pi_{i+1}^G) = \lim_{T \to \infty} \min_{\substack{\widetilde{M} \in \{G,B\}^n \text{ with } \widetilde{M}_i = G \\ \widetilde{s}_{i+1} \in \mathcal{S}_{i+1}}} U_{i+1}^T(\widetilde{s}_{i+1}, \overline{s}_{-(i+1)}^{\widetilde{M}}); \tag{12}$$

π_{i+1}^G *is bounded away from* $-\infty$*, i.e. there exists* $\underline{\pi}$ *(independent of* $\overline{s}$*) such that* $\pi_{i+1}^G \geq \underline{\pi}$*, and* π_{i+1}^G *depends continuously on* $\overline{s}$*.*

(c) *Moreover, every strategy in Phase* 4 *yields player* $i+1$ *the same payoff conditional on each of the following two events:*

1*. The message profile reported by player* i *in Phase* 2 *does not coincide with the intention profile of players other than* $i+1$ *revealed in the first period of Phase* 5*;*

2*. The message profile reported by player* i *in Phase* 2 *does coincide with the intention profile of players other than* $i+1$ *revealed in the first period of Phase* 5*, but the message profile sent by player* $i+1$ *in Phase* 3 *does not coincide with the message profile reported by player* i *in Phase* 2*.*

Proof: in Appendix.

Increasing π_{i+1}^B by a constant that depends only on the message profile sent in the first period of Phase 5 if necessary, we can assume, instead of equation (10), that

$$U_{i+1}^A(s_{i+1}, \overline{s}_{-(i+1)}^M, \pi_{i+1}^B) = \underline{v}_{i+1}$$

for every $M = (M_1, M_2, ..., M_n) \in \{G, B\}^n$ such that $M_i = B$. Similarly, decreasing π_{i+1}^G by a constant that depends only on the message profile sent in first period of Phase 5, we may assume, instead of (12), that

$$U_{i+1}^A(s_{i+1}, \overline{s}_{-(i+1)}^M, \pi_{i+1}^G) = v_{i+1}$$

for every $M = (M_1, M_2, ..., M_n) \in \{G, B\}^n$ such that $M_i = G$. We may now prove Theorem 1 in full generality.

Proof of Theorem 1: Define $\overline{s}_i \mid H_i^E$ for $i = 1, 2, ..., n$ as the first n coordinates of a fixed point of a correspondence F from the set of all strategies $s_i \mid H_i^E$ and non-positive transfers π_{i+1}, $i = 1, 2, ..., n$, into itself. Note that the set of all strategies $s_i \mid H_i^E$ can be identified with a convex subset of a finite-dimensional Euclidean space; similarly (non-positive) transfers π_{i+1} can be identified with a point of a finite-dimensional cube assuming they are bounded away from $-\infty$ by $\underline{\pi}$.

Consider the correspondence F defined by

$$F\left(s_i \mid H_i^E, \pi_{i+1}, i = 1, 2, ..., n\right) = \left\{\left(s_i' \mid H_i^E, \pi_{i+1}', i = 1, 2, ..., n\right)\right\}$$

as the set of (restricted) strategies and transfers such that: $s'_{i+1} \mid H^E_{i+1}$ is a strategy of player $i+1$ that is a best-response to his opponents' strategy $\left(s^M_{-(i+1)} \mid H^N_{-i}, \overline{s}_{-(i+1)} \mid H^E_{-(i+1)}\right)$ in the auxiliary scenario, where M is the sequence of signals about player i's report obtained by player $i+1$ in Phase 2, and to transfers π_{i+1}. The definition of $s'_{i+1} \mid H^E_{i+1}$ is correct, because no history in Phases 1 and 2 is erroneous, and therefore player $i+1$ knows M at every history $h^t_{i+1} \in H^E_{i+1}$. The transfer π'_{i+1} is defined as the (non-positive) transfer π^G_{i+1} whose existence is established in Lemma 4(b) for $\overline{s} \mid H^E = (s_i \mid H^E_i, i = 1, 2, ..., n)$.

The set $F\left(s_i \mid H^E_i, \pi_{i+1}, i = 1, 2, ..., n\right)$ is non-empty and convex, and F is upper hemi-continuous by the same argument as in the case of two players.

