

Folk Theorem with Communication¹

Ichiro Obara
Department of Economics
University of Minnesota and University of California Los Angeles

First Version August 5, 2005
This version February 7, 2008

¹I am grateful to the associate editor, an anonymous referee, Galit Ashkenazi-Golan, Michihiro Kandori, and George Mailath for their helpful comments. I also thank the seminar participants at 2005 SAET conference at Vigo and University of Southern California.

Abstract

This paper proves a new folk theorem for repeated games with private monitoring and communication, extending the idea of delayed communication in Compte [6] to the case where private signals are correlated.

The sufficient condition for the folk theorem is generically satisfied with more than two players, even when other well-known conditions are not. The folk theorem also applies to some two-players repeated games.

Keywords: Communication, Folk Theorem, Private Monitoring, Repeated Games

JEL classification codes: C72, C73, D82

1 Introduction

We have observed a significant progress in repeated games with private monitoring in the last few years. It started with a series of papers which proved a folk theorem with communication, such as Ben-Porath and Kahneman [4], Compte [6], and Kandori and Matsushima [11]. They are very important contributions to the theory of long term relationships, especially because repeated games with private monitoring are very difficult to analyze without communication. These folk theorems, however, do not cover all the interesting cases because of the specific assumptions they require on private monitoring structure. This note proves a new folk theorem with communication to expand the range of environments to which the folk theorem applies.

The main contribution of this note is to extend the idea of delayed communication in Compte [6] to the case where private signals are correlated. Compte [6] focuses on T -public equilibria in which players play the same action for T periods and announce their accumulated private signals truthfully only every T periods. These private signals in each T -period block are used to “test” whether each player has deviated or not within the same block. A player is punished at the end of the block with a lower continuation payoff when the private signals reported by the other players look “bad”. In conducting this statistical test, it is important that a player does not learn the likelihood of her punishment from her own private signals. If she is confident that she will not be punished in the end of a T -period block, she may start deviating toward the end of the block. Compte [6] avoids this problem by assuming conditional independence between players’ private signals. However, conditional independence is a nongeneric assumption.¹ Furthermore, it is difficult to introduce even a slight correlation of private signals. This is because T must go to infinity to obtain the exact folk theorem, hence players may be able to obtain a large amount of information from their accumulated private signals even if each signal has a limited information.²

This note proposes a new condition which serves the same purpose as conditional independence, thus making it possible to apply the idea of delayed communication even when private signals are correlated. The condition is generically satisfied for most of stage games when the number of players is more than two. When the number of actions and signals is the same across players, this condition is generically satisfied even more easily than the sufficient condition proposed by Kandori and Matsushima [11].

There are a few recent contributions proving folk theorems for repeated games with private monitoring and communication. They pay extra attention to the case with two players: the case that was not extensively analyzed in the initial contribu-

¹Compte [6] does allow for some correlation of private signals off the equilibrium path.

²As observed by Abreu, Milgrom and Pearce [1], this type of statistical test becomes more effective as T becomes large. In fact, the cost of expected punishment converges to 0 as T goes to infinity and players become infinitely patient.

tions.³ Fudenberg and Levine [9] proves a Nash-threat folk theorem when players' private signals are highly correlated. Ashkenazi-Golan [2] assumes that deviations are perfectly observable by at least one player with positive probability and proves a Nash-threat folk theorem. These results, as well as the result of this note, apply to repeated games with two or more players. Finally, McLean, Obara and Postlewaite [15] proves a folk theorem when private signals are correlated and can be treated like a public signal once aggregated. But this result requires at least three players.

I also should mention that many folk theorem results without communication have been obtained recently. However, most of them assume almost perfect monitoring (Bhaskar and Obara [5], Ely and Välimäki [7], Hörner and Olszewski [10], and Mailath and Morris [12]).⁴ One exception is Matsushima [14] that allows for noisy private monitoring. However he assumes a certain type of conditional independence of private signals as in Compte [6]. The result of this note may be useful to deal with noisy correlated private signals even without communication, but that is left for future research.

The next section presents the model briefly. Section 3 introduces the assumptions on monitoring structure. Section 4 presents the main result and Section 5 discusses some extension.

2 Model

Stage game $G = (I, A, g)$ is defined as follows. The set of players is $I = \{1, 2, \dots, n\}$, $n \geq 2$. In each period, player $i \in I$ chooses an action from a finite action set A_i (simultaneously with the other players) and observes a private signal s_i from a finite set S_i . Both a_i and s_i are private information, observable only to player i . Let $p(s|a)$ be the probability of $s \in S = \prod_{i=1}^n S_i$ given $a \in A = \prod_{i=1}^n A_i$. It is assumed that $p(s|a)$ has full support on S for every $a \in A$. The signal distribution on $S_{-i} = \prod_{j \neq i} S_j$ for player i given $a \in A$ (and $s_i \in S_i$) is denoted by $p_{-i}(s_{-i}|a)$ ($p_{-i}(s_{-i}|a, s_i)$). Player i 's (expected) stage game payoff $g_i : A \rightarrow \mathfrak{R}$ is given by $g_i(a) = \sum u_i(a_i, s_i) p(s|a)$ where $u_i : A_i \times S_i \rightarrow \mathfrak{R}$ is a realized payoff of player i . Let $\mathbf{g}(a) = (g_1(a), \dots, g_n(a))$ be the expected payoff profile given $a \in A$ and $V = \text{co}\{\mathbf{g}(a) | a \in A\} \subset \mathfrak{R}^n$ be the feasible payoff set.⁵ Abusing notations slightly, let $g_i(\alpha)$ be player i 's expected stage game payoff given a mixed action profile $\alpha \in \prod_{i=1}^n \Delta A_i$ and $\mathbf{g}(\alpha) = (g_1(\alpha), \dots, g_n(\alpha))$. The set of Nash equilibrium profiles is denoted by $NE \subset \prod_{i=1}^n \Delta A_i$. Let $V^{NE} = \text{co}\{v \in V | \exists \alpha \in NE \text{ s.t. } v \gg \mathbf{g}(\alpha)\}$.

This stage game G is played repeatedly over time. In the end of each period, players send messages $m = (m_1, \dots, m_n) \in M = \prod_{i=1}^n M_i$ simultaneously, which

³However, Theorem 2 in Compte [6] applies to two-player games. Kandori and Matsushima [11] spends one section to prove a folk theorem for the repeated prisoners' dilemma game with conditional independence and private monitoring.

⁴Also see Mailath and Morris [13]

⁵For any $X \subset \mathfrak{R}^n$, $\text{co}X$ is the convex-hull of X .

are publicly observed. Private history and public history at period $t = 1, 2, \dots$ are $h_i^t = (a_{i,1}, s_{i,1}, \dots, a_{i,t-1}, s_{i,t-1}) \in (A_i \times S_i)^{t-1}$, $i \in I$ and $h^t = (m_1, \dots, m_{t-1}) \in M^{t-1}$ respectively.⁶ Player i 's strategy σ_i consists of an *action strategy* σ_i^a and a *report strategy* σ_i^m . Player i 's action strategy σ_i^a is a collection of period t behavior strategies $(\sigma_{i,t}^a, t = 1, 2, \dots)$, which map (h_i^t, h^t) into A_i . Player i 's report strategy σ_i^m is a collection of period t behavior strategies $(\sigma_{i,t}^m, t = 1, 2, \dots)$, which map (h_i^{t+1}, h^t) into M_i . They also depend on a realization of a randomization device. But I avoid introducing additional notations here to simplify the exposition. It will be mentioned when the public randomization device is introduced.⁷

The players discount their future payoffs by common discount factor $\delta \in (0, 1)$. Their discounted average payoffs are $E[(1 - \delta) \sum_{t=1}^{\infty} \delta^{t-1} g_i(a_t) | \sigma]$, $i \in I$.

