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

Johannes Hörner† Wojciech Olszewski‡

October 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.

*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. We would also like to thank two anonymous referees and David Levine for their comments that led to substantial improvements in the exposition.

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

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

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 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.

[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.

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:

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

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 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 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 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 player

$-i$ uses one of his two strategies, his opponent receives at least v_i whichever of the two strategies he uses himself; if he uses the other strategy, his opponent gets no more than v_i, no matter which strategy he uses in the T-finitely repeated game. In this case, player $-i$ may be called upon to minmax his opponent. There are thus two kinds of "punishments": from one block to the next, player $-i$ may use the strategy that gives his opponent a high or a low payoff; within each block in which he chooses the strategy giving his opponent a low payoff, player $-i$ may need to minmax him.

Therefore, a player is not indifferent over his opponent's choice of strategy in a block. However, by a suitable choice of the probability with which a player uses each strategy within each non-initial block (*the transition probabilities*), as a function of his *recent* history and of his *recent* strategy (i.e. within the previous block), players are indifferent over their two strategies, and weakly prefer them to all others, at the beginning of each block. Further, by suitably choosing the probability with which each strategy is used in the initial block, the payoff vector v is achieved.

This guarantees that beliefs are irrelevant *at the beginning* of each block. Belief-free equilibria obtain for $T = 1$.[2] In this sense, we show that the special features of the prisoner's dilemma ensuring that the folk theorem obtains with $T = 1$ obtain for any stage game, provided one chooses appropriately $T \geq 1$.[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 small enough, he assigns probability almost one to his opponent having observed the same recent history. This is not true for the other recent, *erroneous* histories, as his posterior may then dramatically vary with small differences in the relative likelihood of incorrect signals. As a player's best-reply in

[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).

non-initial periods may depend on his opponent's recent history, specifying best-responses after such histories is less tractable. Worse, a player's belief about his opponent's recent history - and thus his best-response- a priori depends on his belief about the recent strategy used by his opponent, which in turn depends on his own entire private history, rather than his recent one. This means that best-replies may depend on a player's entire history, destroying the recursive structure of our construction.

This problem is solved as follows. For one of his two strategies used in each block, the corresponding transition probabilities are chosen so that a player's opponent is indifferent over *all* strategies within the block, not only over the two strategy he actually chooses from. Thus, to compute his best-reply, a player may always condition on his opponent using the other strategy, independently of the beliefs he actually holds about the recent strategy used by his opponent. As this best-reply depends on what this other strategy specifies, for each possible recent history, as well as on the corresponding transition probabilities, this strategy and the transition probabilities that go along must be determined jointly, 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 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]

[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.

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 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,*

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

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. A player's initial history is the null history $\emptyset$. 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 histories, and $H_i = \cup_t H_i^t$ the set of (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

[5] Theorem 1 does not rule out equilibrium payoffs outside V^* (see footnote 4), but it is possible to show 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 relaxed, or dropped with only two players.

all strategies of player i in the infinitely repeated game is 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 $U_i^A\left(h_{i+1}^T, \pi_i\right) := U_i^T\left(h_{i+1}^T\right) + (1-\delta)\,\delta^T \pi_i\left(h_{i+1}^T\right)$, and its definition extended to strategies $s \in S^T$ in the usual fashion. Continuation payoffs given some private history h_i^t are denoted $U_i(s \mid h_i^t)$, $U_i^T(s \mid h_i^t)$.

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 $B_i(s_{-i}, \pi_i \mid \bar{s}_i)$ 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 $s_i \mid H_i^E = \bar{s}_i \mid H_i^E$.

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

3.1 From belief-free equilibria to block equilibria

As mentioned in the introduction, the folk theorem under private almost-perfect monitoring is already known to hold for the two-player prisoner's dilemma ($G, L \geq 0$):

	C	D
C	$(1,1)$	$(-L, 1+G)$
D	$(1+G, -L)$	$(0,0)$

The key observation that makes the analysis tractable in this case is that each player can ensure, through cooperation, that his opponent gets at least one, and through defection, that he gets at most zero, whether his opponent cooperates or not. More formally, for each payoff vector $v \in (0,1)^2$, there exists a subset of actions $\mathcal{A}_i \subset A_i$, and two elements $\alpha_{-i}^G, \alpha_{-i}^B \in \triangle \mathcal{A}_{-i}$, all i, such that:

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

To see this, pick $\mathcal{A}_i = A_i$ and $\alpha_{-i}^G = C, \alpha_{-i}^B = D$, $i = 1, 2$. The action α_{-i}^G is the "Good" action that secures one to player $-i$, while α_{-i}^B is the "Bad" action that keeps his payoff below zero. Consider now a strategy by player $-i$ that either plays α_{-i}^G or α_{-i}^B. Such a strategy may provide incentives for cooperation if observations pointing towards defection are more likely to trigger play of α_{-i}^B than observations pointing towards cooperation. In fact, by suitably choosing the probability with which player $-i$ sticks to or changes his action as a function of his last signal only, he ensures that player i is indifferent across all his actions in $\mathcal{A}_i$, independently of his private history, provided only that noise and discounting are low enough. In turn, because player i is indifferent across his actions, it is optimal for him to condition his play on his private signal so as to make his opponent indifferent. The payoff v_i is then exactly achieved by specifying appropriately the probability that player $-i$ plays α_{-i}^G in the initial period.

This construction can be generalized. In doing so, one can extend the previous argument to all feasible and individually rational payoffs of the prisoner's dilemma. The set $\mathcal{A}_i$ may be a proper subset of A_i, and it may depend on calendar time (although it cannot depend on the specific history). One obtains thereby the set of all *belief-free equilibria*. Given a belief-free equilibrium $s = (s_1, s_2)$, there exists a sequence of subsets $\{\mathcal{A}_i^t\}_{t=0}^{\infty}$ of A_i such that any strategy s_i' of player i that adheres to this sequence from period t on, that is, for which

$$\forall_{r \geq t}, \forall_{h_i^r} \quad s_i'(h_i^r) \in \mathcal{A}_i^r,$$

is an optimal continuation strategy, independently of player $-i$'s history h_{-i}^t.

In many games, the set of belief-free equilibrium payoffs is larger than the convex hull of the static Nash equilibrium payoffs. But the prisoner's dilemma is exceptional: in most games, this set does not converge to V^* as noise and discounting vanish. If v_i is a belief-free equilibrium payoff close to player i's minmax level v_i^*, then player $-i$ must be able to use an action close to the minmax action α_{-i}^* "most of the time", implying that, in any such period, $\mathcal{A}_{-i}^t$ includes the minmax action's support. This support may include actions yielding player $-i$ low stage-game payoffs -say, below his own minmax payoff-, independently of i's action. This is impossible, as it must be individually rational for player $-i$ to use this action in these periods, independently of his own history.

We claim that the payoff structure described above may be recovered, for any game, and any feasible and individually rational payoff, provided that the stage game is replaced by the normal form of the finitely repeated game. That is, given $v \in V^*$, there exists $\left\{T, \mathcal{S}_i, s_i^G, s_i^B\right\}_{i=1,2}$, with $T \in \mathbb{N}$, $\mathcal{S}_i \subset S_i^T$, and s_i^G, $s_i^B \in \mathcal{S}_i$ such that, for all δ close enough to one,

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

Strategy s_{-i}^G is the "Good" strategy that secures player i at least v_i on average over T periods, provided only player i uses some strategy within $\mathcal{S}_i$. Strategy s_{-i}^B is the "Bad" strategy that keeps player i's average payoff below v_i, independently of i's strategy $s_i \in S_i^T$. In each *block* of the supergame, player $-i$ uses either s_{-i}^G or s_{-i}^B. By suitably choosing the probability with which player $-i$ sticks to or changes his finitely-repeated game strategy from one block to the next, as a function of his observations in the last block only, he ensures that player i is indifferent across all the elements in $\mathcal{S}_i$ at the beginning of each block, independently of his private history, provided only that noise and discounting are low enough. In turn, because player i is indifferent across these elements, it is optimal for him to condition his choice of s_i^G or s_i^B within each block on his observations in the last block. The payoff v_i is then exactly achieved by specifying appropriately the probability that player $-i$ plays s_{-i}^G in the initial block.

Thus, the horizon of the game is divided into T-period blocks. We construct equilibria such that any strategy of player i that adheres within each future block to an element of $\mathcal{S}_i$ is optimal, independently of player $-i$'s history. More precisely, let $s_i^{\prime n} \mid h_i^{nT}$ denote the restriction of $s_i^{\prime} \mid h_i^{nT}$ to the $(n+1)$-st block. Given s_{-i}, any strategy $s_i^{\prime}$ such that

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

for all histories h_i^{mT} following history h_i^{nT}, yields an optimal continuation strategy $s_i \mid h_i^{nT}$, independently of h_{-i}^{nT}.

In the prisoner's dilemma, it is enough to pick $T = 1$. In general, T depends both on the stage game and the payoff vector v. When $T > 1$, a *block equilibrium* need not be belief-free, as a player's set of optimal actions within a block may depend on his private history. However, this dependence is limited to the *recent history* - the finite, terminal segment of the player's private history of those actions taken and signals observed within the current block.

Because block equilibria need not be belief-free, sequential rationality within each block raises difficulties under imperfect private monitoring, affecting the way $\mathcal{S}_i$, s_i^G, s_i^B and T are defined.

3.2 Perfect Monitoring

In this subsection, monitoring is perfect. Therefore, i's signal is $-i$'s actual action, and h_i^t equals h_{-i}^t up to the ordering of signals and actions. Time t refers to the number of periods elapsed in a block, not in the supergame. Thus, h_i^t denotes a recent history, or history for short.

Fix a stage game and a payoff vector $v \in V^*$ throughout.