Let $\left(\overline{s}_i \mid H^E_i, \pi^G_{i+1}, i = 1, 2, ..., n\right) \in F\left(\overline{s}_i \mid H^E_i, \pi^G_{i+1}, i = 1, 2, ..., n\right)$ be any fixed point. Notice that any strategy of player $i+1$ is his best-response conditional on s^B_i being the intention of player i (by Lemma 4(a)), as well as conditional on s^G_i being the intention of player i but $s^M_{-(i+1)}$, where M is the sequence of signals about player i's report obtained by player $i+1$ in Phase 2, not being the intention of player $i+1$'s opponents (by Lemma 4(c)). This yields, by the definition of $s'_{i+1} \mid H^E_{i+1}$, that playing any strategy s_{i+1} such that $s_{i+1} \mid H^N_{i+1} = \widetilde{s}_{i+1} \mid H^N_{i+1}$ for some $\widetilde{s}_{i+1} \in \mathcal{S}_{i+1}$ and $s_{i+1} \mid H^E_{i+1} = \overline{s}_{i+1} \mid H^E_{i+1}$ is a best-response against $\overline{s}^{\overline{M}}_{-(i+1)}, \pi^B_{i+1}$ and $\overline{s}^{\overline{M}}_{-(i+1)}, \pi^G_{i+1}$ for every set of intentions $\overline{M}$. It yields the payoffs no higher than $\underline{v}_{i+1}$ and no lower than v_{i+1}, respectively, if ε is sufficiently close to 0.

We show that the payoff set $\prod_{i=1}^n [\underline{v}_i, v_i]$ can be achieved under almost perfect private monitoring. Divide the horizon of the infinitely repeated game into large enough T-period blocks. Construct a strategy for player i as follows. The state of player i's strategy at the beginning of a block is player $i+1$'s continuation payoff. Player i's initial block-state is v_{i+1}. At the beginning of each block, in state $u \in [\underline{v}_{i+1}, v_{i+1}]$, player i performs an initial randomization: for $q \in [0, 1]$ such that $u = qv_{i+1} + (1-q)\underline{v}_{i+1}$, he picks strategy $\overline{s}^G_i$ with probability q and strategy $\overline{s}^B_i$ with probability $1-q$. Thus, player i uses one or the other strategy throughout the block as a function of the randomization. Pick $\bar{\delta}$ close enough to one such that $v_{i+1} + (1-\bar{\delta}^T)\pi^G_{i+1} > \underline{v}_{i+1}$ and $\underline{v}_{i+1} + (1-\bar{\delta}^T)\pi^B_{i+1} < v_{i+1}$ for all histories. If he plays $\overline{s}^G_i$, then at the end of the block he transits to the state $v_{i+1} + (1-\delta^T)\pi^G_{i+1} \in [\underline{v}_{i+1}, v_{i+1}]$; if he plays $\overline{s}^B_i$, then at the end of the block he transits to the state $\underline{v}_{i+1} + (1-\delta^T)\pi^B_{i+1} \in [\underline{v}_{i+1}, v_{i+1}]$.

It follows from the one-stage deviation property that, given the strategy of player i, any strategy for player $i+1$ such that, in every block, $s_{i+1} \mid H^N_{i+1} = \widetilde{s}_{i+1} \mid H^N_{i+1}$ for some $\widetilde{s}_{i+1} \in \mathcal{S}_{i+1}$ and $s_{i+1} \mid H^E_{i+1} = \overline{s}_{i+1} \mid H^E_{i+1}$ is a best-response. The payoff of player $i+1$ is equal to the weighted average $qv_{i+1} + (1-q)\underline{v}_{i+1} = u$. *Q.E.D.*

5 Extensions and concluding comments

Throughout, attention has been restricted to the case in which $\Sigma_i = A_{-i}$. As our focus is on almost-perfect monitoring, it makes little sense to consider signal spaces for which, for some i,

$\#\Sigma_i < \#A_{-i}$. However, convergence to perfect monitoring can be defined for signal spaces that have more signals than opponents' action profiles.

Following Ely and Välimäki (2002), given $\{m(\cdot \mid a) : a \in A\}$ (and its finite domain Σ), we say that a monitoring structure is ε-perfect if, for each player i, there exists a partition of Σ_i into $\{\Sigma_i^{a_{-i}} : a_{-i} \in A_{-i}\}$ such that, for all $a_i \in A_i$,

$$\sum_{\sigma_i \in \Sigma_i^{a_{-i}}} m_i(\sigma_i \mid (a_i, a_{-i})) \geq 1 - \varepsilon.$$

A close look at the proof should convince the reader that for $n = 2$ our folk theorem (Theorem 1) remains valid if we use the property given above in the definition of convergence. In fact, this partition could further depend on a_i. If Σ_i is infinite (but A_i is finite), Theorem 1 still holds for $n = 2$ with this more general definition.