3 Assumptions on Monitoring Structure

I introduce and discuss a few important conditions in this section. First, it is assumed that player i 's deviation is statistically detectable from private signals of the other players. For given $a \in A$, let $Q_i(a)$ be the convex-hull of probability vectors $\{p_{-i}(\cdot | a'_i, a_{-i}) \in \mathfrak{R}_+^{S_{-i}} | a'_i \neq a_i\}$. Then this assumption can be formally stated as follows.

Assumption 1

$$p_{-i}(\cdot | a) \notin Q_i(a) \text{ for every } a \in A \text{ and } i \in I.$$

This assumption is standard. In Kandori and Matsushima [11], it is implied by assumption (A2). In Compte [6], this is satisfied when there exists an unbiased monitor for each player.

As is well known, this guarantees that $a_i \in A_i$ is strictly enforceable through transfers contingent on s_{-i} .

Lemma 1 *Assumption 1 is satisfied at $a \in A$ for $i \in I$ if and only if there exists $x_i : S_{-i} \rightarrow \mathfrak{R}$ such that, for all $a'_i \neq a_i$,*

$$\sum_{s_{-i} \in S_{-i}} p_{-i}(s_{-i} | a) \cdot x_i(s_{-i}) > \sum_{s_{-i} \in S_{-i}} p_{-i}(s_{-i} | a'_i, a_{-i}) \cdot x_i(s_{-i})$$

The next condition is the key for the folk theorem. Fix $a \in A$ and $s_i \in S_i$ and consider all the conditional beliefs on S_{-i} when player i observes a different signal or

⁶I follow the convention and set $h_i^t = h^t = \emptyset$ in $t = 1$.

⁷It is possible to generate any public randomization device endogenously through (direct) communication by redefining i 's message space as $M_i \times [0, 1]$ for each $i \in I$ and using a jointly-controlled lottery (Aumann, Maschler and Stearns [3]).

deviates from a_i . Let $R_i(a, s_i)$ be the convex-hull of all such conditional probability vectors, i.e. $co \left\{ p_{-i}(\cdot | a'_i, a_{-i}, s'_i) \in \mathfrak{R}_+^{|S_{-i}|} | (a'_i, s'_i) \neq (a_i, s_i) \right\}$.

Definition 1 *Player i is an informed player at $a \in A$ if*

$$p_{-i}(\cdot | a, s_i) \notin R_i(a, s_i) \text{ for any } s_i \in S_i \quad (\text{IP})$$

When player i is informed at a , her conditional belief given any private signal is not a linear combination of her conditional beliefs given different signals or unilateral deviations. See the following example.

Example

Let $S_i = \{\bar{y}_i, \underline{y}_i\}$, $A_i = \{C, D\}$, and $I = \{1, 2\}$. Player 1 is an informed player at (C, C) if and only if (after relabeling \bar{y}_2 and \underline{y}_2 if necessary)

$$\begin{aligned} p_2(\bar{y}_2 | CC, \bar{y}_1) &> \max \left\{ p_2(\bar{y}_2 | CC, \underline{y}_1), p_2(\bar{y}_2 | DC, \bar{y}_1), p_2(\bar{y}_2 | DC, \underline{y}_1) \right\} \\ p_2(\underline{y}_2 | CC, \underline{y}_1) &> \max \left\{ p_2(\underline{y}_2 | CC, \bar{y}_1), p_2(\underline{y}_2 | DC, \bar{y}_1), p_2(\underline{y}_2 | DC, \underline{y}_1) \right\} \end{aligned}$$

Note that player 1 cannot be an informed player at (D, C) when she is informed at (C, C) .⁸

This condition is satisfied in the following example. Suppose that there is a hidden signal $y \in \{\bar{y}, \underline{y}\}$. Each player's private signal $s_i, i = 1, 2$ is a noisy conditionally independent observation of y . Player i 's private signal s_i is correct (\bar{y}_i for \bar{y} and \underline{y}_i for \underline{y}) with probability $1 - \varepsilon'$ when $a_i = C$ and correct with probability $1 - \varepsilon''$ when $a_i = D$. Assume also that $\varepsilon' < \varepsilon'' < 1/2$. That is, player i 's private signal is more informative when C is chosen. When $\frac{\varepsilon'}{\varepsilon''}$ is small enough, the conditional probability of $\bar{y}_j(\underline{y}_j), j \neq i$ is largest when $a_i = C$ is played and $\bar{y}_i(\underline{y}_i)$ is observed. Thus player 1 (and player 2, too) is an informed player at (C, C) . Note that player 1 is also informed at (C, D) .

A few remarks are in order. First, player i 's private signal given $a \in A$ must be correlated with the other players' private signals when she is informed at a . If private signals are conditionally independent, then $p_{-i}(\cdot | a, s_i)$ does not depend on s_i . Hence this condition is complementary to the independence assumption in Compte [6]. Second, (IP) is very easily satisfied with three players or more. Indeed it is generically satisfied for every player at every action profile when $|A_i| \times |S_i| \leq \prod_{j \neq i} |S_j|$, $i \in I$ (with $|S_j| \geq 2$ for every $j \in I$) because $\{p_{-i}(\cdot | a, s_i) : a_i \in A_i, s_i \in S_i\}$ are generically linearly independent in that case. Kandori and Matsushima [11]'s condition also holds generically with three players or more. To compare my result to

⁸If there are at least three signals for player 2, then there are monitoring structures for which player 1 is informed at all action profiles.

theirs, consider a stage game with $m = |A_i|$ and $K = |S_i|$ for all $i \in I$. Kandori and Matsushima [11] requires $|A_i| + |A_j| - 1 \leq |S_{-i,j}|$ for every i, j for their condition to be satisfied generically, which reduces to $2m - 1 \leq K^{n-2}$, whereas the above inequality reduces to a weaker condition: $m \leq K^{n-2}$. Third, (IP) is less easily satisfied with two players. Indeed $|A_1| \times |S_1| \leq |S_2|$ and $|A_2| \times |S_2| \leq |S_1|$ cannot be satisfied at the same time, thus there must exist some action profile where some player is not informed. However, my folk theorem may still hold even in such cases because (1) some action profiles are not relevant for the proof of the folk theorem and (2) not all players have to be informed. I just need one player to be informed at relevant action profiles. The last section discusses the possibility to relax the sufficient condition regarding informed players even further. Finally, this condition does not presume any type of almost public monitoring or almost common knowledge (Fudenberg and Levine [9], Mailath and Morris [12], and McLean, Obara and Postlewaite [?]).