3.2.1 Payoffs and Actions

Pick four action profiles $\left\{a^{XY}\right\}_{X,Y\in\{G,B\}}$, and corresponding payoff vectors $\left\{w^{XY}\right\}_{X,Y\in\{G,B\}}$:

$$w_i^{XY} = g_i\left(a^{XY}\right),\ i = 1, 2,\ X, Y \in \{G, B\}. \tag{1}$$

The i^{th} superscript (G for "Good", B for "Bad") refers to i^{th} payoff, and indicates whether this payoff is strictly above or below v_i. See Figure 1. Formally:

$$w_i^{GG} > v_i > w_i^{BB},\ i = 1, 2;\ w_1^{GB} > v_1 > w_1^{BG} \text{ and } w_2^{BG} > v_2 > w_2^{GB}.$$

Therefore, there exists $\underline{v}_i$, $\bar{v}_i$, with $v_i^* < \underline{v}_i < v_i < \bar{v}_i$, such that:

$$[\underline{v}_1, \bar{v}_1] \times [\underline{v}_2, \bar{v}_2] \subset \text{ Interior of } \text{co}\left\{w^{GG}, w^{GB}, w^{BG}, w^{BB}\right\}.$$

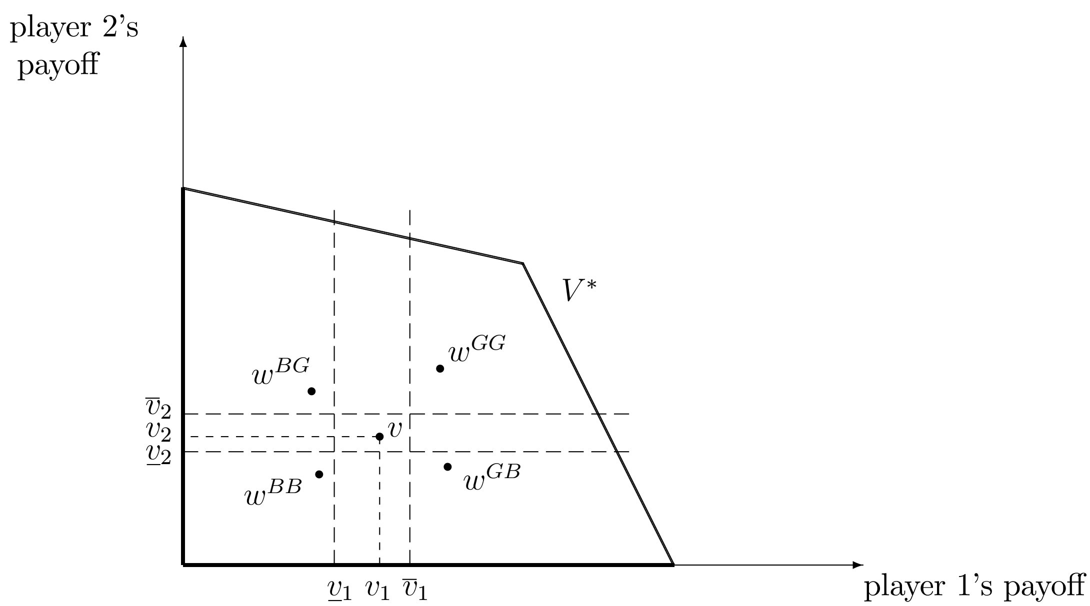

Figure 1: Payoffs

Typically, pure action profiles satisfying the desired inequalities do not exist. However, there always exist an integer m and four finite sequence $\left\{\left(a_1^{XY},...,a_m^{XY}\right)\right\}_{X,Y\in\{B,G\}}$ such that each vector w^{XY}, the average discounted payoff vector over the sequence $\left(a_1^{XY},...,a_m^{XY}\right)$, satisfies the appropriate inequalities, provided δ is close enough to one. The construction that follows must then be modified, by replacing each action profile a^{XY} by the finite sequence of action profiles $\left(a_1^{XY},...,a_m^{XY}\right)$. Details are omitted.

We will show that each payoff in the set $[\underline{v}_1,\bar{v}_1]\times[\underline{v}_2,\bar{v}_2]$ is achieved by some block equilibrium.

3.2.2 The set of strategies $\mathcal{S}_i$

In the first period of a block, actions are used to coordinate continuation play. To this end, partition the set of i's actions into two non-empty subsets, G and B. Player i *sends message* $M_i = G$ if he picks an action in G in the first period; otherwise, player i *sends message* $M_i = B$. As we shall see, players control each other's payoff. That is, i's message refers to $-i$'s average payoff, and the action profile that "corresponds" to the pair of messages (M_1, M_2) is thus $a^{M_2M_1}$. Thus, messages refer to the first period's action choices.

The set of strategies $\mathcal{S}_i$ restricts player i's action only when he observes $M_{-i} = G$; even then, the restriction only applies as long as neither player has deviated from the action profile corresponding to the observed pair of messages (M_1, M_2). In that case, player i must use action $a_i^{M_2M_1}$. Conversely, any strategy in S_i^T that satisfies this restriction is an element of $\mathcal{S}_i$.

Formally, $s_i \in S_i^T$ is in $\mathcal{S}_i \subset S_i^T$ if and only if:

$$(*_2)\ \forall h_i^t = \left(a, \left(a_i^{M_2M_1}, a_{-i}^{M_2M_1}\right), \ldots, \left(a_i^{M_2M_1}, a_{-i}^{M_2M_1}\right)\right),\ a \in M_i \times G,\ t \geq 1\text{:}\ s_i\left[h_i^t\right] = a_i^{M_2M_1}.$$

Observe that the set $\mathcal{S}_i$ imposes no restriction on the action specified by strategy $s_i \in \mathcal{S}_i$ in the initial period, after any history along which player i observed $M_{-i} = B$, and after any history along which some player has deviated from $\left(a_i^{M_2M_1}, a_{-i}^{M_2M_1}\right)$.

In fact, given any history h_i^t, the set $\mathcal{S}_i$ imposes either no restriction on $s_i\left[h_i^t\right]$, or it restricts

$s_i\left[h_i^t\right]$ to a single action. That is, for each history h_i^t, define

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

Then $\mathcal{A}_i\left(h_i^t\right)$, the set of actions *prescribed by* $\mathcal{S}_i$, is either A_i or a singleton $a_i^{M_2M_1}$. With some abuse of notation, we say that $\mathcal{S}_i$ *prescribes action* $a_i^{M_2M_1}$ if $\mathcal{A}_i\left(h_i^t\right) = \left\{a_i^{M_2M_1}\right\}$.

It is useful to consider strategies that assign positive probability to any action that is not ruled out by $(^*{}_2)$. To this end, we define, for $\rho > 0$, the set

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

That is, strategy $s_i \in \mathcal{S}_i$ is in $\mathcal{S}_i^\rho$ if after every history h_i^t, it assigns probability at least ρ to every action in $\mathcal{A}_i\left(h_i^t\right)$. Given $(s_i, s_{-i}) \in \mathcal{S}_i^\rho \times \mathcal{S}_{-i}^\rho$, the set of histories that are on- and off the equilibrium path is independent of the particular choice (s_i, s_{-i}). We define an *erroneous history* h_i^t as any (recent) history that is off the equilibrium path for some (and therefore, every) strategy profile in $\mathcal{S}_1^\rho \times \mathcal{S}_2^\rho$. Otherwise, h_i^t is a *regular history.* Write $H_i^{E,t}$ for the set of all erroneous t-length histories, and let $H_i^{R,t} = H_i^t \backslash H_i^{E,t}$ denote the complement of $H_i^{E,t}$. We define:

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

3.2.3 The strategies s_i^B, $s_i^G \in \mathcal{S}_i$ and T

We define the strategies s_i^B, s_i^G in two steps. First, define s_i^g (for now, on some histories) so that:

$$s_i^g[\emptyset] \in \triangle G \text{ and}$$
$$\forall h_i^t = \left(a, \left(a_i^{M_2M_1}, a_{-i}^{M_2M_1}\right), \ldots, \left(a_i^{M_2M_1}, a_{-i}^{M_2M_1}\right)\right),\ a \in M_i \times M_{-i},\ t \geq 1 : s_i^g\left[h_i^t\right] = a_i^{M_2M_1}.$$

That is, s_i^g sends message G and specifies then the action determined by the pair of messages, as long as player i observes no deviation from this action profile. Similarly, define s_i^b as follows:

$$s_i^b[\emptyset] \in \triangle B \text{ and}$$
$$\forall h_i^t = \left(a, \left(a_i^{M_2M_1}, a_{-i}^{M_2M_1}\right), \ldots, \left(a_i^{M_2M_1}, a_{-i}^{M_2M_1}\right)\right),\ a \in M_i \times M_{-i},\ t \geq 1 : s_i^b\left[h_i^t\right] = a_i^{M_2M_1};$$

moreover, let

$$s_i^b\left[h_i^t\right]=\alpha_{-i}^*$$

for every history h_i^t which is a continuation of a history

$$h_i^r=\left(a,\left(a_i^{M_2M_1},a_{-i}^{M_2M_1}\right),\ldots,\left(a_i^{M_2M_1},a_{-i}^{M_2M_1}\right),\left(a_i^{M_2M_1},a'\right)\right),$$
$$a\in B\times M_{-i},\, a'_{-i}\neq a_{-i}^{M_2M_1},\, t\geq r>1.$$

That is, s_i^b differs from s_i^g in two respects. First, it sends message B rather than G. Second, it specifies minmaxing whenever i's opponent was the first to deviate from the action profile determined by the pair of messages - provided i chose message B.

So far, strategies s_i^b and s_i^g are only defined for some histories. On all other histories, define them arbitrarily. Clearly, s_i^b and s_i^g are in $\mathcal{S}_i$. Observe that, if player $-i$ uses strategy s_{-i}^b, player i gets a stage game payoff strictly below $\underline{v}_i$ in every period but at most two: the initial period, and the one in which he deviates from $\left(a_i^{M_2M_1},a_{-i}^{M_2M_1}\right)$. This is true regardless of the strategy $s_i\in S_i^T$ he may be using. Similarly, if player $-i$ uses strategy s_{-i}^g, player i gets strictly more than $\bar{v}_i$ in every period but the first, for any $s_i\in\mathcal{S}_i$ used by player i.

We may thus pick T such that, for some $\underline{\delta}<1$ and all $\delta\geq\underline{\delta}$, i's average payoff (relative to δ) within the block from any strategy $s_i\in\mathcal{S}_i$ against s_{-i}^g strictly exceeds $\bar{v}_i$, while his average payoff from any strategy $s_i\in S_i^T$ against s_{-i}^b is strictly below $\underline{v}_i$. From now on, assume $\delta\geq\underline{\delta}$.

Pick now some small $\rho>0$ and perturb slightly s_i^g,s_i^b so as to get a pair of strategies s_i^G,s_i^B in $\mathcal{S}_i^\rho$. By choosing ρ small enough, we may assume:

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

Under imperfect monitoring, the small perturbation ρ should be thought of as sufficiently large relative to the monitoring noise.[6]

[6] This perturbation is unnecessary under perfect monitoring. See Remark 2.

3.2.4 The result

Besides stating and proving the result, it is useful to define two further strategies, $r_i^G \in \mathcal{S}_i$ and $r_i^B \in S_i^T$ and functions, or *transfers*, $\pi_i^G : H_{-i}^T \to \mathbb{R}_-$, and $\pi_i^B : H_{-i}^T \to \mathbb{R}_+$.

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 strategies. By enlarging if necessary the rectangle $[\underline{v}_1, \bar{v}_1] \times [\underline{v}_2, \bar{v}_2]$, we may assume without loss of generality that:

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

Given h_{-i}^T, for all $t = 1, \ldots, T$, let θ_t^B denote the difference between i's unnormalized continuation payoff (within a block) from playing r_i^B from period t on, and i's unnormalized continuation payoff from playing i's action - as observed by player $-i$ in period t along h_{-i}^T - followed by reversion to r_i^B from period $t+1$ on. By definition of r_i^B, $\theta_t^B \geq 0$. Set $\pi_i^B \left(h_{-i}^T\right) := \delta^{-T} \sum_{t=1}^T \delta^{t-1} \theta_t^B$.

Given h_{-i}^T, for all $t = 1, \ldots, T$, let θ_t denote the difference between i's unnormalized continuation payoff (within the block) from playing r_i^G from period t on and i's unnormalized continuation payoff from choosing i's action - as observed by player $-i$ in period t along h_{-i}^T - followed by reversion to r_i^G from period $t+1$ on. By definition of r_i^G, $\theta_t \leq 0$ whenever the observed action is in $\mathcal{A}_i\left(h_i^t\right)$, where h_i^t is player i's t-history that corresponds to h_{-i}^T. Define $\theta_t^G := \min\{0, \theta_t\} \leq 0$ and set $\pi_i^G \left(h_{-i}^T\right) := \delta^{-T} \sum_{t=1}^T \delta^{t-1} \theta_t^G$.