For $n > 2$, our proof requires, however, an additional assumption on monitoring structure; namely, for any player $i + 1$, any action profile $(a_{i+1}, \alpha_{-(i+1)})$, where a_{i+1} is a pure action and $\alpha_{-(i+1)}$ a (fixed) totally mixed action profile of other players, and for every signal $\sigma_{i+1} \in \Sigma_{i+1}$ observed with positive probability, player $i + 1$ assigns a probability that tends to 1 as $\varepsilon \to 0$ to $\sigma_i \in \Sigma_i^{a_{-i}}$ for some a_{-i} whose i-th coordinate is a_{i+1} (conditional on $(a_{i+1}, \alpha_{-(i+1)})$ being played and σ_{i+1} being the signal of player $i + 1$). In words, player $i + 1$ must be sure that player i received a signal that is evidence of the action he actually took, independently of his own signal.

We need this additional requirement to give player $i + 1$ an incentive to repeat truthfully in Phase 3 the observed report of player i from Phase 2 (see the proof of Lemma 4 for details). The requirement is always satisfied for canonical signal spaces. When it is satisfied Theorem 1 holds even for infinite signal spaces Σ_i, although the requirement itself seems much stronger when Σ_i is infinite. We do not know whether Theorem 1 holds under the weaker notion of convergence. We do not know either of any tractable modification of our proof that would apply to the case in which the action sets A_i themselves are infinite.[9]

It may seem feasible to combine Theorem 1 with the result of Matsushima (2004), to establish the folk theorem for all games, not only for almost-perfect monitoring, but also for monitoring structures that are not almost perfect, but satisfy conditional independence. The obvious route would consist in considering rounds of blocks, with the same strategy being used in each block of a given round, and players switching or retaining that strategy at the end of the round by using some summary statistics obtained from the round. While this may be possible, we must point out a serious difficult in this endeavor. In Matsushima (2004), conditional independence is useful because it ensures that, within a round, a player's signal does not affect his belief about his probability of failing or passing the statistical test, as this probability depends on his rival's signals. If each period within the round is replaced by a block, this property is not preserved:

[9] The difference between the two cases is somewhat similar to the distinction in Mailath and Morris (2004) between ε-close and strongly ε-close monitoring structures, the benchmark in our case being perfect monitoring, rather than public monitoring.

in the second period of a block, the signal observed by a player affects his belief over the signal received by his opponent in the previous period, since his opponent's continuation strategy did depend on that signal. Hence, within a round, the signals of a player affect his probability of failing or passing the statistical test, even when signals are conditionally independent. This suggests that it may preferable to first generalize the folk theorem for the two-player prisoner's dilemma, to monitoring structures that satisfy weaker requirements, before considering more general stage games.

References

Abreu, D., D. Pearce, and E. Stacchetti (1990): "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring," *Econometrica*, 58, 1041–1063.

Aoyagi M. (2002): "Collusion in Dynamic Bertrand Oligopoly with Correlated Private Signals and Communication," *Journal of Economic Theory* 102 (2002), 229–248.

Bhaskar, V., and I. Obara (2002): "Belief-Based Equilibria in the Repeated Prisoners' Dilemma with Private Monitoring," *Journal of Economic Theory*, 102, 40–69.

Compte O. (1998): "Communication in Repeated Games with Imperfect Private Monitoring," *Econometrica* 66, 597–626.

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 (2004), "Belief-free Equilibria in Repeated Games", *Econometrica*, forthcoming.

Fudenberg D. and D. Levine (1991): "An Approximate Folk Theorem with Imperfect Private Information," *Journal of Economic Theory*, 54, 26–47.

Fudenberg D. and D. Levine (2002): "The Nash Threats Folk Theorem With Communication and Approximate Common Knowledge in Two Player Games", Mimeo.

Fudenberg, D., D. Levine, and E. Maskin (1994): "The Folk Theorem with Imperfect Public Information," *Econometrica*, 62(5), 997–1040.

Fudenberg D. and E. Maskin (1986): "The Folk Theorem in Repeated Games with Discounting or with Incomplete Information," *Econometrica*, 54, 533–554.

Kandori, M., and H. Matsushima (1998): "Private Observation, Communication and Collusion," *Econometrica*, 66, 627–652.

Kandori, M., and I. Obara (2004): "Efficiency in Repeated Games Revisited: The Role of Private Strategies," Mimeo.

Lehrer, E. (1990): "Nash Equilibria of n-player Repeated Games with Semi-Standard Information," *International Journal of Game Theory*, 19, 191–217.

Mailath, G. J., and S. Morris (2002): "Repeated Games with Almost-Public Monitoring," *Journal of Economic Theory*, 102, 189–228.

