# *CARESS Working Paper 98-07*
# Repeated Games with Imperfect Private Monitoring: Notes on a Coordination Perspective*

George J. Mailath
Department of Economics
University of Pennsylvania
3718 Locust Walk
Philadelphia, PA 19104 USA
gmailath@econ.sas.upenn.edu

Stephen Morris
Cowles Foundation
Yale University
30 Hillhouse Avenue
New Haven, CT 06520 USA
smorris@econ.yale.edu

July 3, 1998

**Abstract**

In repeated games with imperfect public monitoring, players can use public signals to perfectly coordinate their behavior. Our study of repeated games with imperfect *private* monitoring focuses on the coordination problem that arises without public signals. We present three new observations. First, in a simple twice repeated game, we characterize the private signalling technologies that allow non-static Nash behavior in pure strategy equilibria. Our characterization uses the language of common $p$-belief due to Monderer and Samet (GEB, 1989). Second, we show that in the continuum action convention game of Shin and Williamson (GEB, 1996), for *any* full support private monitoring technology, equilibria of the finitely repeated convention game must involve only static Nash equilibria. By contrast, with sufficiently informative public monitoring, the multiplicity of Nash equilibria allows a finite folk theorem. Finally, for finite action games, we prove that there are full support private monitoring technologies for which a Nash reversion infinite horizon folk theorem holds.

**key words:** repeated games, imperfect monitoring, coordination, punishments

*This draft was written while Morris was visiting Northwestern University and Mailath was visiting the Australian National University. We thank both institutions for their hospitality.

# Repeated Games with Imperfect Private Monitoring: Notes on a Coordination Perspective

by George J. Mailath and Stephen Morris

## 1. Introduction

Repeated games with imperfect *public* monitoring are well-understood. When public signals provide information about past actions, punishments contingent on public signals provide dynamic incentives to choose actions that are not static best responses (see Green and Porter [12] and Abreu, Pearce, and Stacchetti [1]). Moreover, if the public signals satisfy an identifiability condition, a folk theorem holds: if the discount rate is sufficiently close to one, any individually rational payoff can be supported as the average payoff of an equilibrium of the repeated game (Fudenberg, Levine, and Maskin [11]). Perfect public equilibria of games with imperfect public monitoring have a recursive structure that greatly simplifies their analysis (and plays a central role in Abreu, Pearce, and Stacchetti [1] and Fudenberg, Levine, and Maskin [11]).[1] In particular, any perfect public equilibrium can be described by an action profile for the current period and continuation values that are necessarily equilibrium values of the repeated game. However, for this recursive structure to hold, all players must be able to coordinate their behavior after any history (i.e., play an equilibrium after any history). If the relevant histories are public, then this coordination is clearly feasible.

Repeated games with imperfect *private* monitoring have proved less tractable. Since the relevant histories are typically private, equilibria do not have a simple recursive structure.[2] Consider the following apparently ideal setting for supporting non-static Nash behavior. There exist "punishment" strategies with the property that all players have a best response to punish if they know that others are punishing; and private signals provide extremely accurate information about past play, so that punishment strategies contingent on those signals provide the requisite dynamic incentives to support action profiles that are not static Nash. Even in these circumstances, there is no guarantee that non-static Nash behavior can be supported in equilibrium. The problem is that even when one player is almost sure that another has deviated and would want to punish if he believed that

[1] A strategy is public if it only depends on the pubic history, and a perfect public equilibrium is a profile of public strategies that induces a Nash equilibrium after every public history.

[2] Amarante [3] provides a large state space recursive characterization of the equilibrium set of repeated games with imperfect private monitoring.

**Stage 1**

| | Player 2: $L$ | Player 2: $R$ |
|---|---|---|
| | $1,1$ | $4,0$ |

**Stage 2**

| | Player 2: $G$ | Player 2: $B$ |
|---|---|---|
| Player 1: $G$ | $3,3$ | $0,0$ |
| Player 1: $B$ | $0,0$ | $1,1$ |

Figure 1.1: A two stage game.

others were punishing, he is not sure that others are almost sure that someone has deviated. With private signals, unlike public signals, there is not common knowledge of the histories that trigger punishments. Common knowledge is a sufficient condition for coordinating behavior. But is it a necessary condition?[3]

To illustrate the issues we explore, consider the two stage game in Figure 1.1. If player 1 observes player 2's choice (i.e., there is perfect monitoring), there is a subgame perfect equilibrium in which player 2 plays $R$ in the first stage. Now suppose player 2's choice is not observed, but there is a public signal $\omega \in \{l, r\}$, with the signal correctly revealing the action taken with probability $1-\varepsilon$. Consider the public profile: player 2 plays $R$ in the first stage, and $(G, G)$ is played if $r$ is observed and $(B, B)$ is played otherwise. This will be an equilibrium if $\varepsilon \leq 1/4$. Note that independent of the signal observed, the profile specifies a Nash equilibrium of the continuation game. This is just a reflection of the recursive structure of perfect public equilibria discussed above. Note also that in equilibrium, both players play $B$ after $l$ only because they expect the other player to do so.[4] Now suppose that the signal about player 2's stage 1 choice is *private*: only player 1 observes the signal $\omega_1 \in \{l, r\}$. The only stage 1 action consistent with pure strategy equilibria in this case is $L$. Since player 2 does not observe the signal, her stage 2 behavior must be independent of it. Since we are considering pure strategies, her second period choice is thus deterministic and independent of the signal. But this means the sequentially-rational choice for player 1 is also

[3] An alternative solution to the co-ordination problem is to allow players to (publicly) communicate, thus automatically generating the required common knowledge histories (see Compte [9], Kandori and Matsushima [16], and Aoyagi [4]). Although the coordination problem is automatically solved by public communication, a new set of problems arises from players' truth-telling contraints.

[4] Player 2 is not being punished for deviating since, in equilibrium, the signal $l$ is observed with positive probability. Player 1 assigns probability 1 to the event that player 2 had followed the equilibrium specification of $R$ after observing $l$. Of course, if $(B, B)$ is not specified after $l$, player 2 has no incentive to play $R$.

independent of the signal, and so player 2 has no incentive to play $R$.[5] Finally, suppose that player 2 also received a private signal, $\omega_2 \in \{l, r\}$, correlated with player 1's private signal (player 2 of course knows her own stage 1 action, but additional information about player 1's signal is valuable). The second stage is now a game of incomplete information, with different values of $\omega_i$ corresponding to different types (the game is degenerate in that payoffs are independent of type). Clearly, there is an equilibrium of this incomplete information game where all types of each player choose $B$ and another one where all types choose $G$. Neither of these will support $R$ in the first stage. But are there also equilibria where both actions $G$ and $B$ are each chosen by some types? In this case, we could say there is *contingent coordination*: players coordinate their behavior, contingent on payoff-irrelevant types, *despite* a lack of common knowledge. It would clearly be sufficient if the signals $\omega_1$ and $\omega_2$ were almost perfectly correlated (i.e., if the signal structure was almost public).

The literature on higher-order beliefs has explored the role of common knowledge in coordination in static games of incomplete information (Monderer and Samet [18], Morris, Rob and Shin [19], Kajii and Morris [14]; see Kajii and Morris [13] for a survey). A basic insight from that literature can be illustrated by modifying our earlier example. Ignore the first stage and suppose each player $i$ observes a payoff-irrelevant signal, $\omega_i$, from some finite set, $\Omega_i$, before playing the second stage. Write $\pi$ for the joint probability distribution on $\Omega_1 \times \Omega_2$ and assume it has full support. Thus the two players are engaged in a incomplete information game. There exists a *pure strategy* equilibrium where both $G$ and $B$ are chosen if and only if there exist two disjoint subsets of the state space $\Omega_1 \times \Omega_2$ that can be approximately commonly known (in the sense of Monderer and Samet [18]). More precisely, fix such an equilibrium and write $\Omega_i^G$ for the set of types of player $i$ who choose action $G$ and $\Omega_i^B$ for the set $\Omega_i \backslash \Omega_i^G$, i.e., the set of types who choose action $B$. A necessary condition for equilibrium is that each type $\omega_1 \in \Omega_1^G$ believe (under $\pi$) with probability at least $\frac{1}{4}$ that $\omega_2 \in \Omega_2^G$, and vice versa. If this condition holds, the event $\Omega_1^G \times \Omega_2^G$ is said to be "$\frac{1}{4}$-evident": *whenever* the event occurs, both players believe it with probability at least $\frac{1}{4}$. This implies that the event $\Omega_1^G \times \Omega_2^G$ is "common $\frac{1}{4}$-belief" at all states in $\Omega_1^G \times \Omega_2^G$: Monderer and Samet [18] showed that $\Omega_1^G \times \Omega_2^G$ is exactly the set of states where both players believe $\Omega_1^G \times \Omega_2^G$ with probability at least $\frac{1}{4}$, both players believe with probability at least $\frac{1}{4}$ that both believe $\Omega_1^G \times \Omega_2^G$ with probability at least $\frac{1}{4}$, and so on ad

[5]The restriction to pure strategies is important. Suppose the payoff from the profile $(G, G)$ is $k$ for each player, rather than 3. If $k \in (2, 2\frac{1}{2}]$ and $\varepsilon = (k-2)/(k-1)$, there is a mixed strategy equilbrium in which player 2 chooses $L$ with probability $\alpha \in [\varepsilon, 2(1-\varepsilon)/(2-\varepsilon)]$. We return to the issue of mixed strategies in Section 2.

infinitum. An analogous argument shows that $\Omega_1^B \times \Omega_2^B$ is $\frac{3}{4}$-evident. Observe that this condition is quite stringent: there must be *two, disjoint* events such that *whenever* either occurs, both players attach some minimal probability to it having occurred. If $\Omega_i^G$ and $\Omega_i^B$ are singletons, the condition is equivalent to requiring sufficiently correlated signals (as we described in the previous paragraph).