The crucial step of the proof is to construct a binary signal for player i from reported signal profiles, which is used to punish her when she deviates in action and/or in announcements. The binary signal has the following two properties: (1) it is informative about player i 's action and message and (2) player i does not update her belief about the signal based on her private signal s_i in equilibrium, i.e. her private signal is not informative about whether she will be punished or not. To be more precise, let me introduce the following condition.

Definition 2. *Player i can be secretly screened at $a \in A$ if there exists $q_i : S \rightarrow [0, 1]$ such that*

$$E[q_i(s) | a] < E[q_i(\rho(s_i), s_{-i}) | a'_i, a_{-i}] \quad (1)$$

when either $a'_i \neq a_i$ or $\rho : S_i \rightarrow S_i$ is not the identity function, and

$$E[q_i(s) | a] = E[q_i(s) | a, s_i] \text{ for all } s_i \in S_i \quad (2)$$

This q_i can be interpreted as the probability of the “bad” outcome of the binary signal. The first condition means that the probability of a bad outcome goes up when player i either deviates from a_i or sends a false message. The second condition means that the expected value of q_i is unaffected by player i 's private information when she plays a_i and announces her signal truthfully. The following key lemma proves that (IP) holds at $a \in A$ if and only if player i can be secretly screened at $a \in A$.

Lemma 2 *Player i is an informed player at $a \in A$ if and only if player i can be secretly screened at $a \in A$.*

Proof. Suppose that player i is an informed player at $a \in A$. By the hyperplane theorem, there exists $q_i : S \rightarrow \mathbb{R}$ to satisfy

$$\sum_{s_{-i}} q_i(s) p_i(s_{-i} | a, s_i) < \sum_{s_{-i}} q_i(s) p_i(s_{-i} | (a'_i, a_{-i}), s'_i)$$

for all $(a'_i, s'_i) \neq (a_i, s_i)$ and for all $s_i \in S_i$.

Without loss of generality, I can choose $q_i(s)$ so that $E[q_i(s)|a, s_i] = \pi$ for all s_i for some $\pi \in (0, 1)$ (add some appropriate constant to $q_i(s)$ for each s_i). Furthermore, $q_i(s)$ can be chosen from $(0, 1)$ for all $s \in S$. If not, then redefine $q_i(s)$ as $\tilde{q}_i(s) = (q_i(s) + z\pi) \frac{1}{z+1}$ for some large $z > 0$. Then this $\tilde{q}_i(s)$ still satisfies all the strict inequalities and lies in $(0, 1)$ if z is large enough.

For this $q_i(s)$, (2) is satisfied because

$$\begin{aligned} E[q_i(s)|a] &= \sum_{s_i} E[q_i(s)|a, s_i] \Pr(s_i|a) \\ &= \sum_{s_i} \pi \Pr(s_i|a) \\ &= \pi. \end{aligned}$$

The above strict inequalities imply that $\sum_{s_{-i}} q_i(\rho(s_i), s_{-i}) p_i(s_{-i}|(a'_i, a_{-i}), s_i)$ is strictly larger than π when either $a'_i \neq a_i$ or $s_i \neq \rho(s_i)$. This implies that (1) holds when either $a'_i \neq a_i$ or $\rho: S_i \rightarrow S_i$ is not an identity function, because

$$\begin{aligned} &E[q_i(\rho(s_i), s_{-i})|a'_i, a_{-i}] \\ &= \sum_{s_i} \sum_{s_{-i}} q_i(\rho(s_i), s_{-i}) p_i(s_{-i}|(a'_i, a_{-i}), s_i) \Pr(s_i|a'_i, a_{-i}) \\ &> \pi \\ &= E[q_i(s)|a]. \end{aligned}$$

Conversely, suppose that player i can be secretly screened at $a \in A$. Then, for each s_i , (a_i, s_i) is the unique solution of

$$\max_{(a'_i, s'_i) \in A_i \times S_i} \sum_{s_{-i}} q_i(s) p_i(s_{-i}|(a'_i, a_{-i}), s'_i).$$

Then $p_i(s_{-i}|a, s_i)$ is an extreme point of $co\{p_{-i}(\cdot|a'_i, a_{-i}, s'_i) \in \mathfrak{R}_+^{|S_{-i}|} | (a'_i, s'_i) \in A_i \times S_i\}$, which implies that $p_{-i}(\cdot|a, s_i) \notin R_i(a, s_i)$. ■

The key condition for the main result is (2). Compte [6] uses a similar binary signal for which q_i does not depend on s_i . In this case, no learning ((2)) follows automatically from conditional independence, (IP) can be regarded as an alternative, possibly generic, condition, which guarantees that (2) is satisfied for some q_i even when private signals are correlated.

The Role of Informed Players

Since (1) and (2) are central to the proof of the folk theorem, it may be useful to explain their roles in the formal proof of the folk theorem.

As is often the case, for each utility weight $\lambda \in \mathfrak{R}^n$, I try to support $a \in A$ that maximizes $\lambda \cdot g(a)$ while minimizing “efficiency losses” of continuation payoffs (in the direction of λ).

Suppose that player i is informed at $a \in A$ and $\lambda_i > 0$. The incentive of player $j \neq i$ is provided through transfers $x_j(s_{-j})$ (by Lemma 1) period by period. Note that player j 's revelation constraints are satisfied by definition because x_j does not depend on s_j . To avoid efficiency losses, $-\frac{\lambda_j}{\lambda_i} x_j(s_{-j})$ is added to player i 's transfer. This is as if player i 's stage game payoff is transformed into $g_i(a) - \sum_{j \neq i} \frac{\lambda_j}{\lambda_i} x_j(s_{-j})$. Note that this creates an incentive problem for player i because she may have incentive to misrepresent her signal to control the transfers from j . This problem can be taken care of by using q_i as explained below. This is another reason (in addition to “no learning”) why it is useful to make q_i dependent on s_i in addition to s_{-i} . Compte [6] avoids this problem by making $x_j(s_{-i,j})$ independent of s_i , but this does not allow the case with two players.

Player i 's incentive is provided through punishments. This is where (1) is used. The probability of a bad signal, hence the probability of punishment, increases when the distribution of s_{-i} suggested by player i 's message does not match with the true distribution of s_{-i} .⁹ This punishment is costly because it occurs with positive probability (and $\lambda_i > 0$). To reduce the expected cost of punishments, private signals are stored and revealed only in every T periods. The efficiency of monitoring is improved by using such accumulated private signals. Indeed the expected cost of punishment goes to 0 as $T \rightarrow \infty$ (and $\delta \rightarrow 1$). For this type of “review strategies” to work, player i must be kept from learning the likelihood of punishment within each T period review phase. Condition (2) guarantees that such learning is not happening.

It is also possible to use player i with $\lambda_i < 0$ who can be secretly screened at a . In this case, player i 's incentive needs to be provided through rewards rather than punishments. This extension is discussed in the last section.