Mailath, G. J., and S. Morris (2004): "Coordination Failure in Repeated Games with Almost-Public Monitoring," Cowles Foundation, Working Paper No. 1479, Yale University.

Matsushima, H. (2004): "Repeated Games with Private Monitoring: Two Players," *Econometrica*, 72, 823-852.

Piccione, M. (2002): "The Repeated Prisoner's Dilemma with Imperfect Private Monitoring," *Journal of Economic Theory*, 102, 70–83.

Radner R. (1986): "Repeated Partnership Games with Imperfect Monitoring and No Discounting," *Review of Economic Studies*, 53, 43–58.

Sekiguchi, T. (1997): "Efficiency in Repeated Prisoner's Dilemma with Private Monitoring," *Journal of Economic Theory*, 76(2), 345–361.

Yamamoto, Y. (2004): "Efficiency Results in N Player Games with Imperfect Private Monitoring," Mimeo.

Appendix:

Proof of Lemma 4: (**a**) Let $J_{-(i+1)}$ denote the information revealed by all players other than player $i+1$ in Phase 5 (except the first period of Phase 5, information revealed in that period will be irrelevant for part (**a**)). That is, (i) the intentions of players other than $i+1$ revealed in the first period of Phase 5; (ii) the signals about the message profile M sent in Phase 1; (iii) the announcements whether the observed action profile has been $a^{\widetilde{M}}$ in every period of Phase 4; (iv) the announcements whether the first deviation from action profile $a^{\widetilde{M}}$ was unilateral, (v) the announcements who was first deviated from $a^{\widetilde{M}}$ and in which period this occurred. Note that $J_{-(i+1)}$ reveals all mixed actions of all players other than player $i+1$ in Phase 4; that is, if one knows the part of $J_{-(i+1)}$ revealed by player j, then one also knows (for every period t of Phase 4) the (mixed) action taken by player j in period t. Let $I_{-(i+1)}$ denote the information contained in player i's signals from Phase 5. [Under perfect monitoring, $I_{-(i+1)}$ would coincide with $J_{-(i+1)}$.]

The transfer we define have the form:

$$\pi^B_{i+1}(h_i^T) = \frac{1}{\delta^T}\left[\sum_{t=1}^{T}\delta^{t-1}\theta(h_i^t, a^t, I_{-(i+1)})\right],$$

for some function θ to be defined, where h_i^t, a^t denote the truncation of h_i^T to $h_i^{t+1} = (h_i^t, a^t)$. The values of θ are defined by backward induction with respect to t.

Begin with the periods of Phase 5. For those periods, θ depends neither on h_i^t, nor on $I_{-(i+1)}$, but only on a^t. Assume that the values $\theta(a^{\widetilde{t}})$, all $\widetilde{t} > t$, make player $i+1$ indifferent across all sequences of action profiles $a^{\widetilde{t}}$, $\widetilde{t} > t$, and pick as $\theta(a^t)$ the values that make player $i+1$ indifferent over all action profiles of the stage game in period t. Those indifference conditions impose a system of linear equations on the values $\theta(a^t)$, which satisfies the necessary rank condition because monitoring is almost perfect. Moreover, these values $\theta(a^t)$ may be chosen to be positive, and, as the noise level tends to zero, they may be chosen to be bounded by any number larger than $\max_{a\in A} u_{i+1}(a) - \min_{a\in A} u_{i+1}(a)$. By construction, these values $\theta(a^{\widetilde{t}})$, $\widetilde{t} \geq t$, make player $i+1$ indifferent across all sequences of action profiles $a^{\widetilde{t}}$, $\widetilde{t} \geq t$. Thus, player $i+1$ is indifferent over all strategies in Phase 5. The values of θ assigned in Phase 5 will have a small affect on the auxiliary-scenario average payoff of player $i+1$ provided T is sufficiently large.

By the same argument, we may pick the values $\theta(h_i^t, a^t, I_{-(i+1)})$ for all the periods of Phases 2 and 3 such that player $i+1$ is indifferent over all sequences of action profiles in the two phases when he disregards the stage-game payoffs in the periods of Phases 4 and 5 and the values of θ assigned in the two phases (i.e. the values $\theta(h_i^t, a^t, I_{-(i+1)})$, where t is a period of Phase 4 or 5). However, the stage-game payoffs of Phases 4 and 5 and the values of θ assigned in those phases can be disregarded because, by construction, the stage-game payoffs of Phases 4 and 5 are independent of the outcomes in Phases 2 and 3; in addition, the values $\theta(h_i^t, a^t, I_{-(i+1)})$ have already been defined for all periods in Phase 5 , and we shall define the values $\theta(h_i^t, a^t, I_{-(i+1)})$ for all the periods of Phase 4 independently of the outcomes of Phases 2 and 3.