This example illustrates the type of problem that arises in repeated games with imperfect private monitoring. Players' types (at each date) are their private and payoff-irrelevant histories. There is not common knowledge of types. How much approximate common knowledge is required to achieve contingent coordination? This comparison suggests that we might want to look for a result of the following form:

> *A folk theorem holds for a repeated game with imperfect private monitoring if and only if there exist approximate common knowledge sets of private signals (i.e., "almost public" events) with the property that those almost public events are sufficiently informative* (in the sense of the imperfect public monitoring folk theorems).

Unfortunately complications arise in trying to make this vague conjecture precise. First, the connection between approximate common knowledge and equilibria coordinated on payoff irrelevant events can be made precise in examples (as we demonstrated above), but it is more complex in the continuation games at arbitrary histories of a repeated game. Mixed strategies present special problems, as illustrated by footnote 5 and their role in the imperfect public monitoring folk theorem (see Fudenberg, Levine, and Maskin [11, footnote 13]). Further complicating matters is the possibility of purification (see the end of Section 2). Second, in static games, the information structure is given by the (exogenous) probability distribution over types. In repeated games with imperfect private monitoring, the probability distribution over types at any date depends on not only the (exogenous) monitoring technology but also the (endogenous) strategies. Thus it will in general be hard to guarantee that a given property of the monitoring technology implies that the distribution over types does (or does not) entail enough approximate common knowledge to allow contingent coordination. Posterior distributions over types are especially hard to analyze with mixed strategies.

Nonetheless, we believe that this "coordination perspective" may be valuable in understanding repeated games with imperfect private monitoring. The purpose of this paper is to describe a number of positive (i.e., folk theorem type) results and negative (i.e., anti-folk theorem type) results in examples. We, like much of this literature, focus on noisy imperfect private monitoring. There is also a small literature on monitoring where only a subset of players perfectly observe the behavior of some player (see, for example, Ben-Porath [5] and Ahn [2]).

Section 2 considers the simplest possible two stage game in which we can explore the coordination perspective. In particular, we suppose that two players play the prisoner's dilemma in the first period and a coordination game in the second period (this is essentially a symmetric version of Figure 1.1). In this simple case, we are able to characterize completely pure strategy equilibria (for arbitrary monitoring technologies) and demonstrate that approximate common knowledge in the monitoring technology is both necessary and sufficient for co-operation in the first stage prisoner's dilemma. But we also demonstrate how mixed strategies in the first period can be used to generate the required correlation and thus approximate common knowledge of types in the second period, *even if the monitoring technology generates independent signals conditional on any pure action profile*. This section builds on Bhaskar and van Damme [1997] who demonstrated the role of correlated signals in supporting pure strategy equilibria in a similar context.

In section 3, we present an "anti-folk" theorem. Shin and Williamson [21] introduced and analyzed a simple continuous action "convention game," where there was a continuum of Nash equilibria but nonetheless it was impossible to coordinate different conventions contingent on any event that was not full common knowledge (common $p$-belief for *any* $p < 1$ is never sufficient). This is thus a very extreme game. But exploiting the coordination perspective outlined above, we are able to show that for any full support private monitoring technology, equilibria of the finitely repeated convention game must involve only static Nash equilibria. By contrast, with perfect monitoring or sufficiently informative public monitoring, the multiplicity of Nash equilibria allows a finite folk theorem.

This anti-folk theorem suggests that we cannot be sure to obtain positive (i.e., folk theorem) results even if we are allowed to pick any full support private monitoring technology. But in Section 4, we show that for a particular type of private monitoring technology, there is an infinite horizon folk theorem for *finite action* games. The finite action restriction is crucial as it allows a non-trivial degree of approximate common knowledge to be sufficient for coordination. Our private monitoring technology has players' signals highly correlated *whatever action profile was chosen in the previous period*. One interpretation of this monitoring technology is that we are in a world of almost (but not complete) imperfect public monitoring. We prove a folk theorem for two different strategy profiles, although both have a Nash reversion flavor. In Section 4.1, we consider strategy profiles where players' actions depend *only* on the private signal they observed in the previous period. Together with the highly correlated signal structure, this ensures

| | $D$ | $C$ |
|---|---|---|
| $D$ | $0,0$ | $x+1,-1$ |
| $C$ | $-1,x+1$ | $x,x$ |

Figure 2.1: A Prisoner's Dilemma

| | $G$ | $B$ |
|---|---|---|
| $G$ | $k,k$ | $0,0$ |
| $B$ | $0,0$ | $1,1$ |

Figure 2.2: A Coordination Game

that there is always approximate common knowledge of the actions other players will be taking at any history. In Section 4.2, we consider Nash reversion in which a particular Nash equilibrium "threat point," once reached is never left. Unlike the profile of Section 4.1, this profile does depend upon the entire history. However, as long as the correlation in the signals is sufficiently strong, the required approximate common knowledge still holds.

## 2. A Two Stage Example

In our first example, two players are involved in a two stage game. In the first period, they play the prisoner's dilemma in Figure 2.1,where $x > 0$. In the second period, they play the coordination game in Figure 2.2, where $k > 2$. There is no discounting. Under these assumptions, with perfect monitoring, it is possible to support $(C, C)$ in the first period with pure strategies.

Now consider the case with imperfect private monitoring. Before choosing his second period action, player $i$ observes a signal $\omega_i$ from a finite set $\Omega_i$ concerning the first period action profile; write $\pi((\omega_1, \omega_2)|a)$ for the positive probability that signals $(\omega_1, \omega_2)$ are observed when first period actions are $a \in \{C, D\}^2$. To keep the example simple, we assume $\pi$ is symmetric, i.e., $\pi((\omega_j, \omega_i)|(a_j, a_i)) = \pi((\omega_i, \omega_j)|(a_i, a_j))$.

When the players play according to a Nash equilibrium of this two stage game with imperfect private monitoring, the second stage can be viewed as a game of incomplete information, with player $i$ having $T_i \equiv \{C, D\} \times \Omega_i$ as his type space. Moreover, there is a joint distribution over this type space induced by first period behavior and $\pi$.

In order to characterize the critical properties of $\pi$, we introduce belief operators (see Monderer and Samet [18]). For any $E \subset \Omega_1 \times \Omega_2$, say that $i$ *$p$-believes* $E$ at $\omega = (\omega_1, \omega_2)$ if $\pi(E\,|\omega_i) \geq p$. The *belief operator* for player $i$ identifies the signals at which $i$ $p$-believes $E$, i.e.,

$$B_i^p(E; a) = \{\omega \in \Omega_1 \times \Omega_2 : \pi(E|\omega_i, a) \geq p\}.$$

The event $E$ is *$p$-evident (given $a$)* if $E \subset B_i^p(E; a)$ for $i = 1, 2$. Note that if $E$ is $p$-evident, then (since belief operators are monotonic, in the sense that $E' \subset E$ implies $B_i^p(E'; a) \subset B_i^p(E; a)$) $E \subset B_i^p(E; a) \subset B_i^p(B_j^p(E; a); a)$, i.e., $i$ also assigns a probability of at least $p$ to $j$ assigning probability of at least $p$ to $E$.

**Proposition 1.** *There exists a pure strategy equilibrium with cooperation in the first period if and only if each $\Omega_i$ can be partitioned into sets $\left\{\Omega_i^G, \Omega_i^B\right\}$ such that*

1. *$\Omega_1^G \times \Omega_2^G$ is $\frac{1}{1+k}$-evident (given $CC$),*
2. *$\Omega_1^B \times \Omega_2^B$ is $\frac{k}{1+k}$-evident (given $CC$), and*
3. $$k\pi\left\{\left(\Omega_1^G \times \Omega_2^G \middle| CC\right) - \pi\left(\Omega_1^G \times \Omega_2^G \middle| DC\right)\right\} + \left\{\pi\left(\Omega_1^B \times \Omega_2^B \middle| CC\right) - \pi\left(\Omega_1^B \times \Omega_2^B \middle| DC\right)\right\} \geq 1.$$

This follows almost immediately from the definition of equilibrium. Suppose there is a pure strategy equilibrium $(\widehat{s}_1, \widehat{s}_2)$ with

$$\widehat{s}_i^1 = C \text{ and } \widehat{s}_i^2\left(C, \omega_i\right) = \begin{cases} G, & \text{if } \omega_i \in \Omega_i^G, \\ B, & \text{if } \omega_i \in \Omega_i^B. \end{cases}$$

Properties [1] and [2] in the lemma follow from the requirement that second period strategies constitute a (Bayesian) Nash equilibrium of the second period game following $CC$ (this is a degenerate incomplete information game). Property [3] of the lemma is required to ensure that it is optimal to play $C$ in the first period.

The pure strategy restriction is key to the above analysis. The role of mixed strategies in this context has been explored by Bhaskar and van Damme [7].[6] They consider a twice repeated game with imperfect private monitoring. Our two stage game is essentially a stripped down version of their twice repeated game. They note that co-operation (i.e., an efficient but dominated action) is impossible in

[6] Kandori [15] shows how with private monitoring, mixed strategies may allow non-static Nash behavior even in finitely repeated games with a unique Nash equilibrium.

| | $Cc$ | $Cd$ | $Dc$ | $Dd$ |
|---|---|---|---|---|
| $Cc$ | $\alpha^2(1-\varepsilon)^2$ | $\alpha^2\varepsilon(1-\varepsilon)$ | $\alpha(1-\alpha)\varepsilon(1-\varepsilon)$ | $\alpha(1-\alpha)\varepsilon^2$ |
| $Cd$ | $\alpha^2\varepsilon(1-\varepsilon)$ | $\alpha^2\varepsilon^2$ | $\alpha(1-\alpha)(1-\varepsilon)^2$ | $\alpha(1-\alpha)\varepsilon(1-\varepsilon)$ |
| $Dc$ | $\alpha(1-\alpha)\varepsilon(1-\varepsilon)$ | $\alpha(1-\alpha)(1-\varepsilon)^2$ | $(1-\alpha)^2\varepsilon^2$ | $(1-\alpha)^2\varepsilon(1-\varepsilon)$ |
| $Dd$ | $\alpha(1-\alpha)\varepsilon^2$ | $\alpha(1-\alpha)\varepsilon(1-\varepsilon)$ | $(1-\alpha)^2\varepsilon(1-\varepsilon)$ | $(1-\alpha)^2(1-\varepsilon)^2$ |

Figure 2.3: The distribution over $T_1 \times T_2$ generated by mixing with probability $\alpha$ on $C$.

| | $Cc$ | $Cd$ | $Dc$ | $Dd$ |
|---|---|---|---|---|
| $Cc$ | $(1-\varepsilon)^2$ | $\varepsilon(1-\varepsilon)$ | 0 | 0 |
| $Cd$ | $\varepsilon(1-\varepsilon)$ | $\varepsilon^2$ | 0 | 0 |
| $Dc$ | 0 | 0 | 0 | 0 |
| $Dd$ | 0 | 0 | 0 | 0 |

Figure 2.4: The distribution if $\alpha = 1$.

pure strategy equilibria with independent signals, but note that co-operation is possible with correlated signals. Our Lemma 1 provides an exact description of how great the deviation from independence must be to allow co-operation.