Since T is fixed, we may restrict attention to δ close enough to one so that

$$(1-\delta)\, \pi_i^B \left(h_{-i}^T\right) < \bar{v}_i - \underline{v}_i \text{ and } (1-\delta)\, \pi_i^G \left(h_{-i}^T\right) > \underline{v}_i - \bar{v}_i.$$

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: The strategies we specify can be described by automata, which revises states and actions at the beginning of every block (that is, at the beginning of periods 1, $T+1$, $2T+1, \ldots$). An action of the automaton is the finitely repeated game strategy to be used by the player in the block. The automaton can be described as follows. For each $i = 1, 2$:

State space: The state u of player $-i$'s automaton is an element of $[\underline{v}_i, \bar{v}_i]$, player's i continuation payoff in the repeated game.

Initial state: Player $-i$ starts in state $u = v_i$, the payoff to be achieved.

Actions: In state u, player $-i$ picks strategy s^G_{-i} with probability q and strategy s^B_{-i} otherwise, where $q \in [0,1]$ solves $u = q\bar{v}_i + (1-q)\underline{v}_i$. Thus, in each block, player $-i$ performs an initial randomization, and then sticks to the resulting strategy s^G_{-i} or s^B_{-i} throughout the block.

Transitions: If the action of the automaton is s^B_{-i}, and player $-i$'s history (in the block) is h^T_{-i}, then, at the end of the block, player $-i$ transits to state:

$$\underline{v}_i + (1-\delta)\,\pi^B_i\left(h^T_{-i}\right), \tag{2}$$

which is in $[\underline{v}_i, \bar{v}_i]$.

If the state of the automaton is s^G_{-i}, and player $-i$'s history (in the block) is h^T_{-i}, then, at the end of the block, player $-i$ transits to state:

$$\bar{v}_i + (1-\delta)\,\pi^G_i\left(h^T_{-i}\right), \tag{3}$$

which is in $[\underline{v}_i, \bar{v}_i]$.

We claim that these strategies form a subgame-perfect equilibrium. It follows from equations (7)-(8) and the one-shot deviation property that, given $-i$'s strategy, any strategy s_i such that its restriction to any given block is in $\mathcal{S}_i$ is a best-reply. Player i's payoff is equal to the weighted average of the payoff of playing r^G_i against s^G_{-i} and the payoff of playing r^B_i against s^B_{-i}, with respective weights q and $1-q$. Both the average payoff within the block and the continuation payoff from playing r^G_i against s^G_{-i} are equal to $\bar{v}_i$, and both the average payoff within the block and the continuation payoff from playing r^B_i against s^B_{-i} 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 $q\bar{v}_i + (1-q)\underline{v}_i = u$. *Q.E.D.*

We conclude this section with a few remarks pointing out those features that play an essential role under imperfect monitoring.

Remark 1: If player $-i$ uses s_{-i}^B, and computes therefore transfers according to π_i^B, then player i is indifferent across *all* strategies in S_i^T. This implies that, for the sake of computing best-replies, player i can always condition on player $-i$ using s_{-i}^G (and π_i^G).

Remark 2: Since $\mathcal{S}_{-i}$ imposes no restriction on $-i$'s action in the first period and in any ensuing period after message B, and since both s_{-i}^G and s_{-i}^B are in $\mathcal{S}_{-i}^\rho$, the probability that player i's recent history $h_i^t \in H_i^T$ is regular provided he uses s_i^B is (almost) one, even if player $-i$ deviates to any strategy $s_{-i} \in S_{-i}^T$. This would not be true if s_{-i}^G or s_{-i}^B were merely required to be in $\mathcal{S}_{-i}$. This is why it is possible to specify explicitly s_i^B after (some) regular histories only, yet guarantee that $-i$'s payoff against s_i^B does not exceed $\underline{v}_{-i}$ independently of i's strategy. This feature is especially important under imperfect monitoring, as it is then not possible to specify play after erroneous histories. This is the motivation for considering $\mathcal{S}_i^\rho$ and $s_i^G, s_i^B \in \mathcal{S}_i^\rho$.

Remark 3: Recall that s_i^G and s_i^B are defined arbitrarily after some histories, including all erroneous ones. While the specification on erroneous histories affects r_i^G and r_i^B as well as π_i^G and π_i^B, it is irrelevant for the argument: the transfers π_i^B compensate player i for any course of action that is suboptimal against s_{-i}^B in the block, while transfers π_i^G penalize player i for any course of action that improves upon r_i^G against s_{-i}^G. We could instead start by setting $\pi_i^G\left(h_{-i}^T\right) \leq 0$ arbitrarily on all erroneous histories h_{-i}^T; define then $\left(s_1^G, s_2^G\right)$ on erroneous histories $H_1^E = H_2^E$ as any subgame-perfect equilibrium of the game restricted to erroneous histories[7], in which transfers π_i^G are added to the payoffs of the finitely-repeated game; set $s_i^B = s_i^G$ on any such history; given s_{-i}^B, let $\pi_i^B \geq 0$ be the transfers that make player i indifferent across all strategies in S_i^T against $\left(s_{-i}^B, \pi_i^B\right)$. In the first case, play on erroneous histories is given and suitable transfers are derived. In the second, transfers are given and play is derived. Neither approach generalizes to imperfect private monitoring. The transfers for which the strategies in $\mathcal{S}_i$ are optimal after regular histories

[7] One can think about the restricted game as a game in which actions on regular histories are given, and players choose their actions only on erroneous histories.

then depend on strategies on erroneous histories, and conversely. Therefore, we define strategies on such histories and transfers jointly, by applying Kakutani's fixed-point theorem.

3.3 Imperfect Private Monitoring

In this subsection, we prove the two-player folk theorem for almost-perfect monitoring, constructing block equilibria similar to those from the perfect-monitoring case. The generalization to imperfect monitoring must overcome a significant difficulty.

Note first that, under the canonical signal structure, the definitions of $\mathcal{S}_i$ and $\mathcal{S}_i^\rho$ are still valid, as the domain of private histories is the same under perfect and imperfect monitoring. So are the definitions of regular and erroneous (recent) histories.[8] When the monitoring is almost perfect, conditional on a regular history h_i^t, player i is almost sure that his opponent has observed the *corresponding* history, along which $-i$'s signals coincide with i's actions along h_i^t, and $-i$'s actions coincide with i's signals along h_i^t. For such histories h_i^t, it is relatively straightforward to give player i incentives to play the prescribed action conditional on h_i^t. Loosely speaking, this can (almost) be done by penalizing or rewarding player i after a block when $-i$'s signal in period t, along the history corresponding to h_i^t, differs from, or coincides with the prescribed action.

On the other hand, conditional on an erroneous history, player i can be sure that at least one player, in at least one period, has observed an incorrect, or *erroneous* signal in the block. This may or may not be player i. For instance, it could be that $-i$'s signal in the previous period was erroneous, leading him to pick an action consistent with his own strategy given that signal, but different from the one he should have chosen if his signal had been correct. Which scenario is more likely depends on the fine details of the monitoring structure $\{m(\cdot \mid a) : a \in A\}$: merely assuming that the noise level is small hardly restricts i's beliefs over histories h_{-i}^{t-1} conditional on erroneous histories. It is then difficult to prescribe actions through penalties or rewards.

[8] Observe, however, that given a strategy profile $(s_1, s_2) \in \mathcal{S}_1^\rho \times \mathcal{S}_2^\rho$, erroneous histories need no longer be off the equilibrium path.

Yet it is essential that there exists some best-reply after erroneous histories h_i^t that is independent of $s_{-i} \in \left\{s_{-i}^G, s_{-i}^B\right\}$, as otherwise, i's best-reply after h_i^t would depend on his belief about $-i$'s current strategy, and thus, on his own history in the entire supergame, destroying thereby the recursive structure of block equilibria. To handle this difficulty, Lemma 1 proves that, if player $-i$ uses s_{-i}^B (more precisely, plays s_{-i}^B on regular histories), transfers π_i^B (and so transition probabilities in the supergame) can be defined so that player i is indifferent across all actions, conditional on any history, yet holding the payoff of player i below the target level $\underline{v}_i$. This guarantees that any best-reply after h_i^t conditional on s_{-i}^G is also a best-reply conditional on s_{-i}^B. Thus, player i may always assume, for the sake of computing best-replies, that his opponent is playing s_{-i}^G, independently of his own history.

Recall the definition of the auxiliary scenario, introduced in Section 2. Given π_i and h_{-i}^T, player i's payoff in the auxiliary scenario is defined as $U_i^A\left(h_{-i}^T, \pi_i\right) := U_i^T\left(h_{-i}^T\right) + (1-\delta)\,\delta^T \pi_i\left(h_{-i}^T\right)$. This definition is extended to $s \in S^T$ in the usual fashion, and denoted $U^A(s, \pi_i)$. Given (s_{-i}, π_i), the set of i's best-responses in the auxiliary scenario is $B_i(s_{-i}, \pi_i)$.

In the statement of the following lemmata, $U_i^T(s)$ 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.

Lemma 1 *For every strategy $\overline{s} \mid H^E$, there exists $\overline{\varepsilon} > 0$ such that for all $\varepsilon < \overline{\varepsilon}$:*

There exist a non-negative transfer $\pi_i^B : H_{-i}^T \to \mathbf{R}_+$ such that

$$S_i^T = B_i(\overline{s}_{-i}^B, \pi_i^B), \tag{4}$$

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

$$\lim_{\varepsilon \to 0} U_i^A(s_i, \overline{s}_{-i}^B, \pi_i^B) = \max_{\widetilde{s}_i \in S_i^T} U_i^T(\widetilde{s}_i, \overline{s}_{-i}^B). \tag{5}$$

Proof: Given a history h_{-i}^{T}, let $\left(h_{-i}^{t}, a_i\right)$ denote the truncation of h_{-i}^{T} to h_{-i}^{t} and the signal obtained by player $-i$ in period t. The transfer will have the form:

$$\pi_i^B(h_{-i}^T) = \frac{1}{\delta^T}\left[\sum_{t=1}^{T} \delta^{t-1}\theta(h_{-i}^{t-1}, a_i)\right],$$

for some function $\theta(\cdot,\cdot)$ to be specified.

We will achieve (4) defining the values, or transfers $\theta(\cdot,\cdot)$, by backward induction. First, for every history h_{-i}^{T-1}, we pick transfers $\theta(h_{-i}^{T-1}, a_i)$ that make player i indifferent between all his actions conditional on the event that h_{-i}^{T-1} is the history of player $-i$ at the beginning of period T. To see that it is possible, consider, for every action $a_i \in A_i$, as a row vector the probabilities $m_{-i}(\cdot \mid \alpha)$ of signals $\sigma_{-i} \in \Sigma_{-i} = A_i$ observed by player $-i$ when he plays $\alpha_{-i} = \overline{s}_{-i}^B[h_{-i}^{T-1}]$ and player i plays a_i (i.e. α_i assigns probability 1 to a_i). Construct a matrix D^{T-1} by stacking the row vectors for every action $a_i \in A_i$. When monitoring is almost perfect, this matrix is invertible (it is actually almost an identity matrix if, for every a_i, the row corresponding to the action a_i has the same number as the column corresponding to the signal a_i). Therefore the system of equations with the coefficient matrix D^{T-1}, the column vector of unknowns $\theta(h_{-i}^{T-1}, a_i)$ and the right-hand side vector of $g_i(\overline{s}_{-i}^B[h_{-i}^{T-1}], a_i^*) - g_i(\overline{s}_{-i}^B[h_{-i}^{T-1}], a_i)$, where a_i^* is a stage-game best response to $\overline{s}_{-i}^B[h_{-i}^{T-1}]$, must have a solution.