It therefore remains to define the values $\theta(h_i^t, a^t, I_{-(i+1)})$ for the periods of Phases 1 and 4 independently of the outcomes in Phases 2 and 3, such that player $i+1$ is indifferent over all strategies, and such that (10) is satisfied, when we disregard the stage-game payoffs in the periods of Phases 2, 3 and 5 and the values of θ assigned in these three phases.

We denote by $J^t_{-(i+1)}$ (respectively $I^t_{-(i+1)}$) the component of $J_{-(i+1)}$ (respectively $I_{-(i+1)}$) that reveals the information of all players other than $i+1$ that pertains to all periods up to t; more precisely, it reveals their signals about the message profile M sent in Phase 1, if a deviation from $a^{\widetilde{M}}$ was observed in some period $\widetilde{t} < t$, and, if so, it also reveals if the first observed deviation from $a^{\widetilde{M}}$ was unilateral, who deviated from $a^{\widetilde{M}}$ first and in which period. We write $J^t_{-(i+1)} \in \mathcal{J}^t_{-(i+1)}$ (respectively $I^t_{-(i+1)} \in \mathcal{I}^t_{-(i+1)}$) when either all players played $a^{\widetilde{M}}$ up to period t or player $i+1$ was the first to unilaterally deviate in some period $\widetilde{t} < t$ (both according to the signals of all players other than $i+1$; in particular, all players other than $i+1$ obtained the same signal about the message profile M sent in Phase 1).

We will define the values $\theta(h_i^t, a^t, I_{-(i+1)})$ such that:

3. Player $i+1$ is indifferent across all his strategies from period t on (until the end of Phase 4) conditional on every $J^t_{-(i+1)}$ (both from $\mathcal{J}^t_{-(i+1)}$ and from the complement of $\mathcal{J}^t_{-(i+1)}$);

4. His payoff from period t on, augmented by the transfers assigned from period t on (until the end of Phase 4), conditional on every $J^t_{-(i+1)} \in \mathcal{J}^t_{-(i+1)}$, converges when $\varepsilon \to 0$ to the maximum of his payoffs over all continuation strategies under perfect monitoring (until the end of Phase 4), conditional on the same $J^t_{-(i+1)}$.

Condition 3 will then guarantee that player $i+1$ is indifferent over all his strategies and condition 4 will guarantee that (10) is satisfied. Remember that we apply backward induction with respect to t. Each $J^t_{-(i+1)}$ determines the actions in period t of players other than $i+1$. This implies that, given $J^t_{-(i+1)}$, both player $i+1$'s stage-game payoff and the distribution over $J^{t+1}_{-(i+1)}$ in period $t+1$ are determined by player $(i+1)$'s action in period t. It suffices to pick $\theta(h_i^t, a^t, I_{-(i+1)})$ such that player $i+1$ is indifferent across all consequences (the stage-game payoff and the induced distribution over all $J^{t+1}_{-(i+1)}$) of all his actions in period t. Indeed, since player $i+1$ is also indifferent over all his strategies from period $t+1$ on conditional on every $J^{t+1}_{-(i+1)}$, he must be indifferent over all his strategies from period $t+1$ on conditional on every distribution over all $J^{t+1}_{-(i+1)}$, in particular the distribution induced by his action in period t.

If the values θ depended directly on $J^t_{-(i+1)}$ (not only on $I^t_{-(i+1)}$) making player $i+1$ indifferent between the consequences of all his actions in period t would be straightforward. For small enough ε, it would be possible just to pick values of θ (that would depend on $J^t_{-(i+1)}$ and a^t_{i+1}) to make the continuation payoffs of player $i+1$ equal across all his actions.

The values $\theta(h_i^t, a^t, I_{-(i+1)})$ do not directly depend on $J^t_{-(i+1)}$; they depend only on $I^t_{-(i+1)}$, an imperfect signal of $J^t_{-(i+1)}$. However, there is a one-to-one correspondence between $J^t_{-(i+1)}$ and $I^t_{-(i+1)}$, and each $J^t_{-(i+1)}$ induces a probability distribution over all $I^t_{-(i+1)}$ with the probability assigned to $I^t_{-(i+1)}$ corresponding to $J^t_{-(i+1)}$ converging to 1 as $\varepsilon \to 0$. Thus, the matrix D obtained by stacking these probability distributions as row vectors converges to the identity matrix as $\varepsilon \to 0$. This means that D is invertible and we can define the vector of values $\theta(h_i^t, a^t, I_{-(i+1)})$ as the vector of values θ when they depended directly on $J^t_{-(i+1)}$ multiplied by D^{-1}.