Bhaskar and van Damme go on to show that mixed strategies allow co-operation in the first period, even with independent signals.[7] We now illustrate this point in our example. In doing so, we will note how mixed strategies generate the requisite approximate common knowledge even when the private monitoring technology generates no correlation and thus no approximate common knowledge of the signals.

Consider an independent monitoring technology, with $\Omega_i = \{c, d\}$, in which each player observes his opponent's action correctly with independent probability $1-\varepsilon$, and incorrectly with probability $\varepsilon$. We denote the signal of the opponent's action by a lower case letter. Thus,

$$\pi((\omega_1, \omega_2) \,|\, (a_1, a_2)) = \begin{cases} (1-\varepsilon)^2, & \text{if } \omega_1 = a_2 \text{ and } \omega_2 = a_1, \\ \varepsilon(1-\varepsilon), & \text{if } \omega_1 = a_2 \text{ and } \omega_2 \neq a_1, \\ \varepsilon(1-\varepsilon), & \text{if } \omega_1 \neq a_2 \text{ and } \omega_2 = a_1, \\ \varepsilon^2, & \text{if } \omega_1 \neq a_2 \text{ and } \omega_2 \neq a_1. \end{cases}$$

[7] The probability of co-operation is bounded away from one as the noise goes to zero. But the additional use of public sunspots allows co-operation with probability approaching one using mixed strategies, as the noise goes to zero.

| | $Cc$ | $Cd$ | $Dc$ | $Dd$ |
|---|---|---|---|---|
| $Cc$ | $\alpha^2$ | 0 | 0 | 0 |
| $Cd$ | 0 | 0 | $\alpha(1-\alpha)$ | 0 |
| $Dc$ | 0 | $\alpha(1-\alpha)$ | 0 | 0 |
| $Dd$ | 0 | 0 | 0 | $(1-\alpha)^2$ |

Figure 2.5: The distribution if $\varepsilon = 0$.

Suppose that each player co-operates with probability $\alpha$ in the first period, and defects with probability $1-\alpha$. The induced distribution over the type space $T_1 \times T_2$ is given in Figure 2.3. Setting $\alpha = 1$ gives the pure strategy outcome of first period cooperation, with an induced distribution over types given in Figure 2.4, which cannot give pure strategy contingent coordination (using Lemma 1). In particular, $\{(c,c)\}$ is $(1-\varepsilon)$-evident, while $\{(d,d)\}$ is only $\varepsilon$-evident. The difficulty, of course, is that specifying $B$ after a realization of $c$ (as would be required by condition [2]) removes any incentive to choose $C$ in the first stage. From the induced probability distribution on the type space $T_1 \times T_2$ we also see that for $\alpha = 1$, the types are independent.

On the other hand, mixed strategies *generate* correlated types. As $\varepsilon \to 0$ (holding $\alpha$ constant), the distribution over types tends to the distribution given in Figure 2.5. In other words, for *any* $p < 1$, $\{Cc\} \times \{Cc\}$ and $\{Cd, Dc, Dd\} \times \{Cd, Dc, Dd\}$ are both $p$-evident sets for $\varepsilon$ small. Thus, for small $\varepsilon$ and mixed first period strategies, there is no problem coordinating punishments in the second period.

It is straightforward to construct symmetric mixed strategy equilibria using the above insight. Consider the following mixed strategies:

$$\sigma_i^1\left[a_i^1\right] = \begin{cases} \alpha, & \text{if } a_i^1 = C, \\ 1-\alpha, & \text{if } a_i^1 = D, \end{cases}$$

and

$$\sigma_i^2\left(a_i^1, \omega_i\right)\left[a_i^2\right] = \begin{cases} 1, & \text{if } a_i^2 = G \text{ and } a_i^1 = \omega_i = C, \\ 1, & \text{if } a_i^2 = B \text{ and } \textit{either } a_i^1 = D \textit{ or } \omega_i = D, \\ 0, & \text{otherwise.} \end{cases} .$$

Second period optimality requires

$$\frac{\alpha(1-\varepsilon)^2}{\alpha(1-\varepsilon) + (1-\alpha)\varepsilon} \geq \frac{1}{1+k};$$

$$\begin{aligned}\frac{\alpha\varepsilon^2+(1-\alpha)(1-\varepsilon)}{\alpha\varepsilon+(1-\alpha)(1-\varepsilon)} &\geq \frac{k}{1+k}; \\ \frac{\alpha(1-\varepsilon)^2+(1-\alpha)\varepsilon}{\alpha(1-\varepsilon)+(1-\alpha)\varepsilon} &\geq \frac{k}{1+k}; \\ \text{and } \frac{\alpha\varepsilon(1-\varepsilon)+(1-\alpha)(1-\varepsilon)}{\alpha\varepsilon+(1-\alpha)(1-\varepsilon)} &\geq \frac{k}{1+k}.\end{aligned} \tag{2.1}$$

For the mixed strategy to be optimal in the first period, we must have the payoff to co-operating,

$$\alpha\left\{x+(1-\varepsilon)^2\cdot k+\varepsilon^2\cdot 1\right\}+(1-\alpha)\{-1+(1-\varepsilon)\cdot 1\}=\alpha x+\alpha(1-\varepsilon)^2k+\alpha\varepsilon^2-(1-\alpha)\varepsilon,$$

equal to the payoff from defecting,

$$\alpha\{1+x+(1-\varepsilon)\cdot 1\}+(1-\alpha)\{1\}=\alpha x+\alpha(1-\varepsilon)+1.$$

Thus we must have

$$\alpha=\tilde{\alpha}(\varepsilon)=\frac{1+\varepsilon}{k(1-\varepsilon)^2+\varepsilon^2+2\varepsilon-1}.$$

As $\varepsilon\to 0$, $\tilde{\alpha}(\varepsilon)\to 1/(k-1)$ and the inequalities (2.1) will be satisfied.[8]

Finally, imperfect monitoring may allow the mixed strategy equilibrium to be purified. As an extreme example, add an earlier stage to the game. This earlier stage, like the other two stages, is a simultaneous move stage game with two actions, $L$ and $R$. Suppose the payoffs in this stage game are such that $L$ is a dominant action in the overall game (not just in the first stage). Suppose, moreover, that imperfect monitoring in this stage generates continuously distributed signals $\omega_i^1\in[0,1]$. Let $x$ satisfy $F_L(x)=\tilde{\alpha}(\varepsilon)$, where $F_L$ is the distribution function of $\omega_i^1$ when both players choose $L$. Then, the following pure strategy profile is an equilibrium: In the first stage, both players choose $L$. In the second stage, player $i$ chooses $C$ if he observed a signal $\omega_i^1\leq x$, and $D$ otherwise. In the last stage, each player chooses $G$ if he had chosen $C$ in the second stage and observed $c$, and chooses $B$ otherwise. Note that by choosing $R$ in the first stage, a player can influence (perhaps advantageously) the distribution of his opponent's play in the second stage. The assumption that $L$ is dominant in the overall game is a crude assumption ruling out the profitability of such a deviation. A similar effect is obtained by finitely repeating the prisoner's dilemma before playing the

[8]Since both players are randomizing in the first stage, a player's probability of $C$ is used to make the opponent indifferent, and so the mixed strategy equilibrium exists for all small $\varepsilon$. In contrast, the mixed strategy equilibrium of footnote 5 required a particular value of $\varepsilon$.

coordination game, discounting payoffs, and only using the first period signal to purify behavior in the penultimate stage. This example also requires continuously distributed signals to ensure that the appropriate probability of $C$ in the second stage can be achieved. Since the signals are not directly payoff relevant, each player must be indifferent between the two choices in the second stage, and this is guaranteed by the appropriate probability of $C$ by the opponent.

## 3. An Anti-Folk Theorem

In this section, we present an example of a stage game with multiple Nash equilibria with the property that for *any* number of finite repetitions and *any* full support private monitoring technology, every equilibrium of the finitely repeated game consists only of repetitions of static Nash equilibria.

We will be concerned with a two player "convention" game. There are two players, 1 and 2. Player $i$ chooses an action $a_i \in [0,1]$. Payoff functions are

$$g_1(a_1,a_2) = a_2 - \beta (a_1 - a_2)^2$$

and

$$g_2(a_1,a_2) = 1 - a_1 - \beta (a_1 - a_2)^2,$$

where $\beta > 0$. This is a special case of a class of games analyzed by Shin and Williamson [21]. There is a continuum of Nash equilibria: since each player's best response is to copy his opponent's action, $(a_1, a_2)$ is a Nash equilibrium if and only if $a_1 = a_2$. Thus the sum of players' payoffs is 1 in any equilibrium. But if player 1 chooses action 0 and player 2 chooses action 1, both players receive payoff $1 - \beta$. So for small $\beta$, there is a feasible action profile Pareto-dominating the symmetric Nash equilibrium.

We first analyze what happens when this game is played once, but each player has access to some payoff irrelevant signal; specifically, each player observes a payoff irrelevant signal $\omega_i \in \Omega_i$, where each $\Omega_i$ is finite and $(\omega_1, \omega_2)$ is drawn according to some full support distribution $\pi \in \Delta(\Omega_1 \times \Omega_2)$. For any $x \in [0,1]$, this game has an equilibrium where each type of each player chooses action $x$. Shin and Williamson [21] showed that there are no other equilibria. The argument is elementary. Let $\bar{x}$ be the largest action chosen by any type in an equilibrium. This is a best response only if every type of the other player chooses action $\bar{x}$. Thus play contingent on (full support) payoff irrelevant signals is inconsistent with equilibrium in this example.