Suppose now that all transfers $\theta(h_{-i}^{\tau}, a_i)$ for $\tau > t$ are already defined, so that player i is indifferent across all his strategies from period $t+1$ on (conditional on each history h_{-i}^{t+1} of player $-i$). Then, for every history $h_{-i}^{t-1} \in H_{-i}^{t-1}$, the i's auxiliary-scenario continuation payoff, defined as the sum of the repeated-game continuation payoff and the expected value of the transfer

$$\frac{1}{\delta^T}\left[\sum_{s=t}^{T} \delta^{s-1}\theta(h_{-i}^{s-1}, a_i)\right],$$

conditional of the event that h_{-i}^{t-1} is $-i$'s history at the beginning of period t, depends only on i's action in period t (i.e. it does not depend on his strategy from period $t+1$ on). For every action $a_i \in A_i$, consider as a row vector the probabilities $m_{-i}(\cdot \mid \alpha)$ of signals $\sigma_{-i} \in \Sigma_{-i} = A_i$ observed by player $-i$ when he plays $\alpha_{-i} = \bar{s}^B_{-i}[h^{t-1}_{-i}]$ and player i plays a_i. Construct a matrix D^{t-1} by stacking the row vectors for every action $a_i \in A_i$. Define transfers $\theta(h^{t-1}_{-i}, a_i)$ as the unique solution of the system of equations with the coefficient matrix D^{t-1}, and the right-hand side vector consisting of the differences between the auxiliary-scenario continuation payoff of player i when he plays an action a^*_i maximizing this continuation payoff and the auxiliary-scenario continuation payoff when he plays action a_i.

In this way, (4) is achieved, as $h^0_1 = h^0_2 = \emptyset$, but two difficulties remain. First, transfers must be non-negative. To this end, notice that, as monitoring becomes perfect, the matrices D^{t-1} tend to the identity matrix, and the right-hand side vectors in the systems of equations determining transfers $\theta(h^t_{-i}, a_i)$ are non-negative. Therefore, the limits of transfers $\theta(h^t_{-i}, a_i)$, as noise vanishes, must be non-negative. Thus, we can make all transfers $\theta(h^t_{-i}, a_i)$ non-negative by adding to all of them a positive constant, and we can assume that this constant tends to 0 as noise vanishes.

Second, (5) must be achieved. To this end, consider the strategy r^B_i of player i such that $(\forall h^{t-1}_i \in H^{t-1}_i)$ $r^B_i[h^{t-1}_i]$ is his stage-game best reply to $\bar{s}^B_{-i}[h^{t-1}_{-i}]$, where h^{t-1}_{-i} denotes the history of player $-i$ corresponding to the history h^{t-1}_i. [Note that this strategy r^B_i typically differs from the strategy denoted by the same symbol in Section 3.2. The two strategies play, however, the same role, and for this reason, we denote them by the same symbol.] As monitoring becomes perfect: (a) by construction, the expected value of the transfer $\pi^B_i(h^T_{-i})$ tends to 0, if players play r^B_i and $\bar{s}^B_{-i}$ respectively, and (b) by the definition of r^B_i, the repeated game payoff of player i tends to the right-hand side of (5); obviously, (a) and (b) yield (5). *Q.E.D.*

The second lemma asserts that, for any fixed pair of strategies on erroneous histories, if player $-i$ plays s^G_{-i} on his regular histories, then he can pick transfers such that any strategy of player i that sticks to some strategy in $\mathcal{S}_i$ on his regular histories is optimal. Given $H^E_i \subset H^T_i$, $(\bar{s}_i, s_{-i}) \in S^T_i \times S^T_{-i}$ and π_i, $B_i(s_{-i}, \pi_i \mid \bar{s}_i)$ denotes the set of strategies maximizing i's auxiliary-scenario payoff against (s_{-i}, π_i) among all strategies $s_i \in S^T_i$ such that $s_i \mid H^E_i = \bar{s}_i \mid H^E_i$.

Lemma 2 *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*

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

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

Proof: The transfer will again have the form:

$$\pi_i^G(h_{-i}^T) = \frac{1}{\delta^T}\left[\sum_{t=1}^{T} \delta^{t-1}\theta(h_{-i}^{t-1}, a_i)\right],$$

Pick $r_i^G \in S_i^T$ to be a strategy that satisfies:

(a) $r_i^G \mid H_i^R = s_i \mid H_i^R$ for some $s_i \in \mathcal{S}_i$;

(b) $r_i^G \mid H_i^E = \overline{s}_i \mid H_i^E$;

(c) $\forall h_i^{t-1} \in H_i^{R,t-1}$, $r_i^G | h_i^{t-1}$ yields the lowest payoff against $\overline{s}_{-i}^G$, in the T-period repeated game, among all strategies with properties (a) and (b).

Without loss of generality, assume that $r_i^G \mid H_i^R$ is a pure strategy. [Note again that this strategy r_i^G formally differs from the strategy denoted by the same symbol in Section 3.2, but the two strategies play the same role, and for this reason, we denote them by the same symbol.]

We now define $\theta(h_{-i}^{t-1}, a_i)$, $\forall h_{-i}^{t-1} \in H_{-i}^{t-1}$, $\forall a_i \in \Sigma_{-i} = A_i$ by backward induction with respect to t. To satisfy (3), it suffices to pick non-positive values for $\theta(h_{-i}^{t-1}, a_i)$ such that (given $\theta(h_{-i}^{\tau-1}, a_i)$, $\tau > t$), the following constraints, or properties, are satisfied:

1. For every history $h_i^{t-1} \in H_i^{R,t-1}$ such that $\mathcal{S}_i$ prescribes A_i, player i is indifferent under the auxiliary scenario between playing all actions $a_i \in A_i$, each followed by switching to r_i^G from period $t+1$ on;

2. For every history $h_i^{t-1} \in H_i^{R,t-1}$ such that $\mathcal{S}_i$ prescribes $a_i = a_i^{GG}$ or a_i^{BG} if $i = 2$, or a_i^{GB} if $i = 1$, the payoff of player i under the auxiliary scenario to playing a_i (weakly) exceeds the payoff to playing any other action, both followed by switching to r_i^G from period $t+1$ on.

[Note that the auxiliary-scenario payoff difference across continuation strategies $s_i \mid h_i^{t-1}$ of player i is independent of $\theta(h_{-i}^{\widetilde{t}-1}, a_i)$ 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 $\theta(h_{-i}^{\widetilde{t}-1}, a_i)$ for $\widetilde{t} > t$, 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 $h_i^{t-1} \in H_i^{R,t-1}$, player i assigns, conditional on observing h_i^{t-1}, probability at least $1 - \nu$ to the event that player 1 observed the corresponding history $h_{-i}^{t-1} \in H_{-i}^{R,t-1}$, along which the signals of player $-i$ coincide with the actions of player i in h_i^{t-1} and the actions of player $-i$ along h_{-i}^{t-1} coincide with the signals of player i in h_i^{t-1}. Note that, given two distinct histories $h_i^{t-1}, h_i^{\prime t-1} \in H_i^{R,t-1}$, the corresponding histories h_{-i}^{t-1} and $h_{-i}^{\prime t-1}$ are distinct. Given any history $h_i^{t-1} \in H_i^{R,t-1}$ and any action $a_i \in A_i$, consider as a row vector the probabilities assigned by player i, conditional on history h_i^{t-1} and on action a_i taken by player i in period t, to the different histories $h_{-i}^{t-1} \in H_{-i}^{t-1}$ and signals $a_i \in \Sigma_{-i} = A_i$ observed by player $-i$ in period t. Construct a matrix D^{t-1} by stacking the row vectors for all regular histories $h_i^{t-1} \in H_i^{R,t-1}$ and actions $a_i \in A_i$.

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

The third lemma shows that, provided the noise level is sufficiently small, the applied transfer π_i^G (from Lemma 2) is arbitrarily close to zero, given $\overline{s}_{-i}^G$, if player i plays r_i^G.

Lemma 3 *In Lemma 2, the non-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)*

$$\lim_{\varepsilon \to 0} U_i^A(s_i, \overline{s}_{-i}^G, \pi_i^G) = \min_{\widetilde{s}_i \in \mathcal{S}_i} U_i^T(\widetilde{s}_i, \overline{s}_{-i}^G); \tag{6}$$

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

Proof: To guarantee (6), the transfers $\theta(\cdot, \cdot)$ from the proof of Lemma 1 must be further specified. In the backward induction with respect to t, we can assume, in addition to properties 1-2, that $\theta(h_{-i}^{t-1}, a_i) = 0$ whenever $h_{-i}^{t-1} \in H_{-i}^{t-1}$ is the history corresponding to some history $h_i^{t-1} \in H_i^{R,t-1}$ and $a_i = r_i^G(h_i^{t-1})$ or h_{-i}^{t-1} is a history not corresponding to any $h_i^{t-1} \in H_i^{R,t-1}$ and $a_i \in \Sigma_{-i} = A_i$.

Indeed, remember that the number of constraints imposed on $\theta(h_{-i}^{t-1}, a_i)$ by properties 1-2 falls below the rank of D^{t-1} by k). Since $\theta(h_{-i}^{t-1}, a_i) = 0$ for every h_{-i}^{t-1} corresponding to $h_i^{t-1} \in H_i^{R,t-1}$ and $a_i = r_i^G(h_i^{t-1})$, the expected value of the transfer $\pi_i^G(h_{-i}^T)$ if he uses the strategy r_i^G tends to 0 as $\varepsilon \to 0$. This yields (i), except that the values $\theta(h_{-i}^{t-1}, a_i)$ may be positive.

We shall now show that the transfers can be picked in such a way that, for any $a_i \in A_i$ and any history h_{-i}^{t-1} corresponding to some $h_i^{t-1} \in H_i^{R,t-1}$, $\theta(h_{-i}^{t-1}, a_i)$ tends to a non-positive value as $\varepsilon \to 0$. This will guarantee that all values $\theta(h_{-i}^{t-1}, a_i)$ 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 $\varepsilon \to 0$.

This follows from property 1 for every history $h_{-i}^{t-1} \in H_i^{R,t-1}$ such that $\mathcal{S}_i$ prescribes A_i on the history h_{-i}^{t-1} corresponding to $h_i^{t-1} \in H_i^{R,t-1}$. Indeed, the value $\theta(h_{-i}^{t-1}, a_i)$ is in the limit equal to the difference between the expected value of the continuation transfer (from period t on), i.e.

the expectation of

$$\frac{1}{\delta^T}\left[\sum_{s=t}^{T}\delta^{s-1}\theta(h_{-i}^{s-1},a_i)\right],$$

and the expected value of the continuation transfer (from period $t+1$ on) when player i plays a_i and switches to r_i^G from period $t+1$.