Again the values $\theta(h_i^t, a^t, I_{-(i+1)})$ depend only on $I^t_{-(i+1)}$ and player $(i+1)$'s own action in period t, but they need not be non-negative even when their counterparts that depended directly on $J^t_{-(i+1)}$ are. However, we can make them non-negative by adding a constant. As D tends to the identity matrix as $\varepsilon \to 0$, we can assume that this constant converges to 0 as $\varepsilon \to 0$.

This yields condition 3. By construction, condition 4 is satisfied for all $J^t_{-(i+1)}$ such that player $i+1$ is the first player who unilaterally deviated in some period $\widetilde{t} < t$, since then all other players minmax player $i+1$ in t and in all following periods of the phase. Suppose therefore that $J^t_{-(i+1)}$ is such that the players played $a^{\widetilde{M}}$ until period t. Then, given any action by player $i+1$, the induced probability distribution assigns a probability that converges to 1 as $\varepsilon \to 0$ to one specific $J^{t+1}_{-(i+1)} \in \mathcal{J}^{t+1}_{-(i+1)}$ (more precisely, it is $J^{t+1}_{-(i+1)}$ such that all players played a^M until period $t+1$ if player $i+1$ takes action $a^{\widetilde{M}}_{i+1}$; and it is $J^{t+1}_{-(i+1)}$ such that player $i+1$ is the first unilaterally deviator from $a^{\widetilde{M}}$ in period t otherwise). It thus follow immediately from the induction hypothesis that player $(i+1)$'s payoff from period t on, augmented by the transfers assigned from period t on, induced by any of his actions in period t, converges (as $\varepsilon \to 0$) to the maximum of his continuation payoffs under perfect monitoring over all continuation strategies conditional on $J^t_{-(i+1)}$.

(**b & c**) Let $J^1_{-(i+1)}$ denote the information revealed by all players other than player $i+1$ in the first period of Phase 5. This simply means the intentions of players other than $i+1$. Let $I^1_{-(i+1)}$ denote the information conveyed in player i's signals from the first period of Phase 5. The signals obtained by player i in other periods of Phase 5 will be irrelevant for parts (**b & c**).

The transfer will have the form:

$$\pi^G_{i+1}(h_i^T) = \frac{1}{\delta^T}\left[\sum_{t=1}^{T} \delta^{t-1}\theta(h_i^t, a^t, I^1_{-(i+1)})\right],$$

where h_i^t, a^t denote the truncation of h_i^T to $h_i^{t+1} = (h_i^t, a^t)$.

We define the values $\theta(\cdot)$ by backward induction with respect to t. We make player $i+1$ indifferent over all strategies in Phase 5 in the same manner as in (**a**). Therefore consider first a period t in Phase 4. We will specify a system of linear equations such that if the values of θ satisfy our system, then (11), (12) and 1 and 2 in (**c**) hold. Next, we will show that there exists a solution to our system such that all values of θ are non-positive.

Begin with condition 1 from (**c**). Here we impose $\#A$ equations on the values $\theta(h_i^t, a^t, I^1_{-(i+1)})$ for any h_i^t and $J^1_{-(i+1)}$, such that $J^1_{-(i+1)}$ does not coincide with h_i^t on the report of player i from Phase 2. As in the proof of (**a**), if the values of θ depended directly on $J^1_{-(i+1)}$ (write then $\theta(h_i^t, a^t, J^1_{-(i+1)})$ instead of $\theta(h_i^t, a^t, I^1_{-(i+1)})$) our equations would be simple and they would be the same for all pairs h_i^t and $J^1_{-(i+1)}$ with the required property. Namely, assume that the values $\theta(h_i^{\widetilde{t}}, a^{\widetilde{t}}, J^1_{-(i+1)})$, $\widetilde{t} > t$, make player $i+1$ indifferent over all sequences of action profiles $a^{\widetilde{t}}$, $\widetilde{t} > t$, when $J^1_{-(i+1)}$ does not coincide with h_i^t on the report of player i from Phase 2. Then impose the equations that make player $i+1$ indifferent over the stage-game payoffs of all action profiles in period t. This yields a set of $\#A - 1$ equations. Impose also an additional equation that one of the values $\theta(h_i^t, a^t, J^1_{-(i+1)})$ is equal to a negative number

$$c < -[\max_{a \in A} u_{i+1}(a) - \min_{a \in A} u_{i+1}(a)]. \tag{13}$$

Our system consists therefore of $\#A$ equations (for every pair h_i^t and $J^1_{-(i+1)}$ with the required property). All except two coefficients of each of the first $\#A - 1$ equations converge (as $\varepsilon \to 0$) to 0, and the other two coefficients converge to 1. The last equation has one non-zero coefficient, which is equal to 1. The system consists obviously of linearly independent equations, and therefore it has a solution. By (13) and the form of our equations, all values $\theta(h_i^t, a^t, J^1_{-(i+1)})$ of the solution must be negative for small enough ε.