Now we consider what happens when the convention game is repeated $T$ times. With perfect monitoring, we could clearly obtain a finite folk theorem along the lines of Benoit and Krishna [6]. We will show that if there is full support imperfect private monitoring, Nash equilibrium play must involve a finite sequence of static Nash equilibria.

Player $i$ observes a signal $\omega_i$ from a finite set $\Omega_i$; let $\pi\left((\omega_1, \omega_2)\,|(a_1, a_2)\right)$ be the probability that signals $(\omega_1, \omega_2)$ are observed if actions are $(a_1, a_2)$; $\pi$ has full support. A player $i$ period $t$ history is $h_i^t = (a_i^\tau, \omega_i^\tau)_{\tau=1}^{t-1}$; write $H_i^t$ for the set of such histories. Write $\Delta^*(S)$ for the set of simple (i.e., finite support) probability distributions on $S$. A player $i$ period $t$ simple mixed strategy is a function $\sigma_i^t : H_i^t \to \Delta^*([0,1])$. A player $i$ strategy is $\sigma_i \equiv (\sigma_i^t)_{t=1}^{T}$. Given $\sigma_i$, write $\widetilde{A}_i^t(\sigma_i)$ for the set of action sequences up to $t$ played with positive probability by player $i$, i.e.,

$$\widetilde{A}_i^t(\sigma_i) = \left\{ (a_i^\tau)_{\tau=1}^t \in [0,1]^t \;\middle|\; \begin{array}{l} \text{for some } (\omega_i^\tau)_{\tau=1}^{t-1},\ \sigma_i^\tau\left(\left(a_i^{\tau'}, \omega_i^{\tau'}\right)_{\tau'=1}^{\tau-1}\right)[a_i^\tau] > 0 \\ \text{for } \tau = 1, .., t \end{array} \right\}.$$

**Proposition 2.** *The strategy profile $(\sigma_1, \sigma_2)$ is a Nash equilibrium of the finitely repeated convention game if and only if there exists $(x^t)_{t=1}^T \in [0,1]^T$, such that $\sigma_i^t\left((a_i^\tau, \omega_i^\tau)_{\tau=1}^{t-1}\right)[x^t] = 1$ whenever $(a_i^\tau)_{\tau=1}^{t-1} \in \widetilde{A}_i^{t-1}(\sigma_i)$.*

**Proof.** Write $x^T$ for the maximum of the set of $x$ such that $\sigma_i^T\left((a_i^\tau, \omega_i^\tau)_{\tau=1}^{T-1}\right)[x] > 0$ for some $i$ and $(a_i^\tau)_{\tau=1}^{T-1} \in \widetilde{A}_i^{T-1}(\sigma_i)$ and $(\omega_i^\tau)_{\tau=1}^{T-1} \in [\Omega_i]^{T-1}$. Since this history is reached with positive probability, the equilibrium type playing $x^T$ must be playing a best response. Since any play by other players is consistent with player $i$'s private history (by the full support assumption), every equilibrium type of every other player must be playing $x^T$. Thus every equilibrium type of player $i$ must choose $x^T$. Now write $x^{T-1}$ for the maximum of the set of $x$ such that $\sigma_i^{T-1}\left(h_i^{T-1}\right)[x] > 0$ for some $i$ and $h_i^{T-1}$ with $(a_i^\tau)_{\tau=1}^{T-2} \in \widetilde{A}_i^{T-2}(\sigma_i)$. But the equilibrium type playing $x^{T-1}$ knows that all types of all players will play $x^T$ next period. Thus his best response depends only on payoffs in period $T-1$. Now the argument iterates to show the result.

■

The argument applies to a larger class of coordination games that satisfy the condition (CC) of Shin and Williamson [21, p. 262]. Suppose the action space for all players is the interval $[0,1]$ and let $F$ be a distribution on $[0,1]$. Let $a_m$ ($a_M$) denote the inf (respectively, sup) of the support of $F$. Then (CC) requires that if $a_m = a_M$ then the best reply to $F$ is $a_m$. In addition, if $a_m < a_M$ then the best reply lies *strictly* between $a_m$ and $a_M$. Clearly, any game that falls into this class after relabelling of actions is also covered.

## 4. Folk Theorems

In this section, we prove a folk theorem for infinitely repeated games with a discrete action set, *if we are allowed to choose a full support private monitoring technology.* Specifically, suppose that there is a pure strategy payoff profile $u^*$ that strictly Pareto-dominates some strict pure strategy Nash equilibrium payoff profile. Then there is an equilibrium of the infinitely repeated game supporting payoffs arbitrarily close to $u^*$ for all discount rates close to 1.

We note two reasons why the construction is not trivial, even though we pick the private monitoring technology. First, the previous section proved an anti-folk theorem (with a finite horizon) for the convention game *for any full support private monitoring technology.* Some special properties—including the continuum of actions—were necessary for the result that no degree of approximate common knowledge was sufficient to achieve contingent coordination. Nonetheless, this result tells us that there are limits to what can be achieved with full support private monitoring, uniformly across full support monitoring technologies.[9]

Second, an important paper by Compte [8] considered trigger-strategy equilibria of the infinitely-repeated prisoner's dilemma. A trigger-strategy equilibrium has the property that if, in equilibrium, a player defects, he then defects forever with probability one. Compte showed that, for some class of full support independent signal private monitoring technologies and discount rates close to one, the average expected payoff is close to the payoff from defection.[10] We show how, without independent signals, simple pure strategies can generate a folk theorem

Our results should be contrasted with positive folk theorem results that work for a range of private monitoring technologies. Sekiguchi [20] proves a folk theorem, using mixed strategies, for a class of repeated prisoner's dilemmas, for sufficiently accurate private monitoring technologies that may be independent. We discuss the relation to this result in more detail below.[11]

Fix a finite stage game with players $\mathcal{I} = \{1, .., I\}$, action sets $\{A_i\}_{i \in \mathcal{I}}$ and payoff functions $\{u_i\}_{i \in \mathcal{I}}$, each $u_i : A \to \Re$ where $A = A_1 \times .. \times A_I$. Let $\underline{a}$ be a strict pure strategy Nash equilibrium and write $\underline{u}_i = u_i(\underline{a})$. Let $\bar{a}$ be another

[9] Recall that we only proved a *finite horizon* anti-folk theorem for the convention game. The arguments of this section do not apply to games with a continuum of actions, and we do not know if the arguments could be adapted to incorporate the convention game.

[10] Matsushima [17] also shows an anti-folk theorem. In particular, suppose that signals are independent and that players are restricted to pure strategies which depend on payoff-irrelevant histories *only* if that payoff-irrelevant history is correlated with other players' future play. These restrictions are enough to prevent co-ordination.

[11] Bhaskar and van Damme [7] prove a folk theorem in an infinitely repeated one-sided moral hazard game using mixed strategies and sunspots.

pure strategy profile with $\overline{u}_i = u_i(\bar{a}) > \underline{u}_i$ for all $i \in \mathcal{I}$.

Let

$$\begin{aligned} \lambda &= \max_{i\in\mathcal{I}} \max_{a,a'\in A} \left|u_i(a) - u_i(a')\right|; \\ \mu &= \min_{i\in\mathcal{I}} \min_{\{a,a'\in A: u_i(a)\neq u_i(a')\}} \left|u_i(a) - u_i(a')\right|. \end{aligned}$$

The game is repeated infinitely often. Each player discounts the future with discount rate $\delta \in (0,1)$. Each player $i$ observes a private signal $\omega_i \in \{H, L\}$ concerning the action profile in the previous period. Write $\Omega$ for the set of possible signal profiles, $\{H, L\}^I$; write $\omega$ for a typical element of $\Omega$, $\mathbf{H}$ for the element where all observe $H$ and $\mathbf{L}$ for the element where all observe $L$. Write $\pi(\omega|a)$ for the probability of signal profile $\omega$ given action profile $a$. We will be interested in a class of full support private monitoring technologies parameterized by three numbers, $\alpha$, $\beta$ and $\varepsilon$, all in the interval $(0,1)$; let

$$\pi(\omega|\bar{a}) = \begin{cases} \alpha(1-2\varepsilon), & \text{if } \omega = \mathbf{H}, \\ (1-\alpha)(1-2\varepsilon), & \text{if } \omega = \mathbf{L}, \\ \frac{\varepsilon}{(2^{I-1}-1)}, & \text{if } \omega \notin \{\mathbf{H}, \mathbf{L}\}, \end{cases}$$

and, for all $a \neq \bar{a}$,[12]

$$\pi(\omega|a) = \begin{cases} \beta(1-2\varepsilon), & \text{if } \omega = \mathbf{H}, \\ (1-\beta)(1-2\varepsilon), & \text{if } \omega = \mathbf{L}, \\ \frac{\varepsilon}{(2^{I-1}-1)}, & \text{if } \omega \notin \{\mathbf{H}, \mathbf{L}\}. \end{cases}$$

Writing $\omega_i^t$ for player $i$'s private signal in period $t$, a period $t$ private history for player $i$ is $h_i^t = (a_i^\tau, \omega_i^\tau)_{\tau=1}^{t-1}$.

### 4.1. An "Alternating Nash" Folk Theorem

Our first result restricts attention to strategies that depend only on one-period histories. Because we restrict attention to such strategies (and finite action spaces), it is possible to impose restrictions on the private monitoring technologies sufficient to ensure the requisite approximate common knowledge at any history. In particular, these restrictions on the monitoring technology are independent of $\delta$. In the next sub-section, with the same private monitoring technology, we prove that Nash reversion is also an equilibrium.

[12] While the number of players, $I$, does not explicitly play a role, it is worth noting that the signaling structure is more "discontinuous" for large $I$—only one player deviating from $\bar{a}_i$ leads to a big change in $\pi$ independent of $I$.