Since $\mathcal{S}_i$ prescribes A_i, any continuation strategy that uses $a_i \neq r_i^G(h_i^{t-1})$ in period t and switches to r_i^G from period $t+1$ on yields, by the definition of r_i^G, at least as high a payoff against $\overline{s}_{-i}^G$ in the continuation of the T-period repeated game as $r_i^G \mid h_i^{t-1}$. Therefore, the expected value of the continuation transfer (from period t on) to player i who plays a_i and switches to r_i^G from period $t+1$ has to tend to a non-positive number, as otherwise property 1 would be violated. On the other hand, the expected value of the continuation transfer (from period $t+1$ on) to player i who uses the continuation strategy $r_i^G \mid h_i^t$ tends, by construction, to 0 for every history $h_i^t \in H_i^{R,t}$, and so $\theta(h_{-i}^{t-1},a_i)$ converges to a non-positive value as $\varepsilon \to 0$.

A similar argument applies to histories $h_{-i}^{t-1} \in H_i^{R,t-1}$ such that $\mathcal{S}_i$ prescribes A_i prescribes $a_i = a_i^{GG}$ or a_i^{BG} if $i=2$, or a_i^{GB} if $i=1$. Indeed, it is straightforward to see that property 2 can without loss of generality be replaced with:

3. For every history as in property 2, if playing a_i yields a continuation 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_i^G from period $t+1$ on, then the auxiliary-scenario continuation payoff of player i to playing the prescribed action is equal to the continuation payoff to playing any such a_i, both followed by switching to r_i^G from period $t+1$ on;

4. For every history as in property 2, if playing a_i yields a continuation payoff in the T-period repeated game *lower* than the payoff to playing the prescribed action, both followed by switching to r_i^G from period $t+1$ on, then the expected value of the continuation transfer to player i to playing the prescribed action is equal to the expected value of the continuation transfer to playing any such a_i, both followed by switching to r_i^G from period $t+1$ on.

[Note that properties 3 and 4 indeed imply property 2.] Now the argument used in the case that $\mathcal{S}_i$ prescribes A_i can be used again, referring to properties 3 and 4 instead of property 1.

To obtain (ii) refer again to backward induction with respect to t. Note first that the system of linear equations: 1, 3, 4, and $\theta(h_{-i}^{t-1}, a_i) = 0$ for every h_{-i}^{t-1} corresponding to some $h_i^{t-1} \in H_i^{R,t-1}$ and $a_i = r_i^G(h_i^{t-1})$ as well as for every h_{-i}^{t-1} not corresponding to any $h_i^{t-1} \in H_i^{R,t-1}$ and any $a_i \in \Sigma_{-i} = A_i$, uniquely determines $\theta(h_{-i}^{t-1}, a_i)$ $\forall h_{-i}^{t-1} \in H_{-i}^{t-1}$, $\forall a_i \in \Sigma_{-i} = A_i$. Obviously, this system of equations depends continuously on $\overline{s}$. Thus, the values $\theta(h_{-i}^{t-1}, a_i)$ depend continuously on $\overline{s}$ as well, and they still do so if we subtract from all of them

$$\max_{h_{-i}^{t-1}, a_i} \theta(h_{-i}^{t-1}, a_i),$$

which is a constant that makes the values $\theta(h_{-i}^{t-1}, a_i)$ non-positive.

Finally, to obtain (iii) consider first the perfect monitoring case. Then it can be assumed that $\theta(h_{-i}^{t-1}, a_i)$ with properties 1, 3, and 4 is at least as small as:

$$-B := -T[\max_a g_i(a) - \min_a g_i(a)] - \Sigma_{s>t}[\max_{h_{-i}^{s-1}, a_i} \theta(h_{-i}^{s-1}, a_i) - \min_{h_{-i}^{s-1}, a_i} \theta(h_{-i}^{s-1}, a_i)].$$

Thus, for $\varepsilon > 0$ small enough, we can pick the values $\theta(h_{-i}^{t-1}, a_i)$ satisfying properties 1, 3, and 4 that all exceed $-2B$. By compactness of the set of strategies $\overline{s}$, it can thus be assumed, again referring to a backward-induction argument, that $\theta(\cdot)$ is bounded away from $-\infty$ by some $\underline{\theta}$ independent of $\overline{s}$. This implies that, for some $\underline{\pi}$ independent of $\overline{s}$, $\pi_i\left(h_{-i}^T\right) > \underline{\pi}$, $\forall_{h_{-i}^T \in H_{-i}^T}$. *Q.E.D.*

Lemma 1 to 3 establish that, fixing arbitrarily s_{-i} on erroneous histories, there always exist suitable transfers: for $s_{-i} = s_{-i}^B$ on regular histories, transfers can be found so that any strategy in S_i^T is a best-reply; for $s_{-i} = s_{-i}^G$ on regular histories, if attention is restricted to strategies in S_i^T that are arbitrarily fixed on erroneous histories, transfers can be found so that strategies that stick to any s_i in $\mathcal{S}_i$ on regular histories are optimal.

Given that the exact incentives specified after regular histories cannot be tampered with, it is tempting to proceed as in Remark 3. That is, one may wish to specify transfers $\left(\pi_1^G, \pi_2^G\right)$ after erroneous histories arbitrarily; then, along the lines of Lemma 2, find transfers after regular histories so that the desired specification of $\left(s_1^G, s_2^G\right)$ is optimal after regular histories; finally, given these transfers, let $\left(s_1^G, s_2^G\right)$ on erroneous histories be any sequential equilibrium of the resulting auxiliary game starting after that history. This approach is no longer applicable: the transfers π_i^G (ensuring that the desired s_i^G is optimal after regular histories) depend on s_{-i}^G after erroneous histories, and conversely. Thus, transfers $\left(\pi_1^G, \pi_2^G\right)$ and equilibrium strategies on erroneous histories must be defined jointly, as a fixed point of the relevant mapping.

Proof of Theorem 1 (2 players): Define first $\overline{s}_1 \mid H_1^E$ and $\overline{s}_2 \mid H_2^E$, and some auxiliary transfers π_i^G and π_i^B, $i = 1, 2$. To this end consider a correspondence 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, defined by

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

where $s_i' \mid H_i^E$ 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}^R 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 2-3, for $\overline{s} \mid H^E = (\overline{s}_1 \mid H_1^E, \overline{s}_2 \mid H_2^E)$.

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 $\underline{\pi}$. 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 3, continuous with respect to $\overline{s} \mid H^E$, F is upper hemi-continuous.

Let $\left(\overline{s}_1 \mid H_1^E, \overline{s}_2 \mid H_2^E, \pi_1^G, \pi_2^G\right) \in F\left(\overline{s}_1 \mid H_1^E, \overline{s}_2 \mid H_2^E, \pi_1^G, \pi_2^G\right)$. Further, let π_i^B, $i = 1, 2$, be defined by Lemma 1, in which the strategy profile on erroneous histories coincides with $\overline{s}_1 \mid H_1^E$ and $\overline{s}_2 \mid H_2^E$, the first two coordinates of the fixed point. Notice that, by construction, playing any strategy s_i such that $s_i \mid H_i^R = \widetilde{s}_i \mid H_i^R$ 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 (in the finitely-repeated game) close to $\bar{v}_i$ and $\underline{v}_i$, respectively, if ε is sufficiently close to 0; slightly perturbing the box $\prod_{i=1}^{2}[\underline{v}_i, \bar{v}_i]$ if necessary, we can assume that the payoffs are exactly equal to $\bar{v}_i$ and $\underline{v}_i$.

We show that the payoff set $\prod_{i=1}^{2}[\underline{v}_i, \bar{v}_i]$ can be achieved in a block equilibrium under almost perfect private monitoring, and the discount factor δ close to 1. In particular, assume that the discount factor δ is close enough to 1 so that $\bar{v}_i + (1-\delta)\pi_i^G\left(h_i^T\right) > \underline{v}_i$ and $\underline{v}_i + (1-\delta)\pi_i^B\left(h_i^T\right) < \bar{v}_i$ for all $h_i^T \in H_i^T$. Similarly to the perfect-monitoring case, the strategies can be described as automata, which revise states and actions at the beginning of every block. For each $i = 1, 2$:

State space: The state u of player $-i$'s automaton is an element of $[\underline{v}_i, \bar{v}_i]$, player's i continuation payoff in the repeated game.

Initial state: Player $-i$ starts in state $u = v_i$, the payoff to be achieved.

Actions: Define first $\overline{s}_{-i}^{G}$ as the strategy that coincides with the perfect-monitoring strategy s_{-i}^{G} on regular histories H_{-i}^{R} and with $\overline{s}_{-i} \mid H_{-i}^{E}$ (defined by the fixed point theorem) on erroneous histories, and $\overline{s}_{-i}^{B}$ as the strategy that coincides with the perfect-monitoring strategy s_{-i}^{B} on regular histories H_{-i}^{R} and with $\overline{s}_{-i} \mid H_{-i}^{E}$ (defined by the fixed point theorem) on erroneous histories. In state $u \in [\underline{v}_i, \bar{v}_i]$, player $-i$ picks the strategy $\overline{s}_{-i}^{G}$ with probability q and the strategy $\overline{s}_{-i}^{B}$ with probability $1-q$, where q solves $q\bar{v}_i + (1-q)\underline{v}_i = u$.

Transitions: If the action of the automaton is $\overline{s}_{-i}^{B}$, and player $-i$'s history (in the block) is h_{-i}^{T}, then, at the end of the block, player $-i$ transits to state:

$$\underline{v}_i + (1-\delta)\,\pi_i^B\left(h_{-i}^T\right), \tag{7}$$

which is in $[\underline{v}_i, \bar{v}_i]$.

If the state of the automaton is $\overline{s}_{-i}^{G}$, and player $-i$'s history (in the block) is h_{-i}^{T}, then, at the end of the block, player $-i$ transits to state:

$$\bar{v}_i + (1-\delta)\,\pi_i^G\left(h_{-i}^T\right), \tag{8}$$

which is in $[\underline{v}_i, \bar{v}_i]$.

It follows from the one-shot deviation property that, given the strategy of player $-i$, any strategy of player i such that, in every block, $s_i \mid H_i^R = \widetilde{s}_i \mid H_i^R$ 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. The payoff of player i equals the weighted average of the payoff to playing a best response against $\overline{s}_{-i}^G$ and the payoff to playing a best response against $\overline{s}_{-i}^B$ with weights q and $1-q$. Thus, the payoff of player i in state u is $q\bar{v}_i + (1-q)\underline{v}_i = u$. *Q.E.D.*

4 More players

One of the convenient features of the proof for two players is that, from one block to the next, each player only keeps track of his opponent's continuation payoff. At the beginning of each block, Player i's continuation strategy is optimal independently of his own continuation payoff, which is controlled by his opponent. This means that, up to any such period, private histories do not play the role of a coordination device.

Since the sets of any two players' opponents are no longer disjoint when there are more two players, coordination issues reappear if, at the end of each block, each player must keep track of all his opponents' continuation payoff simultaneously. In the construction below, this problem is avoided by having each player keep track of one and only one of his opponents' continuation payoff. That is, at the beginning of each block, player i uses one of two continuation strategies, as a function of his recent private history and strategy, that determines whether the continuation payoff of his *successor*, player $i+1$ (identifying player $n+1$ and 1), is high or low. Of course, there are events within each block after which coordination is still necessary, such as all those calling for minmaxing of a particular player.

