Federal Reserve Bank of Minneapolis
Research Department Staff Report 287

June 2002
revised November 2003

Finite Memory and Imperfect Monitoring*

Harold L. Cole

University of California, Los Angeles

Narayana R. Kocherlakota

Stanford University
and Federal Reserve Bank of Minneapolis

ABSTRACT

In this paper, we consider a class of infinitely repeated games with imperfect public monitoring. We look at strongly symmetric perfect public equilibria with memory K: equilibria in which strategies are restricted to depend only on the last K observations of public signals. Define Γ_K to be the set of payoffs of equilibria with memory K. We show that for some parameter settings, $\Gamma_K = \Gamma_\infty$ for sufficiently large K. However, for other parameter settings, we show that not only is $\lim_{K\to\infty} \Gamma_K \neq \Gamma_\infty$, but that Γ_k is a singleton. Moreover, this last result is essentially independent of the discount factor.

*We thank Stephen Morris, Steven Tadelis, and participants of the Yale Theory discussion group for comments and suggestions. We are especially grateful for the extremely helpful comments of the associate editor and an anonymous referee. Kocherlakota acknowledges the support of NSF SES 0076315 and Cole acknowledges the support of NSF SES 0137421.

1. Introduction

For the last ten years, the analysis of Abreu, Pearce, and Stacchetti (APS) (1990) has been the foundation for the analysis of repeated games with imperfect public monitoring. They develop a recursive representation for the set of equilibrium payoffs in such games. However, this representation relies crucially on players' being able to use strategies that depend on arbitrarily long histories of past events. For example, they demonstrate in some games, an equilibrium with high initial payoffs for all players involves a particular realization of the public signal's triggering infinite repetition of a stage-game equilibrium.

There are at least two concerns with such equilibria. The first is obvious: is it plausible that players keep track of a long sequence of events? The second arises from recent work by Mailath and Morris (2002). They perturb repeated games with public monitoring by adding a small amount of idiosyncratic noise. They show that strategies that exhibit infinite history dependence are not necessarily robust to this type of perturbation. These concerns suggest a natural question: does requiring strategies to exhibit finite history dependence radically change the set of equilibria if we allow the extent of history dependence to be arbitrarily long?

This paper seeks to address this question within a simple repeated game with imperfect public monitoring. We examine the extent to which the set of equilibrium payoffs with infinite-memory strategies is a good approximation to the set of equilibrium payoffs with arbitrarily long finite-memory strategies. (Throughout, we use the terms *memory* and *history dependence* equivalently.) In particular, we look at strongly symmetric perfect public equilibria with memory K: equilibria in which strategies are restricted to depend only on the last K observations of public signals. Define Γ_K to be the set of payoffs of equilibria with memory K. We show that for some specifications of the parameters of the stage game, $\Gamma_K = \Gamma_\infty$ for sufficiently large K. However, for all other specifications of the parameters of the stage game, and for any specification of the discount factor, Γ_K is a singleton for any finite K (even though Γ_∞ is not).

There is a vast literature on equilibria in repeated games when players are restricted to use strategies playable by finite automata (see for example Kalai and Stanford (1988)). A finite automaton can condition play on a finite-valued summary statistic of past play. Our finite memory restriction is much more restrictive because it requires agents to condition play on a finite history of the public signals. This restriction is similar to that in the literature on bounded recall in repeated games with perfect monitoring (for example, Lehrer (1988) and Sabourian (1998)).

Our arguments are closest in spirit to those of Bhaskar (1998). He shows that in an overlapping generations economy, with one player in each cohort, there is a unique equilibrium to the Hammond transfer game when players know only a finite number of periods of past play. We obtain a similar uniqueness result for a class of repeated games with imperfect public monitoring (at least for the case of perfect public equilibria). However, we show that unlike in Bhaskar's setup, the uniqueness result is sensitive to the equilibrium concept being used and to the parameters of the stage game.[1]

2. A Class of Games

We describe a class of repeated games with imperfect public monitoring.