Consider the following simple strategy for player $i$:

$$\hat{s}_i^t\left(h_i^t\right) = \begin{cases} \bar{a}_i, & \text{if } t = 1, \\ \bar{a}_i, & \text{if } t > 1 \text{ and } \omega_i^{t-1} = H, \\ \underline{a}_i, & \text{if } t > 1 \text{ and } \omega_i^{t-1} = L. \end{cases}$$

We will be interested in the case where $\alpha$ is close to 1, $\beta$ is close to 0 and $\varepsilon$ is much closer to 0 than either $1-\alpha$ or $\beta$. Considering $\varepsilon$ small implies that whatever signal a player observes, he expects that other players have observed the same signal. This kind of information structure makes sense if there is a public outcome $H$ or $L$ depending on players' actions and with high probability players observe the public outcome correctly. But with small probability (of order $\varepsilon$) players' private observations of the public outcome are corrupted. However, we also take $\alpha$ close to 1 and $\beta$ close to 0. This is necessary for both the profile to be an equilibrium and for it to yield an average expected payoff close to $\bar{u}$. For the case of imperfect *public* monitoring, $\varepsilon = 0$, this strategy profile is not a Nash equilibrium of the repeated game for $\alpha$ small or $\beta$ large. Moreover, the signaling structure violates the pairwise identification and pairwise full rank restrictions of Fudenberg, Levine, and Maskin [11], so that we should not expect a public monitoring folk theorem to hold for arbitrary $\alpha$ and $\beta$.

We start by giving sufficient conditions for $(\widehat{s}_i)_{i\in\mathcal{I}}$ to be an equilibrium. We impose the following joint restrictions on the private monitoring technology $(\alpha, \beta, \varepsilon)$ and the discount factor $\delta$:

$$\alpha > \beta, \tag{4.1}$$

$$\frac{\beta\left(1-2\varepsilon\right)}{\beta\left(1-2\varepsilon\right)+\varepsilon} > \frac{\lambda}{\mu}\left(\frac{1}{\delta\left(1-2\varepsilon\right)\left(\alpha-\beta\right)}-1\right), \text{ and} \tag{4.2}$$

$$\frac{\left(1-\alpha\right)\left(1-2\varepsilon\right)}{\left(1-\alpha\right)\left(1-2\varepsilon\right)+\varepsilon} > \frac{\lambda}{\lambda+\mu\left(1-\delta\left(1-2\varepsilon\right)\left(\alpha-\beta\right)\right)}. \tag{4.3}$$

**Lemma 1.** *If $(\alpha, \beta, \varepsilon, \delta)$ satisfy (4.1-4.3), then $(\hat{s}_i)_{i\in\mathcal{I}}$ is an equilibrium.*

Restriction (4.1) can be viewed as a normalization on the interpretation of the signals. If $\alpha = \beta$, then of course conditioning behavior on the signal cannot discipline behavior. If, on the other hand, $\alpha \neq \beta$ then the signal realized with largest probability after $\bar{a}$ is labeled $H$. Restriction (4.2) comes from the $H$-incentive constraint, i.e., it is the condition that guarantees that it is optimal for player $i$ to choose $\bar{a}_i$ after observing $\omega_i^t = H$ in the previous period. This condition still has bite if $\varepsilon = 0$. Finally, restriction (4.3) comes from the $L$-incentive constraint, i.e., it is the condition that guarantees that it is optimal for player $i$ to choose $\underline{a}_i$ after observing $\omega_i^t = L$ in the previous period. This condition is satisfied for all $\alpha$, $\beta$, and $\delta$ if $\varepsilon = 0$, since $\underline{a}$ is a Nash equilibrium of the stage game and so is statically self-enforcing. Condition (4.3) is needed since, with imperfect private monitoring, a player after observing $L$ does not know whether the other players have also observed $L$.

**Proof.** Write $v_i(\omega)$ for player $i$'s continuation value (under the proposed strategies) when the private signal profile was $\omega$. Define $\tilde{a}_i(H) = \bar{a}_i$, $\tilde{a}_i(L) = \underline{a}_i$, $\tilde{a}(\omega) = (\tilde{a}_i(\omega_i))_{i\in\mathcal{I}}$ and $\tilde{u}_i(\omega) = u_i(\tilde{a}(\omega))$. Then,

$$v_i(\omega) = \tilde{u}_i(\omega) + \delta \sum_{\omega' \in \Omega} \pi(\omega' | \tilde{a}(\omega)) v_i(\omega'). \tag{4.4}$$

We will be concerned with $\Delta_i = v_i(\mathbf{H}) - v_i(\mathbf{L})$. From (4.4), (since $\pi(\omega|\mathbf{H}) = \pi(\omega|\mathbf{L})$ for $\omega \notin \{\mathbf{H}, \mathbf{L}\}$) we have $\Delta_i = \bar{u}_i - \underline{u}_i + \delta(1-2\varepsilon)(\alpha-\beta)\Delta_i$, and so

$$\Delta_i = \frac{\bar{u}_i - \underline{u}_i}{1 - \delta(1-2\varepsilon)(\alpha-\beta)}.$$

Now suppose player $i$ observes $H$ and attaches probability $\xi$ to all his opponents observing signal $H$. A sufficient condition for $\bar{a}_i$ to be a best response is, for any $a \equiv (a_i, a_{-i}) \in A$,

$$\begin{aligned} &\xi\left(\bar{u}_i + \delta \sum_{\omega'\in\Omega} \pi(\omega'|\bar{a}) v_i(\omega')\right) \\ &\qquad + (1-\xi)\left(u_i(\bar{a}_i, a_{-i}) + \delta \sum_{\omega'\in\Omega} \pi(\omega'|(\bar{a}_i, a_{-i})) v_i(\omega')\right) \\ \geq\ &\xi\left(u_i(a_i, \bar{a}_{-i}) + \delta \sum_{\omega'\in\Omega} \pi(\omega'|(a_i, \bar{a}_{-i})) v_i(\omega')\right) \\ &\qquad + (1-\xi)\left(u_i(a_i, a_{-i}) + \delta \sum_{\omega'\in\Omega} \pi(\omega'|(a_i, a_{-i})) v_i(\omega')\right). \end{aligned} \tag{4.5}$$

For $a_{-i} = \bar{a}_{-i}$, (4.5) reduces to $\delta(1-2\varepsilon)(\alpha-\beta)\Delta_i \geq u_i(a_i, \bar{a}_{-i}) - \bar{u}_i$, which is implied by condition (4.2). If $a_{-i} \neq \bar{a}_{-i}$, re-arranging and substituting for $\pi$ gives

$$\delta(1-2\varepsilon)(\alpha-\beta)\Delta_i\xi \geq \xi(u_i(a_i, \bar{a}_{-i}) - \bar{u}_i) + (1-\xi)(u_i(a_i, a_{-i}) - u_i(\bar{a}_i, a_{-i})).$$

The right hand side is at most $\lambda$ and $\Delta_i$ is at least $\mu/\left(1-\delta\left(1-2\varepsilon\right)\left(\alpha-\beta\right)\right)$, so it is enough to have

$$\xi \geq \bar{\xi} = \frac{\lambda}{\mu}\left(\frac{1}{\delta\left(1-2\varepsilon\right)\left(\alpha-\beta\right)}-1\right). \tag{4.6}$$

It would be very complicated to calculate players' posterior probabilities that all other players will observe $H$ under every possible history. But our information structure allows us to bound these posteriors. If a player believed that action profile $\bar{a}$ was chosen last period, he attaches probability $\alpha\left(1-2\varepsilon\right)/[\alpha\left(1-2\varepsilon\right)+\varepsilon]$ to the event that all other players observed $H$ this period. On the other hand, if he believed that some other action profile had been chosen last period, he attaches probability $\beta\left(1-2\varepsilon\right)/[\beta\left(1-2\varepsilon\right)+\varepsilon]$ to that event. Since $\beta<\alpha$ (by condition 4.1), we know that, *whatever his beliefs about the previous action profile*, he attaches probability at least $\beta\left(1-2\varepsilon\right)/\left[\beta\left(1-2\varepsilon\right)+\varepsilon\right]$ to his opponents all observing $H$. But by condition (4.2), this exceeds $\bar{\xi}$.

Now suppose player $i$ observes $L$ and attaches probability $\zeta$ to all his opponents observing signal $L$. A sufficient condition for $\underline{a}_i$ to be a best response is, for any $a\equiv(a_i,a_{-i})\in A$,

$$\begin{aligned}
& \zeta\left(\underline{u}_i+\delta\sum_{\omega'\in\Omega}\pi\left(\omega'\middle|\,\underline{a}\right)v_i\left(\omega'\right)\right) \\
& \qquad +\left(1-\zeta\right)\left(u_i\left(\underline{a}_i,a_{-i}\right)+\delta\sum_{\omega'\in\Omega}\pi\left(\omega'\middle|\left(\underline{a}_i,a_{-i}\right)\right)v_i\left(\omega'\right)\right) \\
\geq\ & \zeta\left(u_i\left(a_i,\underline{a}_{-i}\right)+\delta\sum_{\omega'\in\Omega}\pi\left(\omega'\middle|\left(a_i,\underline{a}_{-i}\right)\right)v_i\left(\omega'\right)\right) \qquad (4.7) \\
& \qquad +\left(1-\zeta\right)\left(u_i\left(a_i,a_{-i}\right)+\delta\sum_{\omega'\in\Omega}\pi\left(\omega'\middle|\left(a_i,a_{-i}\right)\right)v_i\left(\omega'\right)\right).
\end{aligned}$$

For $a_{-i}=\underline{a}_{-i}$, (4.7) reduces to $u_i\left(a_i,\underline{a}_{-i}\right)\leq\underline{u}_i$, which holds because $\underline{a}$ is a Nash equilibrium. If $a_{-i}\neq\underline{a}_{-i}$, re-arranging and substituting for $\pi$ gives

$$\begin{aligned}
\zeta\left(\underline{u}_i-u_i\left(a_i,\underline{a}_{-i}\right)\right) \geq\ & \left(1-\zeta\right)\{\left(u_i\left(a_i,a_{-i}\right)-u_i\left(\underline{a}_i,a_{-i}\right)\right) \\
& +\delta\sum_{\omega'\in\Omega}\left(\pi\left(\omega'\middle|\left(a_i,a_{-i}\right)\right)-\pi\left(\omega'\middle|\left(\underline{a}_i,a_{-i}\right)\right)\right)v_i\left(\omega'\right)\}.
\end{aligned}$$