In the two-player construction, each player may assume that his opponent plays a specific strategy within the block (s_{-i}^G), as he is indifferent across all continuation strategies (within the block) if his opponent does not use that strategy. This feature is essential, as it ensures that best-replies after erroneous histories only depends on a player's recent history. With more than two players, this would require that for each strategy profile of player i's opponents, there is at most one for which the set of his best-replies may depend on his recent history. Yet for half of these 2^{n-1} strategy profiles, player i's continuation payoff is high. This means that for at least $2^{n-2}-1 \geq 1$ strategy profiles of his opponents, his continuation payoff must be high, while he is indifferent across all continuation strategies within the block. As discussed in Section 3.1., this is a demanding requirement on the payoff structure of the game. A notable exception, here again, is the n-prisoner's dilemma, for which such strategy profiles can be constructed. Our solution to this dimensionality problem is described below, but it comes at a cost: the construction does not typically generalize to signal spaces that are richer than the canonical space. (See Section 5, Remark 4.)

Our construction can be roughly and loosely summarized as follows.[9] As discussed, player i uses his private recent history and strategy to decide whether player $i+1$'s continuation payoff is high or low. He announces his decision through a choice of action (a message) in the initial period of the current block. As in the two-player case, these messages allow players to coordinate with high probability on the action profile to be played within the block. However, it is possible that, after some recent histories, player $i+1$ assigns high probability to his opponent having failed to coordinate in this initial period. To guarantee that player $i+1$ may always act as if his opponents successfully coordinated, in the initial period and thereafter, we introduce a second round of messages (through choices of actions) at the end of each block. If coordination among $i+1$'s opponents has indeed failed at some point, player i will learn (with high probability) about it at the end of the block, and will adjust his transfer so that player $i+1$'s choice of continuation strategy within the block is payoff-irrelevant, conditional on coordination having failed. This coordination failure may result because of erroneous signals in the initial period, but also because of "deviations" from the prescribed action profile. Because the specific cause of this failure may matter, the second round of messages must carry a lot of information, and

[9] The actual construction described in the next two sections is slightly more complex.

that requires many periods. Nevertheless, the duration of this communication phase is arbitrarily short, relative to the length of a block.

This does not solve the problem of dimensionality, for we still need to make sure that player $i+1$'s beliefs about which strategy profile his opponents (are supposed to) coordinate upon are irrelevant for the sake of computing best-replies. To make sure that player $i+1$ may always act as if he knew his opponents' strategy profile, even after erroneous histories (defined as before), we add a second round of communication, immediately after the initial round of messages. In this round, player $i+1$ is given incentives to "repeat" what strategy profile he believes his opponents intend to play, given his signals about the initial messages. Conditional on his message being incorrect, transfers are adjusted so as to ensure that player $i+1$ is indifferent over all continuation strategies within the remaining periods of the block. Therefore, player $i+1$ may always act as if players $-i$ coordinate, and coordinate on what he said they were.

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.

4.1 Perfect Monitoring

Fix a stage game and a payoff vector $v \in V^*$.

4.1.1 Payoffs and Actions

Pick 2^n payoff vectors w^M, where $M = (M_1, ..., M_n)$, and $M_i \in \{G, B\}$ such that:

$$w_i^M > v_i \text{ if } M_i = G;\ w_i^M < v_i \text{ if } M_i = B.$$

There exist $\underline{v}_i$ and $\overline{v}_i$ with $\underline{v}_i < v_i < \overline{v}_i$ such that:

$$[\underline{v}_1, \overline{v}_1] \times \cdots \times [\underline{v}_n, \overline{v}_n] \subset \text{Interior of } \operatorname{co}(\{w^M : M \in \{G, B\}^n\}).$$

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 conventions of Section 3. We show that the payoff set $\prod_{i=1}^n [\underline{v}_i, \overline{v}_i]$ can be achieved in block strategies.

4.1.2 The set of strategies $\mathcal{S}_i$

The play in a block will be divided into *five* phases. Phases 2, 3 and 5 play no role under perfect monitoring.

Phase 1 (period 1): Each player i simultaneously sends message G or B. That is, 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, corresponding to message profile $M_{-i} = (M_1, ..., M_{i-1}, M_{i+1}, ..., M_n)$, $M_j \in \{G, B\}$ for 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 $j \neq i$.

Phase 3 (periods $t = n(n-1)+2,...,2n(n-1)+1$): Players "repeat" consecutively their predecessors' reports. That is, player 2 first 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 a message $M_i \in \{G,B\}$. Then player i reveals his signals about the message profile M sent in Phase 1, and whether all players (according to his signals) were playing $a^{\widetilde{M}}$ in every period of Phase 4, where $\widetilde{M} = (M_n, M_1, M_2, \ldots, M_{n-1})$. [$\widetilde{M}$ rather than M is used as superscript, as 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 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. This period can be announced in no more than $1+\log T$ periods: 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 in Section 3.2, 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 conditions:

$(^*{}_{n,1})$ In Phase 3, s_{i+1} repeats truthfully (with probability 1) player i's report from Phase 2 (as inferred from player $i+1$'s private history).

$(^*{}_{n,2})$ 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 every period t of Phase 4 as long as, according to player $i+1$'s private history, players have played $a^{\widetilde{M}}$ in all previous periods $\bar{t} < t$ of Phase 4.

Note that no restriction is imposed on the actions chosen in Phases 1, 2 and 5. To avoid clutter, the conditions $(^*{}_{n,1}) - (^*{}_{n,2})$ are informal in two respects. More precisely, condition $(^*{}_{n,1})$ means that s_{i+1} assigns probability 0 to any action from the set G_{i+1} (respectively, B_{i+1}) in the period, in which he repeats message B (respectively, G) from player i's report. That is, the condition restricts only the support of actions taken 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 conditions for $\mathcal{S}_{i+1}$ remains well-defined under imperfect monitoring. Note also that condition $(^*{}_{n,2})$ is the counterpart of condition $(^*{}_2)$ of the two-player case, in which blocks consist only of Phase 1 and 4.

As with two players, $\mathcal{S}_i$ *prescribes* the set of actions $\mathcal{A}_i(h_i^t)$, given some private history h_i^t, if $\mathcal{A}_i(h_i^t)$ 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 coincides always with G_i or B_i, described in Phase 1, when it is player i who repeats his predecessors' report in a given period, and this prescribed set coincides with A_i otherwise.

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 $a_i \in \mathcal{A}_i(h_i^t)$ assign probability at least ρ to action a_i, i.e.

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

4.1.3 The strategies s_i^B, $s_i^G \in \mathcal{S}_i$ and T

We shall now show that if T is large enough, 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 $\overline{v}_{i+1}$; 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 do not differ in the actions taken in Phases 2, 3, and 5 except the first period of that phase. In Phase 1, s_i^g sends message G and s_i^b sends message B, i.e.

$$s_i^g[\emptyset] \in \Delta G_i \text{ and } s_i^b[\emptyset] \in \Delta B_i$$

(As usual, $s_i^g[\emptyset]$ and $s_i^b[\emptyset]$ are uniform distributions over G_i and B_i, respectively).

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 in Section 4.2. In Phase 4, given the message profile $M = (M_1, M_2, \ldots, M_{n-1}, M_n)$ 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 unilateral deviation from $a^{\widetilde{M}}$. [In particular, both strategies specify $a_i^{\widetilde{M}}$ in case of a simultaneous deviation.] Moreover, if a player $j \neq i$ unilaterally deviates from $a_j^{\widetilde{M}}$, then s_i^b specifies player i's (possibly mixed) action in action profile α_j^*, minmaxing player j. On all other histories h_i^t such that t is a period of Phase 4, define s_i^g and s_i^b arbitrarily. Finally in Phase 5, players first repeat their messages from Phase 1, and they next truthfully communicate what they are supposed to communicate in Phase 5 (see Section 4.2).

For large enough T and δ, the average payoffs of player $i+1$ against s_i^g and s_i^b (and $s_j \in \{s_j^g, s_j^b\}$ for $j \neq i, i+1$), 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 $\overline{v}_{i+1}$.

Perturb slightly strategies s_i^g and s_i^b so that, after any history h_i^t, player i plays each action $a_i \in \mathcal{A}_i\left(h_i^t\right)$ 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 $\overline{v}_{i+1}$, 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 $\underline{v}_{i+1}$.

4.1.4 The result

We denote, for $s_j \in \{s_j^G, s_j^B\}$ for $j \neq i, i+1$, by $r_{i+1}^G(s_j, j \neq i, i+1)$ a 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 ($n \geq 2$, perfect monitoring): Similarly to the two-player case, the strategies can be described by automata, which revise states and actions at the beginning of every block. An action of an automaton is the finitely repeated game strategy in the block.

State space, and initial state: The state u of player i's automaton is an element of $[\underline{v}_{i+1}, \overline{v}_{i+1}]$, player $(i+1)$'s continuation payoff in the repeated game. Player i starts in state $u = v_{i+1}$, the payoff to be achieved.

Actions: At the beginning of each block, in state u, player i performs an initial randomization: for $q \in [0,1]$ such that $u = q\overline{v}_{i+1} + (1-q)\underline{v}_{i+1}$. 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 this initial randomization. For later purposes, we refer to the outcome of this randomization as player i's *intention*.

Transitions: Suppose first that player i intends to play s_i^B. Then he records (at the very end of the block) 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 received from player j 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 $r_{i+1}^B(s_j, j \neq i, i+1)$ from period t on, and player $i+1$'s unnormalized continuation payoff from playing the action actually chosen by player $i+1$, followed by switching to $r_{i+1}^B(s_j, j \neq i, i+1)$ from period $t+1$ on. By definition of $r_{i+1}^B(s_j, j \neq i, i+1)$, $\theta_t^B \geq 0$ for every action of player $i+1$. At the end of the block, player i then transits to the state:

$$\underline{v}_{i+1} + \frac{1-\delta}{\delta^T} \sum_{t=1}^{T} \delta^{t-1} \theta_t^B, \tag{9}$$

which is less than $\overline{v}_{i+1}$ if δ is large enough. Observe that, if player $i+1$ knew that player i was playing s_i^B, every strategy s_{i+1} of player i would be a best-response (given the continuation payoffs), and it would yield a payoff (in the supergame) less than $\underline{v}_{i+1}$. Increasing the transition state by a constant that depends on s_j, $j \neq i, i+1$, we can assume that it would yield exactly the payoff of $\underline{v}_{i+1}$.

Suppose now that player i intends to play s_i^G. 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_i^B.

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 the first period of Phase 5.

Case iii: The message vector reported by player i in Phase 2 coincides with the message vector sent in the first period of 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 $\overline{v}_{i+1} - \zeta \in (\underline{v}_{i+1}, \overline{v}_{i+1})$, and picks then the "actual" transition states as when he intends to play s_i^B, in order to make player $i+1$ indifferent across all strategies $s_{i+1} \in S_{i+1}^T$. As the payoff of player $i+1$ from any strategy in S_{i+1}^T is $\underline{v}_{i+1}$ when player i intends to play s_i^B, player $i+1$'s payoff from any strategy is now $\overline{v}_{i+1} - \zeta$.