4 Folk Theorem with Delayed Communication

To prove a folk theorem, *T-public equilibrium* is employed as in Compte [6] and Kandori and Matsushima [11]. In T -public equilibrium, players reveal their private signals publicly only every T periods.¹⁰ Player i 's action strategy σ_i^a depends on

⁹Player i 's incentive for truthful announcements is not an issue in Compte [6] because he uses a slightly stronger version of Assumption 1 (for Theorem 1): player j 's deviation can be detected from $s_{-i,j}$ for any $i \neq j$ rather than s_{-j} (note that this requires three players). Then x_j can be just a function of $s_{-i,j}$, thus player i does not have an incentive to manipulate player j 's transfer. In this case, q_i does not need to depend on s_i , hence (1) can be replaced by $E[q_i(s_{-i})|a] < E[q_i(s_{-i})|a'_i, a_{-i}]$ for all $a'_i \neq a_i$. Therefore there is no incentive problem regarding announcements in Compte [6] because neither x_j nor q_i does not depend on s_i .

¹⁰This can be interpreted as players announcing meaningless random messages until the end of each T -period block. Note also that player i 's message space needs to be at least as large as S_i^T .

only such publicly announced private signals. Player i 's report strategy σ_i^m depends on public signals as well as her private history within the most recent T -period block. Let $E(\delta, T)$ be the set of all T -public equilibrium payoffs given δ . A T -public equilibrium in which players play the same action every period within each T -period block is called a stationary T -public equilibrium.

A few more notations need to be introduced to state the theorem. Let $\Lambda = \{\boldsymbol{\lambda} \in \mathfrak{R}^n \mid \|\boldsymbol{\lambda}\| = 1\}$ be the space of utility weights. Let $A_\tau(\boldsymbol{\lambda}) \subset A$ be the subset of pure action profiles where *there exists at least one informed player with a positive weight larger than τ* . This τ can be an arbitrarily small positive number. Define the following value for each $\boldsymbol{\lambda} \in \Lambda$,

$$k_\tau(\boldsymbol{\lambda}) \equiv \max \left\{ \max_{a \in A_\tau(\boldsymbol{\lambda})} \boldsymbol{\lambda} \cdot \mathbf{g}(a), \max_{\alpha \in NE} \boldsymbol{\lambda} \cdot \mathbf{g}(a) \right\}.$$

Let $D_\tau(\boldsymbol{\lambda}) \equiv \{\mathbf{v} \in \mathfrak{R}^n \mid \boldsymbol{\lambda} \cdot \mathbf{v} \leq k_\tau(\boldsymbol{\lambda})\}$. I say a convex set in \mathfrak{R}^n is full dimensional if it has an interior point in \mathfrak{R}^n .

Theorem 1 *Fix any $\tau > 0$. Suppose that Assumption 1 is satisfied and $\cap_{\boldsymbol{\lambda} \in \Lambda} D_\tau(\boldsymbol{\lambda})$ is full dimensional. Then, for any smooth set W in the interior of $\cap_{\boldsymbol{\lambda} \in \Lambda} D_\tau(\boldsymbol{\lambda})$, there exists $\underline{\delta} \in (0, 1)$ and an integer T such that $W \subset E(\delta, T)$ for all $\delta \in (\underline{\delta}, 1)$.*

Proof. See Appendix. ■

An immediate corollary of this theorem is the following Nash-threat folk theorem.

Corollary 1 *Suppose that Assumption 1 is satisfied and V^{NE} is full dimensional. Also suppose that every player is informed at every action profile that achieves an extreme point of V . Then, for any $v \in \text{int}V^{NE}$, there exists $\underline{\delta} \in (0, 1)$ and an integer T such that $v \in E(\delta, T)$ for all $\delta \in (\underline{\delta}, 1)$.*

Proof. First note that $V^{NE} \subset \lim_{\tau \rightarrow 0} \cap_{\boldsymbol{\lambda} \in \Lambda} D_\tau(\boldsymbol{\lambda})$. For each $v \in \text{int}V^{NE}$. Pick a smooth set W s.t. $v \in W \subset \text{int}V^{NE}$. Then pick small enough $\tau > 0$ such that $W \subset \text{int}D_\tau(\boldsymbol{\lambda})$ and apply the above theorem. ■

The proof of the theorem consists of two steps. The first step is to transform an infinitely repeated game into a T -period game with side transfers. Consider the best stationary T -period equilibrium in the direction of $\boldsymbol{\lambda}$ for this T -period game. The equilibrium payoff provides an upper bound of the best stationary T -public equilibrium of the original infinitely repeated game. It is well known that this bound is indeed tight, i.e. the area surrounded by a collection of such bounds indexed by $\boldsymbol{\lambda}$ is approximately the set of stationary T -public equilibrium payoffs of the original infinitely repeated game as $\delta \rightarrow 1$. In the second step, it is shown that this area contains $\cap_{\boldsymbol{\lambda} \in \Lambda} D_\tau(\boldsymbol{\lambda})$ for any $\tau > 0$ as $T \rightarrow \infty$.

There is nothing new in the first step. Here I briefly copy the results from Compte [6]. Consider the following T -period game with side transfers: stage game G is played T times and players announce their private signals $m_i = (s_{i,1}, \dots, s_{i,T})$, $i = 1, \dots, n$ at the end of the T periods, on which their side transfers $x_i(m)$, $i \in I$ are based. Let $\sigma_i^{a,T}$ be player i 's T -period action strategy and $\sigma_i^{m,T}$ be player i 's report strategy, which are essentially a T -period truncated version of σ_i^a and σ_i^m respectively. Let $\sigma_i^T = (\sigma_i^{a,T}, \sigma_i^{m,T})$ be player i 's T -period strategy. Player i 's payoff from this T -period game given $\delta_0 \in (0, 1)$ is defined by

$$g_i^{T,\delta_0}(\sigma^T) + E[x_i(m) | \sigma^T]$$

where

$$g_i^{T,\delta_0}(\sigma^T) = \sup_{\delta \in [\delta_0, 1]} \frac{(1 - \delta) \sum_{t=1}^T \delta^{t-1} E[g_i(a_t) | \sigma^T]}{1 - \delta^T} \quad (3)$$

One can regard $g_i^{T,\delta_0}(\sigma^T)$ as player i 's average payoff within the first T periods of the original infinitely repeated game and $x_i(m)$ as the variation of continuation payoffs.¹¹

A T -period strategy profile σ^T is called *stationary* if it specifies the same (mixed) action profile every period and all private signals are announced truthfully at the end of the T periods. Player i 's stationary strategy with $\alpha_i \in \Delta A_i$ is denoted $\sigma_i^T(\alpha_i)$. Let $\sigma^T(\alpha)$ be a profile of stationary strategies. Note that $g_i^{T,\delta_0}(\sigma^T(\alpha)) = g_i(\alpha)$ for any α . I say $v \in \mathfrak{R}^n$ is *generated by* $(\sigma^T(\alpha), x)$ if

$$v_i = g_i(\alpha) + E[x_i(m) | \sigma^T(\alpha)]$$

and $\sigma^T(\alpha)$ is a Nash equilibrium (called *stationary T -period equilibrium*), that is

$$\begin{aligned} v_i &= g_i(\alpha) + E[x_i(m) | \sigma^T(\alpha)] \\ &\geq g_i^{T,\delta_0}(\sigma_i^{T'}, \sigma_{-i}^T(\alpha)) + E[x_i(m) | \sigma_i^{T'}, \sigma_{-i}^T(\alpha)] \\ &\text{for any } T\text{-period strategy } \sigma_i^{T'} \text{ and } i \in I. \end{aligned}$$

Note that $\sigma^T(\alpha)$ is just a Nash equilibrium, not necessarily a sequential equilibrium. However, since full support is assumed, there exists an outcome equivalent sequential equilibrium for any Nash equilibrium.