A. Stage Game

Consider the following stage game, which is similar to the partnership game considered by Radner, Myerson and Maskin (1986). There are two players. Player 1 and player 2's action sets are both $\{C, D\}$. Player i's payoffs are given by:

$$\begin{aligned} y - c, \text{ if } a_i &= C \\ y, \text{ if } a_i &= D \end{aligned}$$

The variable y is random, and has support $\{0, 1\}$. The probability distribution of y depends on action choices:

$$\begin{aligned} \Pr(y &= 1 | a_1 = a_2 = C) = p_2 \\ \Pr(y &= 1 | a_i = D, a_j = C) = p_1 \end{aligned}$$

[1] Aumann and Sorin (1989) obtain strong results on cooperation in a game with common interests when the game is perturbed to allow for a small probability that the other player is playing strategies of bounded recall. They show that the perturbed game has an equilibrium in pure strategies, and such equilibria are close to optimal.

$$\Pr(y = 1|a_1 = a_2 = D) = p_0$$

The payoff matrix for the stage game (which is a prisoner's dilemma) is given by

	C	D
C	$(p_2 - c, p_2 - c)$	$(p_1 - c, p_1)$
D	$(p_1, p_1 - c)$	(p_0, p_0)

Throughout, we assume that:

$$\begin{aligned} 1 &> p_2 > p_1 > p_2 - c > p_0 > p_1 - c \\ p_0 &> 0 \end{aligned}$$

These inequalities guarantee that the probability of receiving a good payoff is increasing in the number of players that choose C. They also guarantee that both players playing C is Pareto superior to their playing D, and that both players' playing D is a unique equilibrium of the stage game.

B. Information Structure and Equilibrium

The game stage is infinitely repeated; players have a common discount factor δ, $0 < \delta < 1$. We also assume that there is a public randomizing device; specifically, let $\{\theta_t\}_{t=0}^{\infty}$ be a collection of independent random variables, each uniformly distributed on the unit interval. We define $\theta^t = (\theta_0, ..., \theta_t)$ and $y^t = (y_1, ..., y_t)$.

We assume that player i's action choices are unobservable, but the outcome of y is observable to both players. Hence, player i's history after period t is given by $h_i^t = ((a_{is})_{s=1}^t, y^t, \theta^t)$. The public history after period t is $h^t = (y^t, \theta^t)$. We denote by $y_s(h^t)$ and $\theta_s(h^t)$, $s \leq t$, the realizations of y_s and θ_s in public history h^t. We use the notation (y_s^r, θ_s^r) to represent $(y_t, \theta_t)_{t=s}^r$.

In this world, a behavioral strategy for player i describes the probability of the player's choosing C as a function of his history. (Henceforth, for convenience, we will leave the modifier "behavioral" as implicit in the definition of a strategy.) More formally, a strategy is a collection of mappings $\{\sigma_{it}\}_{t=1}^{\infty}$, where σ_{it} maps the collection of possible period t histories for player i into $[0, 1]$. A *public behavioral strategy* σ_i is a strategy which maps any two histories for player i with the same public history into the same action.

Given these notions of strategies, we restrict attention to strongly symmetric public equilibria, in which both players use the same public strategy. Thus, an equilibrium is a public strategy σ such that σ is a player's optimal strategy, given that the other player is using σ.

C. Finite Memory Equilibria

We are interested in exploring equilibria in which the players' strategies are restricted to depend only in a limited way upon histories. A public strategy with memory K is a public strategy such that $\sigma(h^t) = \sigma(\widehat{h}^t)$ if $y_{t-s}(h^t) = y_{t-s}(\widehat{h}^t)$ and $\theta_{t-s}(h^t) = \theta_{t-s}(\widehat{h}^t)$, for all s such that $0 \leq s \leq \min(K, t) - 1$. Thus, the strategy can only depend on (at most) the last K realizations of the public signals and calendar time. Correspondingly, an equilibrium with memory K is an equilibrium in which the strategy has memory K. (Thus, definitionally, an equilibrium with infinite memory is the same as an equilibrium.)[2]

In any equilibrium, all players receive the same expected utility at any stage of the game. We use the notation Γ_K to refer to the set of payoffs delivered by equilibria of memory K. The key propositions that follow are about the question: does $\lim_{K\to\infty} \Gamma_K = \Gamma_\infty$?