In Case iii, player i records the periods in which player $i+1$ departs from $r_{i+1}^G(s_j, j \neq i, i+1)$, where again $s_j = s_j^{M_j}$, $M_j \in \{G, B\}$ stands for the message received from player j in the first period of 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 $r_{i+1}^G(s_j, j \neq i, i+1)$ from period $t+1$ on, and the unnormalized continuation payoff of playing $r_{i+1}^G(s_j, j \neq i, i+1)$ from period t on. Let $\theta_t^G = \max\{0, \theta_t\}$. At the end of the block, player i transits to the state:

$$\overline{v}_{i+1} - \frac{1-\delta}{\delta^T} \sum_{t=1}^{T} \delta^{t-1} \theta_t^G, \tag{10}$$

which is more than $\underline{v}_{i+1}$ if δ is large enough. Observe that, if player $i+1$ knew that player i played s_i^G, every strategy $s_{i+1} \in \mathcal{S}_{i+1}$ would be a best-response (given the continuation payoffs), and yield a payoff (in the supergame) higher than $\overline{v}_{i+1}$. Decreasing the transition state by a constant that depends on s_j, $j \neq i, i+1$, we can assume that it would yield exactly the payoff of $\overline{v}_{i+1}$.

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

Note that neither Phase 2, 3, nor Phase 5 played any role in our construction of block-strategy equilibria under perfect monitoring. In particular, we did not have to consider Cases i and ii.

Under imperfect private monitoring, the information transmitted in Phase 5 allows player i to pick the transition probabilities that make player $i+1$ indifferent across all strategies when he intends to play s_i^B. [Phase 5 plays no role when player i intends to play s_i^G.] 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 now 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 in Phase 4 that player i will know these actions and continuation strategies when he determines the transition probabilities. It will be (approximately) achieved (for $\varepsilon \to 0$) through the information transmitted in Phase 5. It is essential that player i determines the transition probabilities based on the information transmitted in Phase 5, not that acquired in Phases 1 through 4, because, conditional on some erroneous histories h_{i+1}^t (in Phase 4) player $i+1$ may happen to believe that player i received erroneous signals about the actions and continuation strategies of players other than $i+1$.

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 are provided below.

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 *regular*. As in the case of two players, let $H_i^{R,t}$ denote the set of all regular t-length histories, and let $H_i^{E,t} = H_i^t \backslash H_i^{R,t}$ denote the complement of $H_i^{R,t}$. Let

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

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^R 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^R and a given strategy on H_i^E), player i can pick the transition probabilities so that after histories from H_{i+1}^R, 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.

We wish to emphasize here that all words in italics refer to actual messages (sent by taken actions) of a given player, not to signals about these messages observed by any other player. This construction will guarantee that player $(i+1)$'s set of best replies (in Phase 4) 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 an auxiliary scenario in which 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 the 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 $M_i = B$*

$$S_{i+1}^T = B_{i+1}(\overline{s}_{-(i+1)}^M, \pi_{i+1}^B), \tag{11}$$

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

$$\lim_{\varepsilon \to 0} U_{i+1}^A(s_{i+1}, \overline{s}_{-(i+1)}^M, \pi_{i+1}^B) = \max_{\widetilde{s}_{i+1} \in S_{i+1}^T} U_{i+1}^T(\widetilde{s}_{i+1}, \overline{s}_{-(i+1)}^M). \tag{12}$$

(***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 $M_i = G$*

$$\begin{aligned} \{s_{i+1} &\in S_{i+1}^T : s_{i+1} \mid H_{i+1}^R = \widetilde{s}_{i+1} \mid H_{i+1}^R \text{ for some } \widetilde{s}_{i+1} \in \mathcal{S}_{i+1} \text{ and} \\ s_{i+1} &\mid H_{i+1}^E = \overline{s}_{i+1} \mid H_{i+1}^E\} \subset B_{i+1}(\overline{s}_{-(i+1)}^M, \pi_{i+1}^G \mid \overline{s}_{i+1}) , \end{aligned} \tag{13}$$

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

$$\lim_{\varepsilon \to 0} U_{i+1}^A(s_{i+1}, \overline{s}_{-(i+1)}^M, \pi_{i+1}^G) = \min_{\widetilde{s}_{i+1} \in \mathcal{S}_{i+1}} U_{i+1}^T(\widetilde{s}_{i+1}, \overline{s}_{-(i+1)}^M); \tag{14}$$

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

(c) *Moreover, every strategy in Phase* 4 *yields the same payoff to player* $i+1$*, 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 (12), 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 (14), that

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

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 first $\overline{s}_i \mid H_i^E$ for $i = 1, 2, ..., n$, and some auxiliary transfers π_{i+1}^{G} and π_{i+1}^{B}. To this end consider a correspondence 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 defined by

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

where $s'_{i+1} \mid H^E_{i+1}$ is a strategy of player $i+1$ that is a best-response to his opponents' strategy $\left(s^M_{-(i+1)} \mid H^R_{-i}, \overline{s}_{-(i+1)} \mid H^E_{-(i+1)}\right)$ 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 $\overline{s} \mid H^E = (\overline{s}_i \mid H^E_i, i = 1, 2, ..., n)$.

Note that the set of all strategies $\overline{s}_i \mid H^E_i$ 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 $\underline{\pi}$. 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 of F. Let π^B_{i+1} be defined as in Lemma 4, in which the strategy profile on erroneous histories coincides with $\overline{s}_i \mid H^E_i$, $i = 1, 2, ..., n$ (the first coordinates of the 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 $s_{i+1} \mid H^R_{i+1} = \widetilde{s}_{i+1} \mid H^R_{i+1}$ 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 payoffs no higher than $\underline{v}_{i+1}$ and no lower than $\overline{v}_{i+1}$, respectively, if ε is sufficiently close to 0.

We show that the payoff set $\prod_{i=1}^n [\underline{v}_i, \overline{v}_i]$ can be achieved under almost perfect private monitoring, if the discount factor is large enough. Divide the horizon of the infinitely repeated game into T-period blocks. Pick $\bar{\delta}$ close enough to one such that $\overline{v}_{i+1} + (1 - \bar{\delta})\pi^G_{i+1} > \underline{v}_{i+1}$ and $\underline{v}_{i+1} + (1 - \bar{\delta})\pi^B_{i+1} < \overline{v}_{i+1}$ for all histories. Construct a strategy for player i as follows:

State space, and initial state: The state u of player i's automaton is an element of $[\underline{v}_{i+1}, \overline{v}_{i+1}]$, player's $(i+1)$ continuation payoff in the repeated game. Player i starts in state $u = v_{i+1}$, the payoff to be achieved.

Actions: At the beginning of each block, in state u, player i performs an initial randomization: for $q \in [0,1]$ such that $u = q\overline{v}_{i+1} + (1-q)\underline{v}_{i+1}$. 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 this initial randomization.

Transitions: If he plays $\overline{s}_i^G$, then at the end of the block he transits to the state

$$\overline{v}_{i+1} + (1-\delta)\pi_{i+1}^G \in [\underline{v}_{i+1}, \overline{v}_{i+1}];$$

if he plays $\overline{s}_i^B$, then at the end of the block he transits to the state

$$\underline{v}_{i+1} + (1-\delta)\pi_{i+1}^B \in [\underline{v}_{i+1}, \overline{v}_{i+1}].$$

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, $s_{i+1} \mid H_{i+1}^R = \widetilde{s}_{i+1} \mid H_{i+1}^R$ for some $\widetilde{s}_{i+1} \in \mathcal{S}_{i+1}$ and $s_{i+1} \mid H_{i+1}^E = \overline{s}_{i+1} \mid H_{i+1}^E$ is a best-response. The payoff of player $i+1$ is equal to the weighted average $q\overline{v}_{i+1} + (1-q)\underline{v}_{i+1} = u$. *Q.E.D.*

To conclude this section, we wish to briefly point out which of the additional features of our construction in the case of more than two players are important, and which are not. It is essential that players repeat messages in Phase 3 sequentially, and also that all players but $i+1$ randomize across all actions when player $i+1$ repeats the observed report of player i from Phase 2, for the same reason as in Remark 4 below: player $i+1$ must be sure, independently of his own signal, that player i received a signal that is "evidence" of the action he actually took. If players repeated messages simultaneously, or if some of the other players played an action with a probability that were small relative to ε, player $i+1$ would assign high probability to some of his signal being erroneous. Since erroneous signals may be correlated, player $i+1$ could then believe that player i received an erroneous signal about his action. It is convenient but inessential to assume that players make sequential reports in Phases 2 and 5, and that all players but the one reporting uniformly randomize.

5 Extensions and concluding comments

Remark 4 (non-canonical signal spaces): 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 $a_i \in A_i$,

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

Under this definition, the reader can check that the proof of the folk theorem (Theorem 1) remains valid for $n = 2$ for any finite signal space, as it does for infinite Σ_i, but finite A_i. This partition can further depend on a_i.

For $n > 2$, our proof requires an additional assumption on the 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 totally mixed action profile of other players, and for every signal $\sigma_{i+1} \in \Sigma_{i+1}$ observed with positive probability (given the action profile), player $i + 1$ assigns a probability that tends to 1 as ε tends to 0 to signals $\sigma_i \in \bigcup \{\Sigma_i^{a_{-i}} : a_{i+1}$ is the i-th coordinate of $a_{-i}\}$. That is, player $i + 1$ must be sure, independently of his own signal, that player i received a signal that is "evidence" of the action he actually took.

This assumption is needed to give player $i + 1$ an incentive to truthfully repeat in Phase 3 his observed report of player i from Phase 2 (see the proof of Lemma 4 for details). It is trivially satisfied for the canonical signal space. When it is satisfied, Theorem 1 also holds for infinite signal spaces Σ_i. We do not know whether Theorem 1 holds for $n > 2$ without this assumption.[10]

We do not know either of any tractable modification of the proof that would apply to the case in which the action sets A_i themselves are infinite.

Remark 5 (Non-vanishing monitoring imperfections): One may wish 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 conditionally independent. 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 in the round. While this may be possible, there is a serious difficulty. 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 only on his rival's signals. If each period within the round is replaced by a block, this property is not preserved: 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 than in Matsushima (2004), before considering more general stage games.

[10] 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, with perfect monitoring, rather than public monitoring, as the benchmark.

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: Let $J_{-(i+1)}$ denote the information revealed by players other than player $i+1$ in Phase 5. That is, (i) the intentions of players other than $i+1$ revealed in the first period of Phase 5; (ii) the signals (of those players) 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) if not, the announcements whether the first "deviation" from action profile $a^{\widetilde{M}}$ was unilateral, (v) the announcements who (as the first) deviated from $a^{\widetilde{M}}$ and in which period this occurred. It is important to emphasize that by "revealed by players other than player $i+1$" we understand the messages actually sent by those players, not the signals about those actions received by any other player.

Note that $J_{-(i+1)}$ reveals all mixed actions of all players other than player $i+1$ in Phase

4; that is, if one knew the part of $J_{-(i+1)}$ revealed by player j, then one would also know (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 obviously coincide with $J_{-(i+1)}$.]