Now consider the following programming problem for each $\lambda \in \Lambda$,

$$\begin{aligned} k(\lambda, \delta_0, T) &= \max_{\mathbf{v}, \alpha, x, \sigma^T(\alpha)} \lambda \cdot \mathbf{v} \\ &\text{s.t. } \mathbf{v} \text{ is generated by } (\sigma^T(\alpha), x) \\ 0 &\geq \sum_{i=1}^n \lambda_i \cdot x_i(m), \forall m \in S^T, \end{aligned} \quad (4)$$

¹¹This can be more easily seen from the following transformation: $x_i(m) = \frac{\delta^T(v_i(m) - v_i)}{1 - \delta^T}$, where v_i is player i 's discounted average payoff and $v_i(m)$ is player i 's discounted average continuation payoff given m .

This problem characterizes the best stationary T -period equilibrium payoffs of the T -period game in the direction of λ . Let $k(\lambda, T) = \lim_{\delta_0 \rightarrow 1} k(\lambda, \delta_0, T)$ and $D(\lambda, T) = \{\mathbf{v} | \lambda \cdot \mathbf{v} \leq k(\lambda, T)\}$.¹² Then it turns out that almost all payoffs in $\cap_{\lambda \in \Lambda} D(\lambda, T)$ can be supported by stationary T -public equilibria as $\delta \rightarrow 1$ in the original infinitely repeated game. The following proposition follows from Fudenberg and Levine [8] and Lemma 2 in Compte [6].

Proposition 1 *Suppose that $\cap_{\lambda \in \Lambda} D(\lambda, T)$ is full dimensional. For any smooth set W in the interior of $\cap_{\lambda \in \Lambda} D(\lambda, T)$, there exists $\underline{\delta}$ such that $W \subset E(\delta, T)$ for all $\delta \in (\underline{\delta}, 1)$*

Given this result, all I need is to show that $\cap_{\lambda \in \Lambda} D(\lambda, T)$ contains $\cap_{\lambda \in \Lambda} D_\tau(\lambda)$ as $T \rightarrow \infty$. Thus the second step is to prove the following lemma.

Lemma 3 *Suppose that Assumption 1 is satisfied. For any $\varepsilon > 0$ and $\tau > 0$, there exists δ' and T' such that $k(\lambda, \delta', T') > k_\tau(\lambda) - \varepsilon$ for all $\lambda \in \Lambda$ and $\delta \in (\delta', 1)$*

Proof. See the appendix. ■

Here I just provide a rough sketch of this second step. Fix $\lambda \in \Lambda$. It is easy to show that $D_\tau(\lambda) \subset D(\lambda, T)$ when $k_\tau(\lambda)$ is achieved by $\alpha \in NE$. There exists a T -period equilibrium with zero transfer where α is played in every period. So suppose that $k_\tau(\lambda)$ is achieved by a non-Nash pure strategy profile $a \in A_\tau(\lambda)$. Below I construct transfers $x_i, i \in I$, for which $\sigma^T(a)$ is a stationary T -period equilibrium and the efficiency loss is small ($E \left[\sum_{j \in I} \lambda_j x_j(m) | \sigma^T(a) \right] \geq -\varepsilon$).

Let player i be an informed player at a with $\lambda_i > \tau$. For $j \neq i$, since Assumption 1 is satisfied, there exists transfer x_j that provides the incentive for player j to play a_j (by Lemma 1). Define $x_j(m) = \frac{1}{T} \sum_{t=1}^T x_j(s_{-j,t})$. These transfers take care of the incentive of player $j \neq i$.

To keep “the budget balanced”, $\tilde{x}_{i,j}(m) = -\frac{\lambda_j}{\lambda_i} x_j(m)$ is added to player i 's transfer.¹³ Then player i may have incentive to deviate or send a false message to manipulate $x_j(m), j \neq i$. To address this problem, I follow Compte [6] to use the following scheme. Private signal profiles (s_1, \dots, s_T) reported in the end of the T periods are translated into Γ binary signals $c = (c_1, \dots, c_T) \in \{g, b\}^T$. The probability of $c_t = b$ at period t is given by some function $q_i(s_t)$. This is where a public randomization device is used. Then player i is punished if and only if $c_t = b$ for $t = 1, 2, \dots, T$. The sum of this punishment and $\sum_{j \neq i} \tilde{x}_{i,j}(m)$ is the total side transfer for player i .

¹²Note that $k(\lambda, \delta_0, T)$ is monotonically increasing in δ_0 .

¹³The role of $\tau > 0$ is to obtain a lower bound for λ_i , which provides an upper bound for these transfers.

Since player i can be secretly screened, this q_i can satisfy (1) and (2). The first condition (1) guarantees that any deviation by player i increases the probability of b in that period.

I like to make the expected probability of punishment as small as possible to avoid efficiency losses. To meet this goal, first I single out the binding incentive constraint out of many incentive constraints. If player i can learn the likelihood of her punishment from her private information, then the binding incentive constraint can be with respect to a very complicated contingent deviation. The second condition (2) is useful here to exclude this possibility: such learning does not occur in equilibrium. Then it turns out that the binding constraint is with respect to the one-period deviation in the first period as in Abreu, Milgrom and Pearce [1]. Note that the deviation gain from such one-period deviation is in the order of $\frac{1}{T}$ compared to the total payoff of the T -period game. Since the expected cost of punishment ($E[x_i(m)] < 0$) is in the same order, it becomes negligible as $T \rightarrow \infty$. This is why $E\left[\sum_{j \in I} \lambda_j x_j(m) \mid \sigma^T(a)\right]$ can be made arbitrarily small and $k_\tau(\boldsymbol{\lambda})$ is approximately achieved for the problem (4) by a stationary T -period equilibrium as $T \rightarrow \infty$.

5 Discussion

The above theorem may not apply to the two-player example in Section 3. Suppose that the stage game payoff is given by the following Prisoner's dilemma.

	C	D
C	1,1	-1,2
D	2,-1	0,0

Since player 1 is not an informed player at (D, C) , $A_\tau(\boldsymbol{\lambda})$ does not contain (D, C) when $\lambda_1 > 0$ and $\lambda_2 < \tau$. Hence $D_\tau(\boldsymbol{\lambda})$ needs to be defined with respect to the Nash equilibrium (D, D) (i.e. $D_\tau(\boldsymbol{\lambda}) \equiv \{(v_1, v_2) \mid \boldsymbol{\lambda} \cdot \mathbf{v} \leq 0\}$) for such $\boldsymbol{\lambda}$. Then clearly $\cap_{\boldsymbol{\lambda} \in \Lambda} D_\tau(\boldsymbol{\lambda})$ does not cover the individually rational and feasible payoff set (in fact it coincides with $(0, 0)$!).