But $\underline{u}_i - u_i(a_i, \underline{a}_{-i}) \geq \mu$ (if $a_i \neq \underline{a}_i$), $u_i(a_i, a_{-i}) - u_i(\underline{a}_i, a_{-i}) \leq \lambda$, and

$$\begin{aligned}\sum_{\omega' \in \Omega} \left(\pi\left(\omega' \middle| (a_i, a_{-i})\right) - \pi\left(\omega' \middle| (\underline{a}_i, a_{-i})\right)\right) v_i(\omega') &\leq (1-2\varepsilon)(\alpha-\beta)\Delta_i \\ &\leq \frac{(1-2\varepsilon)(\alpha-\beta)\lambda}{1-\delta(1-2\varepsilon)(\alpha-\beta)}.\end{aligned}$$

Thus, it is enough to have

$$\frac{\zeta}{1-\zeta} \geq \frac{\lambda\left(1+\frac{\delta(1-2\varepsilon)(\alpha-\beta)}{1-\delta(1-2\varepsilon)(\alpha-\beta)}\right)}{\mu},$$

i.e.,

$$\zeta \geq \overline{\zeta} = \frac{1}{1+\frac{\mu}{\lambda}\left(1-\delta(1-2\varepsilon)(\alpha-\beta)\right)}.$$

But (by a similar argument to that given above) any player observing $L$ attaches probability at least $(1-\alpha)(1-2\varepsilon)/[(1-\alpha)(1-2\varepsilon)+\varepsilon]$ to all others having observed $L$. This expression is more than $\overline{\zeta}$ by condition (4.3).

■

We now give a parameter restriction that ensures that the average expected payoff from the profile $\hat{s}$ is close to $\bar{u}$. Note that if $\varepsilon = 0$, the inequality (4.8) reduces to

$$\frac{1-\alpha}{1-\delta(\alpha-\beta)} < \frac{\eta}{\lambda},$$

i.e., for $\eta$ small, $\alpha$ must be close to 1. As we indicated above, we need this condition while Fudenberg, Levine, and Maskin [11] do not, because our signaling structure when $\varepsilon = 0$ violates pairwise identifiability.

**Lemma 2.** *If, for some $\eta > 0$,*

$$\frac{1-\alpha(1-2\varepsilon)}{1-\delta(1-2\varepsilon)(\alpha-\beta)} + 2\varepsilon < \frac{\eta}{\lambda}, \tag{4.8}$$

*then player $i$'s average payoff under $(\widehat{s}_i)_{i\in\mathcal{I}}$ is at least $\bar{u}_i - \delta\eta$.*

**Proof.** Since $\pi(\cdot|a)$ is constant for all $a \neq \bar{a}$, $v_i(\omega) = v_i(\mathbf{L}) + (\tilde{u}_i(\omega) - \underline{u}_i)$ for all $\omega \neq \mathbf{H}$. Thus (by (4.4))

$$v_i(\mathbf{H}) = \bar{u}_i + \delta v_i(\mathbf{H}) - \delta(1-\alpha(1-2\varepsilon))\Delta_i + \delta \sum_{\omega \notin \{\mathbf{H},\mathbf{L}\}} \frac{\varepsilon}{(2^{I-1}-1)}(\tilde{u}_i(\omega) - \underline{u}_i).$$

But since $\left|\tilde{u}_i(\omega) - \underline{u}_i\right| \leq \lambda$ and substituting for $\Delta_i \leq \frac{\lambda}{1-\delta(1-2\varepsilon)(\alpha-\beta)}$,

$$\bar{u}_i - (1-\delta)\, v_i(\mathbf{H}) \leq \delta \left( \frac{1-\alpha(1-2\varepsilon)}{1-\delta(1-2\varepsilon)(\alpha-\beta)} + 2\varepsilon \right) \lambda.$$

Thus $\bar{u}_i - (1-\delta)\, v_i(\mathbf{H}) < \delta\eta$ by (4.8).

■

The sufficient conditions of Lemmas 1 and 2 can be expressed in a more understandable form as follows. Define

$$\tilde{\alpha}(\beta,\eta) = \max\left\{\beta + \frac{\lambda}{\lambda+\mu}, 1 - \frac{\beta\eta}{\lambda}\right\}, \tag{4.9}$$

$$\tilde{\varepsilon}(\beta,\alpha,\eta) = \max\left\{\varepsilon \;\middle|\; \begin{array}{l} (1).\ \frac{\beta(1-2\varepsilon)}{\beta(1-2\varepsilon)+\varepsilon} \geq \frac{\lambda}{\mu}\left(\frac{1}{(1-2\varepsilon)(\alpha-\beta)} - 1\right) \\ (2).\ \frac{(1-\alpha)(1-2\varepsilon)}{(1-\alpha)(1-2\varepsilon)+\varepsilon} \geq \frac{\lambda}{\lambda+\mu(1-\alpha+\beta)} \\ (3).\ \left(\frac{1-\alpha(1-2\varepsilon)}{1-\alpha+\beta} + 2\varepsilon\right)\lambda \leq \eta \end{array}\right\}, \tag{4.10}$$

and

$$\tilde{\delta}(\beta,\alpha,\varepsilon) = \frac{\lambda(\beta(1-2\varepsilon)+\varepsilon)}{(1-2\varepsilon)(\alpha-\beta)(\beta(1-2\varepsilon)(\lambda+\mu)+\lambda\varepsilon)}. \tag{4.11}$$

The expression $\tilde{\alpha}(\beta,\eta)$ is the lower bound on $\alpha$ that ensures that the three inequalities defining $\tilde{\varepsilon}$ in (4.10) are all satisfied at $\varepsilon = 0$. Inequality (4.10.1) is (4.2) evaluated at $\delta = 1$; this inequality becomes more binding as $\delta$ is lowered from $\delta = 1$. Inequality (4.10.2) is (4.3) evaluated at $\delta = 1$; this inequality becomes less binding as $\delta$ is lowered from $\delta = 1$. And inequality (4.10.3) is (4.8) evaluated at $\delta = 1$; this inequality also becomes less binding as $\delta$ is lowered from $\delta = 1$. Finally, $\tilde{\delta}(\beta,\alpha,\varepsilon)$ is the lower bound on $\delta$ for which (4.2) holds.

Tedious algebra confirms that if $0 < \beta < \mu/(\lambda+\mu)$, then $\tilde{\alpha}(\beta,\eta) < 1$; if in addition, $\tilde{\alpha}(\beta,\eta) < \alpha < 1$, then $\tilde{\varepsilon}(\beta,\alpha,\eta) > 0$; and finally, if in addition, $0 < \varepsilon < \tilde{\varepsilon}(\beta,\alpha,\eta)$, then $\tilde{\delta}(\beta,\alpha,\varepsilon) < 1$. Summarizing:

**Lemma 3.** *If* $0 < \beta < \frac{\mu}{\lambda+\mu}$, $\tilde{\alpha}(\beta,\eta) < \alpha < 1$, $0 < \varepsilon < \tilde{\varepsilon}(\beta,\alpha,\eta)$ *and* $\tilde{\delta}(\beta,\alpha,\varepsilon) < \delta < 1$, *then conditions (4.1) through (4.8) are satisfied.*

From Lemmas 1, 2, and 3 we have:

**Proposition 3.** *If* $0 < \beta < \mu/(\lambda+\mu)$, $\tilde{\alpha}(\beta,\eta) < \alpha < 1$ *and* $0 < \varepsilon < \tilde{\varepsilon}(\beta,\alpha,\eta)$, *then* $\hat{s}$ *is an equilibrium in which each player* $i$ *receives average utility at least* $\bar{u}_i - \eta$, *for all* $\delta$ *sufficiently close to* 1.

### 4.2. Nash Reversion

We now examine the possibility that static Nash reversion may be a Nash equilibrium. We keep the previous signaling technology. Let $\tilde{s}$ denote grim trigger, i.e.,

$$\tilde{s}_i^t(h_i^t) = \begin{cases} \bar{a}_i, & \text{if } t = 1, \\ \bar{a}_i, & \text{if } t \geq 2 \text{ and } h_i^t = \tilde{h}_i^t, \\ \underline{a}_i, & \text{if } t \geq 2 \text{ and } h_i^t \neq \tilde{h}_i^t, \end{cases}$$

where $\tilde{h}_i^t = \overbrace{(\bar{a}_i H, \ldots, \bar{a}_i H)}^{t \text{ periods}}$. Let $m_i = \min_a u_i(a)$ and $M_i = \max_a u_i(a)$. In contrast to the strategy profile of the previous section, this profile depends upon the entire history, not just the last period. None the less, we claim

**Proposition 4.** *Suppose* $(1-\alpha\delta)(1-\delta(1-\beta)) M_i < \bar{u}_i (1-\delta) - \delta[\alpha(1-\delta(1-\beta)) - \beta]\underline{u}_i$. *There exists* $\bar{\varepsilon} > 0$ *such that for all* $\varepsilon \in (0, \bar{\varepsilon})$, $\tilde{s}$ *is a Nash equilibrium.*

It is worth noting that as $\delta \to 1$, $\bar{\varepsilon} \to 0$ (we return to this point below). This is not true of the profile $\hat{s}$ analyzed in the previous section. While $\tilde{s}$ is not sequential, its outcome path can be supported by a sequential equilibrium when it is Nash (Sekiguchi [20, Proposition 3]). It is useful to explain why the profile is not sequential. First note that if player $i$ deviates in a period $t$ and then observes $H$ in that period, then due to the highly correlated information structure player $i$ is reasonably confident that player $j$ did not observe $L$ and so $i$ "got away with it." So at the very least a sequential profile cannot specify $\underline{a}_i$ after $a_i H$ (for any $a_i \neq \bar{a}_i$). Let $\tilde{H}_i$ denote the set of player $i$'s private histories in which player $i$ has never observed the signal $L$ and consider the profile which specifies $\bar{a}_i$ after every history in $\tilde{H}_i$ and $\underline{a}_i$ otherwise. The profile differs from Nash reversion in that it ignores the player's own deviation. Play only switches to static Nash if the player is sufficiently confident that the deviation was detected (indicated by the player also receiving the signal $L$). This again is not sequential. After the private history $(a_i L, \bar{a}_i H, \ldots, \bar{a}_i H) \notin \tilde{H}_i$, to the event that all the other players are still playing $\bar{a}_{-i}$, player $i$ assigns a probability that converges to 1 as $t \to \infty$. Sequential rationality would then require that player $i$ continue to play $\bar{a}_i$.