However, the values $\theta(h_i^t, a^t, I^1_{-(i+1)})$ do not depend directly on $J^1_{-(i+1)}$ but only on $I^1_{-(i+1)}$. Let D_1 be the matrix obtained by stacking the probability distributions over $I^1_{-(i+1)}$ conditional on $J^1_{-(i+1)}$ as row vectors. Note that the matrix D_1 tends to the identity matrix as $\varepsilon \to 0$. For any h_i^t and $J^1_{-(i+1)}$ such that $J^1_{-(i+1)}$ does not coincide with h_i^t on the report of player i from Phase 2, impose the set of $\#A$ equations on the values $\theta(h_i^t, a^t, I^1_{-(i+1)})$ which obtains from that when the values $\theta(\cdot)$ depended directly on $J^1_{-(i+1)}$ by replacing $\theta(h_i^t, a^t, J^1_{-(i+1)})$ with D_1 multiplied by the vector of $\theta(h_i^t, a^t, I^1_{-(i+1)})$.

To summarize, for any h_i^t and $J^1_{-(i+1)}$ such that $J^1_{-(i+1)}$ does not coincide with h_i^t on the report of player i from Phase 2, we impose $\#A$ equations. Each of our equations has non-zero coefficients only at the values $\theta(h_i^t, a^t, I^1_{-(i+1)})$ for a given h_i^t. The non-zero coefficients converge to 1 when $I^1_{-(i+1)}$ corresponds to $J^1_{-(i+1)}$ and a^t is either one of two elements of A (for $\#A - 1$ equations) or a distinguished element of A (for one of the equations); otherwise the non-zero coefficients converge to 0. This implies that our system consists so far of linearly independent equations.

When our system is satisfied, every continuation strategy in Phase 4 is a best-response of player $i+1$ conditional on each history of h_i^t of player i and on the event that the message profile reported by player i in Phase 2 will not coincide with the intention profile of players other than $i+1$. This obviously implies that every strategy in Phases 4 is a best-response of player $i+1$ conditional on the latter event.

An analogous argument guarantees condition 2 in (**c**). First, we assume that the values of θ depend directly on the message profile sent by player $i+1$ in Phase 3, and then we replace those values by a matrix D_2 multiplied by the vector of values that depend only on player i's signals about the message profile sent by player $i+1$ in Phase 3, where the matrix D_2 is obtained by stacking the distributions over player i's signals about message profile sent by player $i+1$ in Phase 3 induced by message profiles sent by player $i+1$ in Phase 3.

Summarizing, for any h_i^t, a message profile sent by player $i+1$ in Phase 3 that differs from h_i^t on the part that corresponds to the report of player i from Phase 2, and $J_{-(i+1)}^1$ such that $J_{-(i+1)}^1$ coincides with h_i^t on the report of player i from Phase 2, we impose $\#A$ equations. Each of these equations has non-zero coefficients only at the values $\theta(\widetilde{h}_i^t, a^t, I_{-(i+1)}^1)$, where $\widetilde{h}_i^t$ may differ from a given h_i^t only in Phase 3. The non-zero coefficients converge to 1 when $I_{-(i+1)}^1$ corresponds to $J_{-(i+1)}^1$, $\widetilde{h}_i^t$ is a history that differs from h_i^t but only in Phase 3, and a^t is either one of two elements of A (for $\#A-1$ equations) or a distinguished element of A (for one of the equations); otherwise the non-zero coefficients converge to 0.

Obviously, this system of equations is linearly independent, even combined with the system that guarantees condition 1 from (**c**); indeed, the coefficients of this system converge to 1 whenever $I_{-(i+1)}^1$ corresponds to $J_{-(i+1)}^1$ that coincides with h_i^t on the report of player i from Phase 2 whereas the coefficients of the system that guarantees condition 1 from (**c**) converge to 1 whenever $I_{-(i+1)}^1$ corresponds to $J_{-(i+1)}^1$ that differs from h_i^t on the report of player i from Phase 2.