3. Equilibrium Payoffs with Infinite Memory

$>$From APS (1990), we know that Γ_∞ is a closed interval. It is also straightforward to show that the minimax payoff in the stage game is p_0, which is also an equilibrium payoff in the stage game. Hence, the lower bound of Γ_∞ is given by $v_{\min} \equiv p_0$.

What about the upper bound, $v_{\max}$, of Γ_∞? We know from APS (1990) that if $v_{\max} > v_{\min}$, then for some π, there is an equilibrium that delivers $v_{\max}$ of the form:

$$(1) \qquad \sigma(h^t) \quad = \quad 0 \text{ if for some } s \leq t,\ y_s(h^t) = 0 \text{ and } \theta_s(h^t) \geq \pi,$$

[2]Note that in this definition of an equilibrium with memory k, we have not imposed limited recall on the players. Hence, players can contemplate using arbitrary functions of past histories, but choose in equilibrium to use strategies that depend only on the last k lags of the public signal. In contrast, in a game with recall limit k, players can only contemplate using strategies that are measurable with respect to what they have seen in the last k periods. We discuss allowing for bounded recall later in the paper.

$$\sigma(h^t) \;=\; 1 \text{ otherwise.}$$[3]

Verbally, we can think of two possible phases in this equilibrium: a "cooperate" phase and a "non-cooperate" phase. Players start in the cooperate phase, and stay there until they observe $y = 0$ *and* a sufficiently high realization of θ. Then they start playing a permanent "non-cooperate" phase in which they both play D forever. The possibility of switching from the cooperative to the noncooperative phase whenever the outcome is $y = 0$ is in effect a punishment for low output, and this punishment is what induces the players to play C in the cooperation phase even though it is costly.

The continuation payoff in the cooperate phase is $v_{\max}$, and the continuation payoff in the non-cooperate phase is $v_{\min}$. Hence, we can see that:

$$(2) \qquad v_{\max} = p_2[(1-\delta) + \delta v_{\max}] + (1-p_2)\delta(\pi v_{\max} + (1-\pi)v_{\min}) - c(1-\delta)$$

For the strategy to be an equilibrium, it must be true that in the non-cooperate phase, players prefer to play D rather than deviate to C

$$v_{\min} \geq p_1[(1-\delta) + \delta v_{\min}] + (1-p_1)\delta v_{\min} - c(1-\delta)$$

but this is satisfied trivially because $(p_1 - c) < p_0$. As well, it must be true that in the cooperate phase, players prefer to play C rather than deviate to D:

$$(3) \qquad v_{\max} \geq p_1(1-\delta) + \delta v_{\max} + (1-p_1)\delta[\pi v_{\max} + (1-\pi)v_{\min}]$$

Moreover, for $v_{\max}$ to be the maximal element of Γ_∞, the latter inequality must be an equality. Otherwise, we can increase π and thereby increase the value of $v_{\max}$ implied by the flow equation (2), without violating the equilibrium requirement (3).

In the strategy supporting $v_{\max}$, the punishment for realizing $y = 0$, which we denote by v_{pun}, is given by

$$v_{pun} = \pi v_{\max} + (1-\pi)v_{\min}.$$

$>$From the above discussion, it follows that, given (p_2, p_1, c), $(v_{\max}, v_{pun})$ are the solutions to the two equations:

$$\begin{aligned} v_{\max} &= p_2[(1-\delta) + \delta v_{\max}] + (1-p_2)\delta v_{pun} - c(1-\delta) \\ v_{\max} &= p_1(1-\delta + \delta v_{\max}) + (1-p_1)\delta v_{pun} \end{aligned}$$

Since $p_2 - c < p_1$, $v_{pun} < v_{\max}$. Hence, $\Gamma_\infty = [v_{\min}, v_{\max}]$ if and only if $v_{pun} \geq v_{\min} = p_0$ (which is true if and only if δ is sufficiently large). It is tedious but simple to show that this is equivalent to assuming that $p_2 - p_1 - c + \delta p_1 c - \delta p_0 p_2 + \delta p_1 p_0 > 0$.