(**a**) The transfer we define have the form:

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

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

Begin with the periods of Phase 5. For those periods, θ will depend neither on h_i^{t-1}, nor on $I_{-(i+1)}$, but only on a^t. Assume that the values $\theta(a^{\widetilde{t}})$, for all $\widetilde{t} > t$, make player $i+1$ indifferent across all sequences of action profiles $a^{\widetilde{t}}$, $\widetilde{t} > t$. Note that, in particular, player $(i+1)$'s auxiliary-scenario continuation payoff is independent of the history (of player i or his own) at the beginning of period $t+1$. To pick values $\theta(a^t)$ that make player $i+1$ indifferent over all action profiles of the stage game in period t, we must just a system of linear equations (on the values $\theta(a^t)$), whose matrix tends to the identity matrix as monitoring becomes 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 (12) 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, i.e. the values $\theta(h_i^{t-1}, a^t, I_{-(i+1)})$ where t belongs to one of these three phases.

Given a period t of Phase 4, we denote by $J_{-(i+1)}^t$ (respectively $I_{-(i+1)}^t$) 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 $\widetilde{t} < t$ of Phase 4, 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 $J_{-(i+1)}^t \in \mathcal{J}_{-(i+1)}^t$ (respectively $I_{-(i+1)}^t \in \mathcal{I}_{-(i+1)}^t$) when all players other than $i+1$ reveal the same information about the play in Phases 1 and 4 up to period t, and according to this information, either they played $a^{\widetilde{M}}$ up to period t or player $i+1$ was the first to unilaterally deviate in some period $\widetilde{t} < t$.

We will define the values $\theta(h_i^t, a^t, I_{-(i+1)})$, which depend only on a^t and $I_{-(i+1)}$, 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 $J_{-(i+1)}^t$ (both from $\mathcal{J}_{-(i+1)}^t$ and from the complement of $\mathcal{J}_{-(i+1)}^t$);

4. His payoff from period t on, augmented by the transfers θ assigned from period t on (until the end of Phase 4), conditional on every $J_{-(i+1)}^t \in \mathcal{J}_{-(i+1)}^t$, 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 $J_{-(i+1)}^t$.

Condition 3 then guarantees that player $i+1$ is indifferent over all his strategies and condition 4 guarantees that (12) is satisfied. Recall that backward induction with respect to t is used. 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 probability 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 $\theta(h^t_i, a^t, I_{-(i+1)})$ such that player $i+1$ is indifferent across all consequences (the stage-game payoff and the transfers θ assigned in period t, and the induced probability distribution over $J^{t+1}_{-(i+1)}$) of all his actions in period t. Indeed, since player $i+1$ is indifferent over all his strategies from period $t+1$ on, conditional on every $J^{t+1}_{-(i+1)}$, by the inductive assumption, he must be indifferent over all his strategies from period $t+1$ on conditional on every probability distribution over $J^{t+1}_{-(i+1)}$, in particular the distribution induced by his action in period t.

To define transfers θ that satisfy conditions 3 and 4, notice first that it is straightforward to define the transfers $\theta^{/}(h^t_i, a^t, J_{-(i+1)})$ that depend directly on $J^t_{-(i+1)}$ (instead of $I^t_{-(i+1)}$) making player $i+1$ indifferent between the consequences of all his actions in period t. Indeed, for small enough values of ε, it is possible to pick $\theta^{/}(h^t_i, a^t, J_{-(i+1)})$ that offset the differences in the stage-game payoffs and the induced probability distribution over $J^{t+1}_{-(i+1)}$ across all actions of player $i+1$.

The values $\theta(h^t_i, a^t, I_{-(i+1)})$ of course can 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 $\theta(h^t_i, a^t, I_{-(i+1)})$ as the vector of values $\theta^{/}(h^t_i, a^t, J_{-(i+1)})$ multiplied by D^{-1}.

The values $\theta(h^t_i, a^t, I_{-(i+1)})$ just defined need not be non-negative. We can make them non-negative by adding a constant tending to 0 as $\varepsilon \to 0$. Indeed, one could directly pick non-negative values for $\theta^{/}(h^t_i, a^t, J_{-(i+1)})$; since the matrix D (and so D^{-1}) tends to the identity matrix as $\varepsilon \to 0$, one then gets the values of $\theta(h^t_i, a^t, I_{-(i+1)})$ converging to non-negative numbers as $\varepsilon \to 0$.

This yields condition 3. Condition 4 is easy to satisfy for all $J^t_{-(i+1)}$ such that player $i+1$ is the first player who unilaterally deviated in some period $\widetilde{t} < t$; then all other players minmax player $i+1$ in period t and in all following periods of Phase 4, and thus, the value $\theta^{/}(h_i^t, a^t, J_{-(i+1)})$ can be set as 0 if a_{i+1}^t is player $(i+1)$'s best response to the minmaxing strategy of other players. Suppose therefore that $J_{-(i+1)}^t$ is such that 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 a specific $J_{-(i+1)}^{t+1} \in \mathcal{J}_{-(i+1)}^{t+1}$ (more precisely, it is $J_{-(i+1)}^{t+1}$ such that all players played a^M until period $t+1$ if player $i+1$ takes action $a_{i+1}^{\widetilde{M}}$; and it is $J_{-(i+1)}^{t+1}$ such that player $i+1$ is the first unilaterally deviator from $a^{\widetilde{M}}$ in period t otherwise). It thus follows 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, can be made to converge (as $\varepsilon \to 0$) to the maximum of his continuation payoffs under perfect monitoring over all continuation strategies conditional on $J_{-(i+1)}^t$; it can be done by setting the transfer $\theta^{/}(h_i^t, a^t, J_{-(i+1)})$ equal to 0 when player $i+1$ plays a_{i+1}^t maximizing this augmented payoff from period t on (given that players played $a^{\widetilde{M}}$ until period t).

Finally, by applying an argument similar to that used for Phases 2, 3, and 5, pick transfers $\theta(h_i^0, a^1, I_{-(i+1)})$ that make player $i+1$ indifferent across all actions in Phase 1 conditional on every $J_{-(i+1)}^1$, and such that (12) is satisfied.

(**b & c**) Let $J_{-(i+1)}^1$ 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 $I_{-(i+1)}^1$ 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:

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

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

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 as in (**a**). Therefore consider first a period t from Phase

4. We will specify a system of linear equations such that if the values of θ satisfy the system, then (13), (14) 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 $\theta(h_i^t, a^t, I_{-(i+1)}^1)$ for any h_i^t and $J_{-(i+1)}^1$, such that $J_{-(i+1)}^1$ 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 $J_{-(i+1)}^1$ (again, write then $\theta^{/}(h_i^t, a^t, J_{-(i+1)}^1)$ instead of $\theta(h_i^t, a^t, I_{-(i+1)}^1)$) our equations would be simple and they would be the same for all pairs h_i^t and $J_{-(i+1)}^1$ with the required property. Namely, assume that the values $\theta^{/}(h_i^{\widetilde{t}}, a^{\widetilde{t}}, J_{-(i+1)}^1)$, $\widetilde{t} > t$, make player $i+1$ indifferent over all sequences of action profiles $a^{\widetilde{t}}$, $\widetilde{t} > t$, when $J_{-(i+1)}^1$ 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 $\theta^{/}(h_i^t, a^t, J_{-(i+1)}^1)$ is equal to a negative number

$$c < -[\max_{a \in A} u_{i+1}(a) - \min_{a \in A} u_{i+1}(a)]. \tag{15}$$

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 (15) and the form of our equations, all values $\theta(h_i^t, a^t, J_{-(i+1)}^1)$ of the solution are negative for small enough ε.

However, the values $\theta(h_i^t, a^t, I_{-(i+1)}^1)$ do not depend directly on $J_{-(i+1)}^1$ but only on $I_{-(i+1)}^1$. Let D_1 be the matrix obtained by stacking the probability distributions over $I_{-(i+1)}^1$ conditional on $J_{-(i+1)}^1$ as row vectors. Note that the matrix D_1 tends to the identity matrix as $\varepsilon \to 0$. For any h_i^t and $J_{-(i+1)}^1$ such that $J_{-(i+1)}^1$ 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 $\theta(h_i^t, a^t, I_{-(i+1)}^1)$ which obtains from that for the values $\theta^{/}$ by replacing $\theta(h_i^t, a^t, J_{-(i+1)}^1)$ with D_1 multiplied by the vector of $\theta(h_i^t, a^t, I_{-(i+1)}^1)$.

In this way, we impose $\#A$ equations for any h_i^t and $J_{-(i+1)}^1$ such that $J_{-(i+1)}^1$ does not coincide with h_i^t on the report of player i from Phase 2. Each of these equations has non-zero coefficients only at the values $\theta(h_i^t, a^t, I_{-(i+1)}^1)$ for a given h_i^t. The non-zero coefficients converge to 1 when $I_{-(i+1)}^1$ corresponds to $J_{-(i+1)}^1$ 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, assume that the values of θ depend directly on the message profile sent by player $i + 1$ in Phase 3, and then replace those values by a matrix D_2 multiplied by the vector of values that depends only on player i's signals about that message profile, where the matrix D_2 is obtained by stacking the probability distributions over player i's signals about the message profile sent by player $i + 1$ in Phase 3 induced by those message profiles.

In this way, we impose $\#A$ equations on the transfers $\theta(h_i^t, a^t, I_{-(i+1)}^1)$ for any h_i^t such that a message profile sent by player $i + 1$ in Phase 3 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. 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 $I^1_{-(i+1)}$ corresponds to $J^1_{-(i+1)}$ that differs from h^t_i about the report of player i in Phase 2.

Consider now a period t of Phase 4 and 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. Then (13) can be guaranteed in a similar manner to the proof of Lemma 1; (14) 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 linear 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 $\theta(h^t_i, a^t, I^1_{-(i+1)})$ such that the part of h^t_i 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 $I^1_{-(i+1)}$ coincides with h^t_i on the report of player i from Phase 2. This guarantees that the system of that we just described consists of linearly independent equations; moreover, this system of equations combined with the system that guarantees (**c**) is still linearly independent.

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 the signals that he observed about player i's report in Phase 2. The latter requirement is easy to achieve, because it suffices to set $\theta(h^t_i, a^t, I^1_{-(i+1)}) = 0$ 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 $\theta(h^t_i, a^t, I^1_{-(i+1)}) = c$, satisfying condition (15), otherwise. Indeed, player $i+1$ cannot then benefit from incorrectly repeating due to a higher flow payoff (the stage-game payoff and the transfer assigned in the current period) 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 repeat truthfully in Phase 3 the observed signals about player

i's report in 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 of 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 messages sent 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 is also essential that in Phase 3 players send messages sequentially, and that signals coincide with action profiles of other players. Suppose instead that messages were not sent sequentially in Phase 3 and erroneous signals are highly correlated. Then it may happen that player $i+1$, upon receiving signals according to which another player did not repeat correctly in Phase 3 his predecessor's message from Phase 2, assigns high probability to the event that player i received an erroneous signal about one of his previous messages 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 conditional on this signal, player i's distribution over his signals is independent of player $i+1$'s action. In this case, player $i+1$ would no longer have an incentive to repeat in Phase 3 player i's report in Phase 2.

Finally, by an argument similar to that from (**a**), we can ensure that player $i+1$ is indifferent over all actions in Phases 1 and 2. To see this, observe that player $(i+1)$'s actions in Phase 2 do not affect his payoff conditional on regular histories, and they can alter the probability of reaching an erroneous history only marginally, i.e. with probability converging to 0 as $\varepsilon \to 0$. Similarly, player $(i+1)$'s action in Phase 1 (period 1) affect only slightly, although not marginally, his total payoff. The difference compared to Phase 2 is that player $(i+1)$'s action in period 1 also affects the action profile a^M 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.*