However, the folk theorem can be still proved for this game by extending Theorem 1. It is possible to support (D, C) in the following way. Since D is the best response at (D, C) , player 1's incentive is satisfied without any transfer. Player 2's incentive to cooperate can be provided through rewards when $\lambda_2 < 0$.¹⁴ Remember that transfers must satisfy $0 \geq \sum_{i=1}^2 \lambda_i \cdot x_i(m)$, hence $x_2(m)$ must be positive given $x_1(m) = 0$ and $\lambda_2 < 0$.

How large can $k(\boldsymbol{\lambda}, \delta_0, T)$ be when $\lambda_1 > 0$ and $\lambda_2 < 0$? Since $x_2(m) \geq 0$ for any m , player 2's payoff is at least as large as $\max_{a_2} g_2(D, a_2) = 0$. Thus $k(\boldsymbol{\lambda}, \delta_0, T)$

¹⁴For $\lambda_2 \in [0, \tau)$, player 2's incentive is still provided through punishments. The lower bound of positive λ_2 was used to limit the size of transfers (from player 1 to player 2, in this case) in the proof of Theorem 1. This bound is not relevant in this case because no transfer from player 1 is needed.

can be at most $\lambda_1 g_1(D, C) + \lambda_2 g_2(D, D) = 2\lambda_1$. The lemma below proves that it is possible to approximate this level of k by making the expected reward given (D, D) almost negligible as $T \rightarrow \infty$. Then $D(\boldsymbol{\lambda}) = \{\mathbf{v} | \boldsymbol{\lambda} \cdot \mathbf{v} \leq 2\lambda_1\}$ for such $\boldsymbol{\lambda}$, which still covers all the individually rational and feasible payoffs, can be approximated when $\lambda_1 \geq 0, \lambda_2 < 0$.

More generally, let $A' \subset A$ be the subset of action profiles where every player is either informed or playing a best response action. For each $a \in A'$ and $\boldsymbol{\lambda} \in \Lambda$, let $I(a, \boldsymbol{\lambda})$ be the set of informed players with a negative weight. Define $k'(a, \boldsymbol{\lambda}) \equiv \sum_{i \in I/I(a, \boldsymbol{\lambda})} \lambda_i g_i(a) + \sum_{i \in I(a, \boldsymbol{\lambda})} \lambda_i \max_{a_i} g_i(a_i, a_{-i})$. Then $k'(a, \boldsymbol{\lambda})$ can be approximated by $k(\boldsymbol{\lambda}, \delta_0, T)$ as $T \rightarrow \infty$ and $\delta_0 \rightarrow 1$.¹⁵

Lemma 4 *For any $\varepsilon > 0$, there exists δ' and T' such that $k(\boldsymbol{\lambda}, \delta, T') > k'(a, \boldsymbol{\lambda}) - \varepsilon$ for every $a \in A'$ and every $\boldsymbol{\lambda} \in \Lambda$ and $\delta \in (\delta', 1)$*

Proof. See the appendix. ■

This means that Theorem 1 holds even if $D_\tau(\boldsymbol{\lambda}) \equiv \{\mathbf{v} \in \mathfrak{R}^n | \boldsymbol{\lambda} \cdot \mathbf{v} \leq k_\tau(\boldsymbol{\lambda})\}$ is modified by redefining $k_\tau(\boldsymbol{\lambda})$ as follows:

$$k_\tau(\boldsymbol{\lambda}) \equiv \max \left\{ \max_{a \in A_\tau(\boldsymbol{\lambda})} \boldsymbol{\lambda} \cdot \mathbf{g}(a), \max_{\alpha \in NE} \boldsymbol{\lambda} \cdot \mathbf{g}(a), \max_{\alpha \in A'} k'(a, \boldsymbol{\lambda}) \right\}.$$

This folk theorem applies to the above prisoners' dilemma example.

The proof of this lemma is slightly more complicated than the proof of Lemma 3 at one point. For Lemma 3, the binding constraint was with respect to the one shot deviation in the first period. This was because the marginal increase of the probability of punishment increases as players deviate more from the equilibrium action. When rewards are being used, the marginal decrease of the probability of reward goes down as players deviate more. Therefore the most tempting deviation is not the one shot deviation. In fact, it may be a very complicated deviation because of learning.¹⁶ The proof of Lemma 4 involves finding an upper bound of expected rewards without identifying the binding incentive constraint.

6 Appendix

Proof of Theorem 1

Given Proposition 1, I just need to show that $\cap_{\boldsymbol{\lambda} \in \Lambda} D(\boldsymbol{\lambda}, T)$ contains $\cap_{\boldsymbol{\lambda} \in \Lambda} D_\tau(\boldsymbol{\lambda})$ as $T \rightarrow \infty$ uniformly in $\boldsymbol{\lambda} \in \Lambda$ when Assumption 1 is satisfied. The following lemma proves this step.

¹⁵See Compte [6] for a more detailed discussion for this expression.

¹⁶When no learning occurs, the binding constraint is with respect to deviations in every period (Compte [6]).

Proof of Lemma 3

Proof. Pick any $\lambda \in \Lambda$. If $k_\tau(\lambda) = \lambda \cdot \mathbf{g}(\alpha)$ for some $\alpha \in NE$ (which is necessarily the case when every element of λ is nonpositive), then $k_\tau(\lambda)$ is trivially achieved by a stationary T -period equilibrium with constant action α and 0 side transfer. Let Λ_τ be the set of λ for which $k_\tau(\lambda)$ is achieved by a non-Nash pure strategy profile $a^\lambda \in A_\tau(\lambda)$. Let $i \in I$ be the informed player with $\lambda_i > \tau$ at a^λ . I am going to find side transfers and a stationary T -period equilibrium where players always play a^λ and announce their private signals truthfully.

To verify if a stationary T -period strategy is a stationary T -period equilibrium, I just need to check *all one-period deviation constraints (in action and/or announcements) on and off the equilibrium path*.¹⁷ Let's start with $j \neq i$. For given a^λ , let $x_j : S_{-j} \rightarrow \mathfrak{R}$ be the function described in Lemma 1. Define player j 's transfer by $x_j(m) = \frac{1}{T} \sum_{t=1}^T x_j(s_{-j,t})$. To avoid efficiency losses, $\frac{1}{T} \sum_{t=1}^T \frac{\lambda_j}{\lambda_i} x_j(s_{-j,t})$ is subtracted from player i 's transfer. For any T' , there exists δ' such that the following condition is satisfied for every $a_j \neq a_j^\lambda$ and $\delta \in (\delta', 1)$:

$$\frac{(1-\delta)}{1-\delta^T} \left(g_j(a_j, a_{-j}^\lambda) - g_j(a^\lambda) \right) < \frac{1}{T} \left\{ E \left[x_j(s_{-j}) | a^\lambda \right] - E \left[x_j(s_{-j}) | a_j, a_{-j}^\lambda \right] \right\}.$$

because $\frac{(1-\delta)}{1-\delta^T} \rightarrow \frac{1}{T}$ as $\delta \rightarrow 1$. This condition implies that player j 's incentive constraint is satisfied period by period on and off the equilibrium path when $x_j(m)$ is used as side transfers for player j in the end of T periods. Furthermore, this $\delta'(T)$ can be chosen uniformly with respect to $\lambda \in \Lambda_\tau$ because there are only finite number of action profiles and finite number of players.