When we switch to finite memory equilibria, we will find that the key to generating $v_{\max}$ is being able to credibly threaten to punish low output levels with v_{pun} while respecting the memory constraint on the equilibrium strategies.

4. Equilibrium Payoffs with Finite Memory

With finite memory, the structure of the analysis is substantially different. Let $\bar{v}_K$ and $\underline{v}_K$ denote the upper and lower bounds of the set of payoffs when the equilibrium strategies only depend upon at most memory K. Clearly since playing D was an equilibrium of the stage game, playing D forever is an equilibrium with memory K. Hence the lower bound of the payoff set is unchanged and $\underline{v}_K = p_0$.

However, we can no longer enforce the playing of C during the cooperative phase by threatening to punish low output levels with the possibility of switching to playing D forever. The reason is that switching to playing D forever after some event would involve us keeping track of the fact that the event had occurred for all of the subsequent periods. Instead, with memory K, if some event triggers switching from playing C to playing D, then if that event does not recur within the next K periods, the players switch back to C. Therefore, the restriction to memory K strategies has two effects: First, it ties together the probability of switching from the cooperative phase to the noncooperative phase, and vice versa. Second, it makes the incentive constraints during the noncooperative phase more difficult to satisfy since players might have an incentive to deviate in order to make the switch to the cooperative phase more likely. These two restrictions make it harder to support $v_{\max}$ with a memory K strategy because of the difficulties associated with being able to credibly threaten a payoff of v_{pun} when $y = 0$ while respecting the memory constraint.

To illustrate these points, consider the following strategy:

$$\sigma(h^t) = 0 \text{ if for some } t - K + 1 \leq s \leq t,\ y_s(h^t) = 0 \text{ and } \theta_s(h^t) \geq \tilde{\pi}, \tag{4}$$

$$\sigma(h^t) \quad = \quad 1 \text{ otherwise.}$$

Again, verbally, we can think of two possible phases in this equilibrium: a "cooperate" phase and a "non-cooperate" phase. Players start in the cooperate phase, and stay there until they observe $y = 0$ *and* a sufficiently high realization of θ. Then they start playing a non-cooperate phase in which they both play D which lasts until they have not observed both $y = 0$ *and* a sufficiently low realization of θ in any of the last K periods.

This strategy is similar to the infinite memory strategy used to support the best equilibrium payoff: (i) high output realizations cause the cooperate phase to be extended for sure, and (ii) low output realizations can cause the players to switch to the noncooperate phase. It differs in that the noncooperative phase is not permanent. This aspect is an inevitable consequence of (i). But a property like (i) is necessary if high payoff levels are to be achieved. Despite property (i), this strategy has the potential to induce cooperative behavior in the same circumstances as the infinite memory strategy since, as K gets large and $\tilde{\pi}$ approaches zero, the continuation payoff from $y = 0$ in the cooperative phase approaches $v_{\min}$. Hence it would seem to offer the prospect of being able to deliver the appropriate level of v_{pun} through an appropriate choice of $\tilde{\pi}$.

However, with this strategy, players can influence the probability of switching from the noncooperative to the cooperative phase. If the current period's outcome was $y = 0$ and $\theta \leq \tilde{\pi}$, then the likelihood that they will switch back to being in the cooperative phase $K+1$ periods from now is $[p_0 + (1-p_0)(1-\tilde{\pi})]^K$. However, if one of the players deviated and played C instead of D in the next period, the likelihood of switching back to the cooperative phase in $K+1$ periods rises to $[p_1 + (1-p_1)(1-\tilde{\pi})][p_0 + (1-p_0)(1-\tilde{\pi})]^{K-1}$. The possibility of influencing the possibility of switching back to cooperation may induce a deviation at some point during the noncooperative phase and hence undercut the possibility of this strategy credibly threatening $v_{\min}$. We turn next to showing that this can lead to the set of equilibrium payoffs under finite memory being much more restricted than under infinite memory.