Finally, (11) for the periods of Phase 4, conditional on the event that the message profile reported by player $i+1$ in Phase 3 coincides with the intention profile revealed in the first period of Phase 5, can be guaranteed in a similar manner to the proof of Lemma 1; (12) can also be guaranteed in a similar manner, if we disregard the stage-game payoffs and transfers assigned in Phases $1-3$.

This requires a system of linearly independent equations with coefficients that tend either to 0 or to 1 as $\varepsilon \to 0$. At least one coefficient of each equation tends to 1, and the only coefficients that may tend to 1 are for values $\theta(h_i^t, a^t, I_{-(i+1)}^1)$ such that the part of h_i^t that corresponds to player i's signal about the message profile sent by player $i+1$ in Phase 3 coincides with player i's report from Phase 2 and $I_{-(i+1)}^1$ coincides with h_i^t on the report of player i from Phase 2. This guarantees that our system combined with the system that guarantees (**c**) consists of linearly independent equations.

It therefore remains to show that player $i+1$ can be made indifferent in Phases 1 and 2, and that he can be made to strictly prefer repeating truthfully in Phase 3 player i's observed report from Phase 2. The latter requirement is easy to achieve, because it suffices to set $\theta(h_i^t, a^t, I_{-(i+1)}^1) = 0$ for every period of Phase 3 where player $i+1$ repeats correctly (according to the signal of player i) the corresponding action of player i's report from Phase 2, and $\theta(h_i^t, a^t, I_{-(i+1)}^1) = c$, satisfying condition (13), otherwise. Indeed, player $i+1$ cannot then benefit from incorrectly repeating due to a higher flow payoff (provided that ε is small enough). On the other hand, the continuation payoff of player $i+1$ at the beginning of Phase 4 is strictly higher when the message profile sent by him in Phase 3 coincides with the message profile reported by player i in Phase 2 (compared to when it does not). The way to make the probability that the two message profiles coincide close to 1 (as $\varepsilon \to 0$) is to repeated truthfully in Phase 3 player i's observed report from Phase 2; this is so because for every history h_{i+1}^t, player $i+1$ assigns a probability that converges to 1 as $\varepsilon \to 0$ to the intersection of the following two events:

(i) the signals received by player $i+1$ in Phase 2 reveal correctly the actions taken by player i;

(ii) the signals received by player i in those periods $\widetilde{t} < t$ of Phase 3 where player $i+1$ repeats the message profile of player i from Phase 2 coincide with the actions by player $i+1$.

Notice that it is essential for (i) and (ii) that player i reports (in Phase 2) each message vector at least with probability ρ, where $\varepsilon/\rho \to 0$ as $\varepsilon \to 0$, as opposed to reporting with probability 1 the message vector that corresponds to the signals he observed in period 1. It should also be emphasized here that this relies on two facts: First, it is essential that in Phase $2-3$ players send messages sequentially. More importantly, we use that signals coincide with action profiles of other players. Suppose for a moment that the messages were not sequentially sent in Phase 3 (a similar argument applies to Phase 2) and erroneous signals are strongly correlated. Then it could happen that player $i+1$, upon receiving signals according to which another player did not repeat correctly in Phase 3 his predecessor's message profile from Phase 2, assigns high probability to the event that player i received an incorrect signal about one of his previous action in Phase 3. Then player $i+1$ could no longer have an incentive to repeat player i's report from Phase 2. Suppose now that the signal set did not coincide with the set of action profiles. Then for some monitoring structures, there could exist a signal of player $i+1$ whose probability is very low for any action profile, but contingent on this signal player i's distribution over his signal is independent of player $i+1$'s action. In such a case, player $i+1$ could again no longer have an incentive to repeat in Phase 3 player i's report from Phase 2.

It is slightly more difficult to ensure that player $i+1$ is indifferent over all actions in Phases 1 and 2. However, player $(i+1)$'s actions in Phase 2 do not affect his payoff conditional on non-erroneous histories, and they can alter the probability of reaching an erroneous history only marginally, that is, with probability converging to 0 as $\varepsilon \to 0$. Thus, player $i+1$ can be made indifferent over all actions in Phase 2 following the same method as in (**a**). Similarly, player $(i+1)$'s actions in Phase 1 (period 1) affects only slightly his total payoff. The key difference compared to Phase 2 is that player $(i+1)$'s action in period 1 also affects the action profile a^M that will be played in Phase 4, but there is only a slight difference in player $(i+1)$'s total payoff across the two different action profiles a^M for any given $M_1, ..., M_i, M_{i+2}, ..., M_n$. *Q.E.D.*