Next I focus on the incentive of player i . First, the following expression characterizes the maximum deviation gain of player i in each period, taking into account the effect of the transfer from j to i defined above:

$$\max_{a_i, \rho} \left\{ \frac{(1-\delta)}{1-\delta^T} g_i(a_i, a_{-i}^\lambda) - \frac{1}{T} \sum_{j \neq i} \frac{\lambda_j}{\lambda_i} E \left[x_j(\rho(s_i), s_{-i}) | a_i, a_{-i}^\lambda \right] \right\} \quad (5)$$

$$- \left\{ \frac{(1-\delta)}{1-\delta^T} g_i(a^\lambda) - \frac{1}{T} \sum_{j \neq i} \frac{\lambda_j}{\lambda_i} E \left[x_j(s) | a^\lambda \right] \right\}.$$

Let $\bar{g}(T)$ be an upper bound of (5) across all $\lambda \in \Lambda_\tau$ and $\delta \in (\delta'(T), 1)$. Such bound can be found because λ_i is bounded below by $\tau > 0$. Note also that, $\bar{g}(T)$

¹⁷This does not guarantee that the strategy is a sequential equilibrium because it may not be sequentially rational for a player to announce her signal truthfully immediately after deviating from the equilibrium action. However, there always exists an outcome equivalent sequential equilibrium for any Nash equilibrium because of the full support assumption.

decreases roughly in the order of $\frac{1}{T}$ by taking $\delta'(T)$ large enough for each T so that $\frac{(1-\delta'(T))}{1-\delta'(T)^T} \approx \frac{1}{T}$.

Player i can be secretly screened by Lemma 2. Let $q_i : S \rightarrow (0, 1)$ be the function to satisfy (1) and (2). Define a random variable c as explained in Section 4. Then define $\tilde{x}_i(m)$ as follows

$$\begin{aligned}\tilde{x}_i(m) &= -\Delta \text{ when } c(m) = (1, \dots, 1) \\ &= 0 \quad \text{otherwise}\end{aligned}$$

where $\Delta > 0$ is to be chosen later. Player i 's total transfer is defined by the sum of $\tilde{x}_i(m)$ and $-\sum_{j \neq i} \frac{\lambda_j}{\lambda_i} x_j(m_{-j})$.

By Lemma 2, when player i deviates in action or in announcement in period t , the probability of her punishment increases by at least

$$\begin{aligned}& \min_{(a_i, \rho) \neq (a_i^\lambda, \rho^*)} \left(E \left[q_i(\rho(s_i), s_{-i}) \mid (a_i, a_{-i}^\lambda) \right] - \pi \right) \\ & \times \Pr(c_{t'} \text{ is } b \text{ at } t' \neq t \mid h_t, h_{i,t}) \\ & \geq \min_{(a_i, \rho_i) \neq (a_i^\lambda, \rho_i^*)} \left(E \left[q_i(\rho_i(s_i), s_{-i}) \mid (a_i, a_{-i}^\lambda) \right] - \pi \right) \times \pi^{T-1} \\ & \geq \Delta \pi \times \pi^{T-1}\end{aligned}$$

where ρ^* is the truth-telling strategy and $\Delta \pi$ is a lower bound of

$$\min_{(a_i, \rho_i) \neq (a_i^\lambda, \rho_i^*)} \left(E \left[q_i(\rho_i(s_i), s_{-i}) \mid (a_i, a_{-i}^\lambda) \right] - \pi \right)$$

with respect to all $\lambda \in \Lambda_\tau$. This is strictly positive by Lemma 2. Note that (1) and (2) guarantee that $\Pr(c_{t'} \text{ is } b \text{ at } t' \neq t \mid h_t, h_{i,t})$ is at least π^{T-1} in the beginning of any period t on and off the equilibrium path.

Since the gain from any one-period deviation (in action and/or announcement) is at most $\bar{g}(T)$, all the one-period deviation constraints on and off the equilibrium path are satisfied if

$$\bar{g}(T) \leq \Delta \pi \times \pi^{T-1} \times \Delta$$

is satisfied. Set Δ so that this inequality holds with equality. Since $x_j(m)$, $j \neq i$ are cancelled out by $-\sum_{j \neq i} \frac{\lambda_j}{\lambda_i} x_j(m)$, $\sum_{j=1}^n \lambda_j E[x_j(m) \mid \sigma^T(a^\lambda)]$ - the expected efficiency loss on the equilibrium path - is $\lambda_i \pi^T \Delta$, which is at most $\bar{g}(T) \frac{\pi}{\Delta \pi}$ from the definition of Δ (note that $\forall i, |\lambda_i| \leq 1$). This efficiency loss can be made smaller than ε because $\bar{g}(T)$ can be made arbitrarily small by choosing large enough T' (and large enough δ'). Notice that the same T' and δ' is being used for all $\lambda \in \Lambda_\tau$.

Hence, given $\varepsilon > 0$ and $\tau > 0$, there exist T' and δ' such that, for every $\lambda \in \Lambda_\tau$,

there exists a T-period stationary equilibrium that satisfies

$$\begin{aligned}
& \sum_{j=1}^n \lambda_j \left\{ g_j(a^\lambda) + E \left[x_j(m) \mid \sigma^{T'}(a^\lambda) \right] \right\} \\
& \geq k_\tau(\lambda) - \bar{g}(T') \frac{\pi}{\Delta\pi} \\
& > k_\tau(\lambda) - \varepsilon.
\end{aligned}$$

Therefore $k(\lambda, \delta, T') \geq k(\lambda, \delta', T') > k_\tau(\lambda) - \varepsilon$ for all $\delta \in (\delta', 1)$. ■

This Lemma proves that $\cap_{\lambda \in \Lambda} D(\lambda, T)$ contains $\cap_{\lambda \in \Lambda} D_\tau(\lambda)$ as $T \rightarrow \infty$ uniformly in $\lambda \in \Lambda$, thus W is contained in the interior of $\cap_{\lambda \in \Lambda} D(\lambda, T)$ for large enough T . Fix such T . Then Proposition 1 implies that there exists $\underline{\delta}$ such that W is contained in $E(\delta, T)$ for any $\delta \in (\underline{\delta}, 1)$. ■

Proof of Lemma 4

Proof. Fix $\varepsilon > 0$. Note that (1) best response actions can be supported without any transfer and (2) the proof of Lemma 3 can be applied to informed players with nonnegative weight. For (2), notice that no lower bound of the positive weight is needed here because no transfer across different players is involved. Hence I focus on the incentive of informed players with negative weight in the following.