There is a close relationship between Proposition 4 and Sekiguchi [20, Proposition 3] (indeed, the proof of Claim 1 below essentially reproduces the analogous argument from that paper) which we discuss below.

**Proof.** We need to verify that player $i$ finds it optimal to start with $\bar{a}_i$, to continue with $\bar{a}_i$ as long as $H$ has been observed, and that if ever $L$ was observed, to play $\underline{a}_i$ thereafter. It is easiest to think of $\tilde{s}_i$ as being in one of two states, $C$ or $D$, with $D$ being absorbing and play starting in $C$. Let $v_i^{\xi}$ be the value of player

$i$ when the state of all the players is described by $\xi \in \{C, D\}^I$. While the values of $v_i^{\xi}$ can be calculated recursively, it suffices for our purposes to calculate

$$v_i^D \equiv v_i^{(D,\dots,D)} = \frac{\underline{u}_i}{(1-\delta)},$$

and

$$v_i^C \equiv v_i^{(C,\dots,C)} = \bar{u}_i + \delta \left\{ \alpha (1-2\varepsilon) v_i^C + (1-\alpha)(1-2\varepsilon) v_i^D + O(\varepsilon) \right\},$$

so that

$$v_i^C = \frac{\bar{u}_i + \delta \left\{ (1-\alpha)(1-2\varepsilon) v_i^D + O(\varepsilon) \right\}}{1-\delta\alpha(1-2\varepsilon)} = \frac{\bar{u}_i (1-\delta) + \delta \{(1-\alpha)(1-2\varepsilon) \underline{u}_i + O(\varepsilon)\}}{(1-\delta\alpha(1-2\varepsilon))(1-\delta)}$$

and

$$\lim_{\varepsilon \to 0} v_i^C = \frac{\bar{u}_i (1-\delta) + \delta (1-\alpha) \underline{u}_i}{(1-\delta\alpha)(1-\delta)}.$$

**Claim 1** $-$ $\bar{a}_i$ is optimal in period $t$ after the history $\tilde{h}_i^{t-1} = (\bar{a}_i H, \dots, \bar{a}_i H)$:

Let $p_i^t = \Pr\left\{ a_{-i}^t = \bar{a}_{-i} \middle| \tilde{h}_i^{t-1} \right\}$ (this also covers the initial period, where $\tilde{h}_i^0 = \{\emptyset\}$ and $p_i^1 = 1$). An upper bound to deviating at $t$ is obtained by noting that with probability $p_i^t$ the other players are still playing $\bar{a}_{-i}$, after choosing $a_i \neq \bar{a}_i$ at $t$ all players receive the signal $L$ with probability $(1-\beta)(1-2\varepsilon)$, and $M_i/(1-\delta)$ is the highest payoff $i$ could earn when players don't play $\underline{a}_{-i}$. Thus, deviating is not profitable at $t$ if

$$p_i^t \left\{ M_i + \delta \left\{ (1-\beta)(1-2\varepsilon) \frac{\underline{u}_i}{(1-\delta)} + (1-(1-\beta)(1-2\varepsilon)) \frac{M_i}{(1-\delta)} \right\} \right\}$$
$$+(1-p_i^t)\frac{M_i}{(1-\delta)} \leq p_i^t v_i^C + \left(1-p_i^t\right) \frac{m_i}{(1-\delta)}.$$

Since $p_i^t$ can be made arbitrarily close to 1 for all $t$ by choosing $\varepsilon$ sufficiently small, this inequality will hold for all $t$ if $\varepsilon$ is small and

$$(1-\alpha\delta)(1-\delta(1-\beta)) M_i < \bar{u}_i (1-\delta) - \delta (\alpha (1-\delta(1-\beta)) - \beta) \underline{u}_i.$$

Now we verify that $p_i^t$ can be made arbitrarily close to 1 for all $t$ by making $\varepsilon$ small. Since

$$\Pr\left\{ a_{-i}^{t-1} = \bar{a}_{-i} \mid \tilde{h}_i^{t-1} \right\} = \frac{\Pr\left\{ a_{-i}^{t-1} = \bar{a}_{-i},\ \omega_i^{t-1} = H \middle| \tilde{h}_i^{t-2},\ a_i^{t-1} = \bar{a}_i \right\}}{\Pr\left\{ \omega_i^{t-1} = H \middle| \tilde{h}_i^{t-2},\ a_i^{t-1} = \bar{a}_i \right\}}$$

$$
\begin{aligned}
&= \frac{\Pr\left\{\omega_i^{t-1} = H \left| a^{t-1} = \bar{a}\right.\right\} \Pr\left\{a_{-i}^{t-1} = \bar{a}_{-i} \left| \tilde{h}_i^{t-2}\right.\right\}}{\Pr\left\{\omega_i^{t-1} = H \left| \tilde{h}_i^{t-2},\ a_i^{t-1} = \bar{a}_i\right.\right\}} \\
&= \frac{\left(\alpha\left(1-2\varepsilon\right)+\varepsilon\right) p_i^{t-1}}{\Pr\left\{\omega_i^{t-1} = H \left| \tilde{h}_i^{t-2},\ a_i^{t-1} = \bar{a}_i\right.\right\}}
\end{aligned}
$$

and

$$
\begin{aligned}
\Pr\left\{\omega_i^{t-1} = H \left| \tilde{h}_i^{t-2},\ a_i^{t-1} = \bar{a}_i\right.\right\} &= \Pr\left\{a_{-i}^{t-1} = \bar{a}_{-i},\ \omega_i^{t-1} = H \left| \tilde{h}_i^{t-2},\ a_i^{t-1} = \bar{a}_i\right.\right\} \\
&\quad + \Pr\left\{a_{-1}^{t-1} \neq \bar{a}_{-i},\ \omega_i^{t-1} = H \left| \tilde{h}_i^{t-2},\ a_i^{t-1} = \bar{a}_i\right.\right\} \\
&= \Pr\left\{\omega_i^{t-1} = H \left| a^{t-1} = \bar{a}\right.\right\} p_i^{t-1} \\
&\quad + \Pr\left\{\omega_i^{t-1} = H \left| a^{t-1} \neq \bar{a}\right.\right\} (1 - p_i^{t-1}) \\
&= \left(\alpha\left(1-2\varepsilon\right)+\varepsilon\right) p_i^{t-1} + \left(\beta\left(1-2\varepsilon\right)+\varepsilon\right)(1 - p_i^{t-1}),
\end{aligned}
$$

we have

$$
\begin{aligned}
p_i^t = \Pr\left\{a_{-i}^t = \bar{a}_{-i} \left| \tilde{h}_i^{t-1}\right.\right\} &= \Pr\left\{a_{-i}^{t-1} = \bar{a}_{-i},\ \omega_{-i}^{t-1} = \mathbf{H}_{-i} \left| \tilde{h}_i^{t-1}\right.\right\} \\
&= \Pr\left\{\omega_{-i}^{t-1} = \mathbf{H}_{-i} \left| a^{t-1} = \bar{a},\ \omega_i^{t-1} = H\right.\right\} \Pr\left\{a_{-i}^{t-1} = \bar{a}_{-i} \left| \tilde{h}_i^{t-1}\right.\right\} \\
&= \frac{\alpha\left(1-2\varepsilon\right) p_i^{t-1}}{\left(\alpha\left(1-2\varepsilon\right)+\varepsilon\right) p_i^{t-1} + \left(\beta\left(1-2\varepsilon\right)+\varepsilon\right)(1 - p_i^{t-1})}.
\end{aligned}
$$

Letting $\ell_i^t$ denote the posterior odds ratio $\left(1 - p_i^t\right)/\,p_i^t$,

$$
\begin{aligned}
\ell_i^t &= \frac{\varepsilon p_i^{t-1} + \left(\beta\left(1-2\varepsilon\right)+\varepsilon\right)\left(1 - p_i^{t-1}\right)}{\alpha\left(1-2\varepsilon\right) p_i^{t-1}} \\
&= \frac{\varepsilon}{\alpha\left(1-2\varepsilon\right)} + \frac{\left(\beta\left(1-2\varepsilon\right)+\varepsilon\right)}{\alpha\left(1-2\varepsilon\right)} \ell_i^{t-1}.
\end{aligned}
$$

Since $\ell_i^1 = 0$, $\ell_i^t$ is monotonically increasing in $t$ and

$$
\ell_i^t \uparrow \frac{\varepsilon}{\left(\left(\alpha-\beta\right)\left(1-2\varepsilon\right)-\varepsilon\right)},
$$

which implies

$$
p_i^t \geq 1 - \ell_i^t > 1 - \frac{\varepsilon}{\left(\left(\alpha-\beta\right)\left(1-2\varepsilon\right)-\varepsilon\right)}.
$$

**Claim 2** – $\underline{a}_i$ is optimal in period $t$ after the history $h_i^t = \left(\tilde{h}_i^{t-1}, \bar{a}_i L\right)$:

By Bayes' rule,

$$\begin{aligned}
q_i^{t+1} &= \Pr\left\{a_{-i}^{t+1} = \underline{a}_{-i} \,\middle|\, \left(\tilde{h}_i^{t-1}, \bar{a}_i L\right)\right\} \\
&\geq \Pr\left\{a_{-i}^{t} = \bar{a}_{-i},\ \omega_{-i}^{t} = \mathbf{L}_{-i} \,\middle|\, \left(\tilde{h}_i^{t-1}, \bar{a}_i L\right)\right\} \\
&= \frac{\Pr\left\{a_{-i}^{t} = \bar{a}_{-i},\ \omega^{t} = \mathbf{L} \,\middle|\, \left(\tilde{h}_i^{t-1}, \bar{a}_i\right)\right\}}{\Pr\left\{\omega_i^{t} = L \,\middle|\, \left(\tilde{h}_i^{t-1}, \bar{a}_i\right)\right\}} \\
&= \frac{(1-\alpha)(1-2\varepsilon)\, p_i^t}{((1-\alpha)(1-2\varepsilon)+\varepsilon)\, p_i^t + ((1-\beta)(1-2\varepsilon)+\varepsilon)(1-p_i^t)} \\
&\to 1 \text{ as } \varepsilon \to 0.
\end{aligned}$$

The payoff to deviating after the history $h_i^t$ by choosing $a_i \neq \underline{a}_i$ and then following some continuation strategy is no more than

$$q_i^{t+1}\left\{\underline{u}_i - \mu + \frac{\delta \underline{u}_i}{(1-\delta)}\right\} + \left(1 - q_i^{t+1}\right)\left\{\frac{M_i}{(1-\delta)}\right\},$$

and so deviating after the history $h_i^t$ is not profitable if

$$q_i^{t+1}\left\{\underline{u}_i - \mu + \frac{\delta \underline{u}_i}{(1-\delta)}\right\} + \left(1 - q_i^{t+1}\right)\left\{\frac{M_i}{(1-\delta)}\right\} \leq q_i^{t+1}\left\{\frac{\underline{u}_i}{(1-\delta)}\right\} + \left(1 - q_i^{t+1}\right)\left\{\frac{m_i}{(1-\delta)}\right\},$$

i.e.,

$$\left(1 - q_i^{t+1}\right)\left\{\frac{(M_i - m_i)}{(1-\delta)}\right\} \leq q_i^{t+1}\mu,$$

which holds for $\varepsilon$ small.

**Claim 3** – $\underline{a}_i$ is optimal in period $T$ after the history $h_i^{T-1} = (\tilde{h}_i^{t-1}, \bar{a}_i L, \underline{a}_i \omega_i^{t+1}, \ldots, \underline{a}_i \omega_i^{T-1})$:

Again from Bayes' rule (and using the assumption that $\Pr\left\{\omega_i^{T-1} \,\middle|\, a^{T-1} = \underline{a}\right\} = \Pr\left\{\omega_i^{T-1} \,\middle|\, a^{T-1} \neq \bar{a}\right\}$),

$$q_i^T = \Pr\left\{a_{-i}^T = \underline{a}_{-i} \,\middle|\, h_i^{T-1}\right\} \geq \Pr\left\{a_{-i}^{T-1} = \underline{a}_{-i} \,\middle|\, h_i^{T-1}\right\} = q_i^{T-1}.$$

Thus, if $\underline{a}_i$ is optimal in period $t$ after the history $h_i^{t+1} = \left(\tilde{h}_i^t, \bar{a}_i L\right)$, then it is optimal in period $T > t$ after the history $h_i^{T-1} = \left(\tilde{h}_i^{t-1}, \bar{a}_i L, \underline{a}_i \omega_i^{t+1}, \ldots, \underline{a}_i \omega_i^{T-1}\right)$. ■

We now compare this result with the impressive result of Sekiguchi [20]. Sekiguchi [20] showed that, for some repeated prisoner's dilemmas, there exists a nearly efficient sequential equilibrium, when private monitoring is arbitrarily accurate and players are patient. There are three features we draw the reader's attention to. First, Sekiguchi's result does not make any assumptions on the nature of the private monitoring (it includes both independent and correlated signals). Second, his equilibrium is in mixed strategies, while ours is in pure. Finally, while his equilibrium builds on grim trigger (Nash reversion in his context), the final equilibrium is not grim trigger.

Crudely summarizing Sekiguchi's argument, he begins by considering a strategy profile that randomizes between always defection, and grim trigger. This profile is an equilibrium (given a payoff restriction) for moderate discount factors and sufficiently accurate private monitoring. Crudely, there are two things to worry about. First, if a player has been cooperating for a long time and has always received a cooperative signal, will the player continue to cooperate? The answer here is yes, given sufficiently accurate private monitoring (this is Claim 1 above). Note that the bound on the accuracy of the monitoring depends on the discount factor.

Second, will a player defect as soon as a defect signal is received? This is where the randomization and upper bound on the discount factor comes in. For illustrative purposes, suppose the players are playing the pure strategy profile of grim trigger. After the initial period, if player $i$ observes the defect signal, then the highest order probability events are that player $j$ did not defect but $i$ received an erroneous signal (in which case $j$ is still cooperative), and that player $j$ in the previous period had received an erroneous signal (in which case $j$ now defects forever). These two events have equal probability, and if players are not too patient (so that they are not willing to experiment), player $i$ will defect. If players are patient, then even a large probability that the opponent is already defecting may not be enough to ensure that the player defects: One more observation before the player commits himself may be quite valuable. Of course, in the initial period player $j$ is not responding to any signal, so in order for player $i$ to assign positive probability to the signal reflecting $j$'s behavior, $j$ must defect in the initial period with positive probability.[13] Since we have assumed that the

[13]In other words, grim trigger is not a Nash equilibrium because players have an incentive to ignore defect signals when first received (players believe their opponents are still cooperating and do not want to initiate the defect phase) and so players have no incentive to cooperate in the initial period.

Of course, the players must be indifferent between cooperation and defection in the initial period, and this determines the randomization probability. Moreover, as long as the discount is close to the value at which a player is indifferent between cooperation and defection against grim trigger in a game with perfect monitoring, then for sufficiently accurate monitoring, this randomization probability assigns small weight to initial defection.

private signals are highly correlated, these latter considerations are irrelevant. As soon as a player receives a defect signal, he assigns very high probability to his opponents having received the same signal, and so will defect. This is why we do not need randomization, nor an upper bound on the discount factor.

Sekiguchi then removes the upper bound on the discount factor by observing (following Ellison [10]) that the repeated game can be divided into $N$ distinct games, with the $k$th game played in period $k + tN$, where $t \in \mathbf{N}$. This gives an effective discount rate of $\delta^N$ on each game. Of course, the resulting strategy profile does not look like grim trigger.

As a final comment on Nash reversion, note that approximating the payoff vector $\bar{u}$ as we consider $\delta$ large may require taking $\varepsilon$ even smaller than the upper bound $\bar{\varepsilon}$ in Proposition 4. For fixed $\alpha$, $\beta$, $\delta$, and $\varepsilon$ satisfying the hypotheses of Proposition 4, there is an expected waiting time till the players switch from $\bar{a}$ which depends on $\varepsilon$, but is independent of $\delta$. Recall however, that $\delta \to 1$ requires $\varepsilon \to 0$ and the expected waiting time becomes infinite as $\varepsilon \to 0$. An alternative is to fix $\alpha$, $\beta$, $\delta$, and $\varepsilon$ satisfying the hypotheses of Proposition 4, and then apply Ellison's trick. This yields a profile for $\delta$ arbitrarily close to 1 while at the same time not requiring $\varepsilon \to 0$.

## References


[1] Dilip Abreu, David Pearce, and Ennio Stacchetti. Toward a theory of discounted repeated games with imperfect monitoring. *Econometrica*, 58:1041–1063, 1990.

[2] Illtae Ahn. Imperfect information repeated games with a single perfect observer. Mimeo, 1996.

[3] Massimiliano Amarante. Discounted repeated games with imperfect private monitoring, 1998. M.I.T.

[4] Masaki Aoyagi. A collusion mechanism in dynamic bertrand oligopoly under private observation of correlated demand shocks. Mimeo, University of Pittsburgh, 1997.

[5] Elchanan Ben-Porath and Michael Kahneman. Communication in repeated games with private monitoring. *Journal of Economic Theory*, 70:281–297, 1996.

[6] Jean-Pierre Benoit and Vijay Krishna. Finitely repeated games. *Econometrica*, 53:905–922, 1985.

[7] V. Bhaskar and Eric van Damme. Moral hazard and private monitoring. Mimeo, University of St. Andrews and CentER Tilburg, 1997.

[8] Olivier Compte. On failing to cooperate when monitering is private. Mimeo, C.E.R.A.S.-E.N.P.C. and CEPRAMAP, 1996.

[9] Olivier Compte. Communication in repeated games with imperfect private monitoring. *Econometrica*, 66:597–626, 1998.

[10] Glenn Ellison. Cooperation in the prisoner's dilemma with anonymous random matching. *Journal of Economic Theory*, 61:567–588, 1994.

[11] Drew Fudenberg, David K. Levine, and Eric Maskin. The folk theorem with imperfect public information. *Econometrica*, 62:997–1040, 1994.

[12] Edward Green and Robert Porter. Noncooperative collusion under imperfect price formation. *Econometrica*, 52:87–100, 1984.

[13] Atsushi Kajii and Stephen Morris. Refinements and higher order beliefs: A unified survey. Northwestern CMS-EMS Discussion Paper 1197, 1997.

[14] Atsushi Kajii and Stephen Morris. The robustness of equilibria to incomplete information. *Econometrica*, 65:1283–1309, 1997.

[15] Michihiro Kandori. Cooperation in finitely repeated games with imperfect private information. Mimeo, Princeton University, 1991.

[16] Michihiro Kandori and Hitoshi Matsushima. Private observation, communication and collusion. *Econometrica*, 66:627–652, 1998.

[17] Hitoshi Matsushima. On the theory of repeated games with private information. part i: Anti-folk theorem without communication. *Economic Letters*, 35:253–256, 1991.

[18] Dov Monderer and Dov Samet. Approximating common knowledge with common beliefs. *Games and Economic Behavior*, 1:170–190, 1989.

[19] Stephen Morris, Rafael Rob, and Hyun Shin. $p$-dominance and belief potential. *Econometrica*, 63:145–167, 1995.

[20] Tadashi Sekiguchi. Efficiency in repeated prisoner's dilemma with private monitoring. *Journal of Economic Theory*, 76:345–361, 1997.

[21] Hyun Shin and Timothy Williamson. How much common belief is necessary for a convention? *Games and Economic Behavior*, 13:252–268, 1996.