Pick any $a \in A'$, $\lambda \in \Lambda$ and pick any $i \in I(a, \lambda)$. Consider a T-period game with side transfers and a stationary T-period strategy where a is played for T-periods. Since player i can be secretly screened by Lemma 2, there exists a function $q_i : S \rightarrow (0, 1)$ that satisfies (1) and (2). As before, define a random variable c as explained in Section 4. This time player i 's transfer is defined by

$$\begin{aligned}
x_i(m) &= \Delta \text{ when } c(m) = (g, \dots, g) \\
&= 0 \quad \text{otherwise}
\end{aligned}$$

If $\Delta > 0$ is large enough, it is optimal to play a_i every period. Δ is chosen so that one incentive constraint is binding. Let $\sigma^{T'}$ be the strategy with the binding incentive constraint. Let $p(h_i^T \mid \sigma^{T'})$ be the distribution on $A_i^T \times S_i^T$ that is generated when player i deviates to $\sigma_i^{T'}$ unilaterally and $d(h_i^T)$ be the number of times player i deviated from a_i along T-period history h_i^T . Then player i 's binding incentive constraint implies

$$\begin{aligned}
& \sum_{h_i^T} \sup_{\delta \in [\delta_0, 1]} \frac{(1 - \delta) \sum_{k=1}^{d(h_i^T)} \delta^{k-1} \bar{g}_i(a)}{1 - \delta^T} p(h_i^T \mid \sigma^{T'}) \\
& \geq \left[\Pr(c(m) = (g, \dots, g) \mid \sigma(a)) - \Pr(c(m) = (g, \dots, g) \mid \sigma_i^{T'}, \sigma_{-i}(a)) \right] \Delta
\end{aligned} \tag{6}$$

where $\bar{g}_i(a) = \max_{a'_i} g_i(a'_i, a_{-i}) - g_i(a)$. The left hand side is larger because it is computed as if $d(h_i^T)$ deviations happened in the first $d(h_i^T)$ periods.

Since

$$\begin{aligned}
& \Pr(c(m) = (g, \dots, g) | \sigma(a)) - \Pr(c(m) = (g, \dots, g) | \sigma_i^{T'}, \sigma_{-i}(a)) \quad (7) \\
&= \prod_{t=1}^T \Pr(c_t(m) = g | \sigma(a)) - \sum_{h_i^T} \prod_{t=1}^T \Pr(c_t(m) = g | \sigma_i^{T'}, \sigma_{-i}(a), h_i^T) p(h_i^T | \sigma_i^{T'}) \\
&\geq (1 - \pi)^T - \sum_{h_i^T} (1 - \pi)^{T-d(h_i^T)} (1 - \pi')^{d(h_i^T)} p(h_i^T | \sigma_i^{T'}) \\
&= \Pr(c(m) = (g, \dots, g) | \sigma(a)) \left(1 - \sum_{h_i^T} \left(\frac{1 - \pi'}{1 - \pi} \right)^{d(h_i^T)} p(h_i^T | \sigma_i^{T'}) \right)
\end{aligned}$$

where $\pi = E[q_i(s) | a]$ and π' satisfies $\pi' > \pi$ and $\pi' \leq E[q_i(s'_i, s_{-i}) | a'_i, a_{-i}, s_i]$ for any $(s'_i, a'_i) \neq (s_i, a_i)$.

By (6) and (7), player i 's average payoff $g_i(a) + \Pr(c(m) = (g, \dots, g) | \sigma(a)) \Delta$ is bounded above by

$$\begin{aligned}
& g_i(a) + \frac{\sum_{h_i^T} \sup_{\delta \in [\delta_0, 1]} \frac{(1-\delta) \sum_{k=1}^{d(h_i^T)} \delta^{k-1} \bar{g}_i(a)}{1-\delta^T} p(h_i^T | \sigma^{T'})}{1 - \sum_{h_i^T} \left(\frac{1-\pi'}{1-\pi} \right)^{d(h_i^T)} p(h_i^T | \sigma_i^{T'})} \\
&\leq g_i(a) + \max_{d \in \{1, \dots, T\}} \frac{\frac{(1-\delta_0) \sum_{k=1}^d \delta_0^{k-1} \bar{g}_i(a)}{1-\delta_0^T}}{1 - \left(\frac{1-\pi'}{1-\pi} \right)^d}.^{18}
\end{aligned}$$

For each T , the second term converges to

$$\max_{d \in \{1, \dots, T\}} \frac{\frac{d}{T} \bar{g}_i(a)}{1 - \left(\frac{1-\pi'}{1-\pi} \right)^d}$$

as $\delta_0 \rightarrow 1$ by the theorem of Maximum. It is easy to show that this is maximized at $d = T$. Therefore I can find large enough (T', δ') such that this bound is approximately $g_i(a) + \bar{g}_i(a) = \max_{a'_i} g_i(a'_i, a_{-i})$. Clearly this approximation can be done uniformly across every $i \in I(a, \lambda)$, $a \in A'$, and $\lambda \in \Lambda$.

Hence there exists δ' and T' such that

$$k'(a, \lambda) = \sum_{i \in I'(a, \lambda)} \lambda_i g_i(a) + \sum_{i \in I'(a, \lambda)} \lambda_i \max_{a_i} g_i(a_i, a_{-i})$$

is approximated within ε by a T-stationary equilibrium for every $a \in A'$ and $\lambda \in \Lambda$ in the T' -period repeated game with side transfers. Therefore $k(\lambda, \delta, T') \geq k(\lambda, \delta', T') > k'(a, \lambda) - \varepsilon$ for every $a \in A'$, $\lambda \in \Lambda$, and \int ■

References

- [1] D. Abreu, P. Milgrom, and D. Pearce. Information and timing in repeated partnerships. *Econometrica*, 59(6):1713–1733, 1991.
- [2] G. Ashkenazi-Golan. Confession and pardon in repeated games with communication. 2004.
- [3] R. J. Aumann, M. Maschler, and R. Stearns. Repeated games of incomplete information: An approach to the non-zero sum case. In R. J. Aumann and M. Maschler, editors, *Repeated Games with Incomplete Information*. MIT Press, 1995. (originally published in 1968).
- [4] E. Ben-Porath and M. Kahneman. Communication in repeated games with private monitoring. *Journal of Economic Theory*, 70:281–297, 1996.
- [5] V. Bhaskar and I. Obara. Belief-based equilibria in the prisoners’ dilemma with private monitoring. *Journal of Economic Theory*, 102:16–39, 2002.
- [6] O. Compte. Communication in repeated games with imperfect private monitoring. *Econometrica*, 66:597–626, 1998.
- [7] J. C. Ely and J. Välimäki. A robust folk theorem for the prisoner’s dilemma. *Journal of Economic Theory*, 102(1):84–105, 2002.
- [8] D. Fudenberg and D. Levine. Efficiency and observability with long -run and short-run players. *Journal of Economic Theory*, 62:103–135, 1994.
- [9] D. Fudenberg and D. Levine. The Nash threats folk theorem with communication and approximate common knowledge in two player games. 2004.
- [10] J. Hörner and W. Olszewski. The folk theorem for games with private almost-perfect monitoring. mimeo, 2004.
- [11] M. Kandori and H. Matsushima. Private observation, communication and collusion. *Econometrica*, 66:627–652, 1998.
- [12] G. J. Mailath and S. Morris. Repeated games with almost-public monitoring. *Journal of Economic Theory*, 102(1):189–228, 2002.
- [13] G. J. Mailath and S. Morris. Coordination failure in repeated games with almost-public monitoring. 2004.
- [14] H. Matsushima. Repeated games with private monitoring: Two players. *Econometrica*, 72:823–852, 2004.
- [15] R. McLean, I. Obara, and A. Postlewaite. Informational smallness and private monitoring in repeated games. mimeo UCLA, 2005.