5. Finite Memory: A Non-Convergence Result

Our first result is to show that there exists an open set of parameters such that $\lim_{K\to\infty}\Gamma_K \neq \Gamma_\infty$. In fact, as the following proposition shows, if p_1 is sufficiently close to p_2, always playing D is the only equilibrium with memory K, for any finite K.

Proposition 1. *If:*

$$p_1 > 0.5(p_2 + p_0) \tag{5}$$

then $\Gamma_K = \{p_0\}$ *for all* K.

To understand the logic of this proposition, consider two different histories of length K of the public signal y:

$$\begin{aligned}(y_{t-K}, ..., y_{t-1}) &= (1, 1, ..., 1, 1)\\ (y'_{t-K}, ..., y'_{t-1}) &= (0, 1.., 1, 1)\end{aligned}$$

Trying to support an equilibrium other than always playing D with K memory strategies means that we must be able to generate different outcomes from two histories such as these. For example, if (4) is an equilibrium, a player must find it optimal to play C with some probability after the first type of history, and play C with some other probability after the second type of history. But these histories differ only in their last entry, and this difference disappears next period. Hence, the players' continuation payoffs are the *same* function of y_t after both these histories. In order to generate these two different continuation equilibria, we need to be able to choose continuation values v_1 and v_0 so as to make it an equilibrium for players to choose C or choose D. The essence of the proposition is that if $p_1 > (p_2 + p_0)/2$, this cannot be done.

Proof. The first part is that the only equilibrium with memory 1 is to always choose D for all public histories. The second part is to assume inductively that the only equilibrium with memory $(K-1)$ is to always choose D, and show by induction, the only equilibrium with memory K is to always choose D.

Part 1: If always playing D is not the only equilibrium with memory 1, then there exists some period t such that $\sigma(h^{t-2}, y_{t-1}, \theta_{t-1}) = q'$ and $\sigma(h^{t-2\prime}, y'_{t-1}, \theta'_{t-1}) = q$, where

$1 \geq q' > q \geq 0$. Define $v_{1,t}$ to be the expected continuation payoff in period $(t+1)$ if $y_t = 1$, and $v_{0,t}$ to be the expected continuation payoff if $y_t = 0$ (where the expectations are over θ_t). Then, because $q' > 0$, we know that in period t either player must weakly prefer playing C to playing D, given that the other player chooses q'. This implies that:

$$\begin{aligned} & q'[p_2(1-\delta+\delta v_{1t}) + (1-p_2)\delta v_{0t} - c(1-\delta)] \\ & +(1-q')[p_1(1-\delta+\delta v_{1t}) + (1-p_1)\delta v_{0t} - c(1-\delta)] \\ \geq\ & q'[p_1(1-\delta+\delta v_{1t}) + (1-p_1)\delta v_{0t}] \\ & +(1-q')[p_0(1-\delta+\delta v_{1t}) + (1-p_0)\delta v_{0t}] \end{aligned}$$

This implies that:

$$q'[(p_2-p_1)(1-\delta+\delta v_{1t}-\delta v_{0t})] + (1-q')[(p_1-p_0)(1-\delta+\delta v_{1t}-\delta v_{0t})] \geq c(1-\delta)$$

Similarly, because $q < 1$, we know that in period t, either player must weakly prefer playing D to playing C, given that the other player chooses q. This implies that:

$$q[(p_2-p_1)(1-\delta+\delta v_{1t}-\delta v_{0t})] + (1-q)[(p_1-p_0)(1-\delta+\delta v_{1t}-\delta v_{0t})] \leq c(1-\delta)$$

By combining these two inequalities, we know that $(1-\delta+\delta v_{1t}-\delta v_{0t}) > 0$ and:

$$(q'-q)[(p_2-p_1)(1-\delta+\delta v_{1t}-\delta v_{0t}) - (p_1-p_0)(1-\delta+\delta v_{1t}-\delta v_{0t})] \geq 0$$

which in turn implies that:

$$(p_2-2p_1+p_0)(1-\delta+\delta v_{1t}-\delta v_{0t}) \geq 0$$

which violates $p_1 > (p_2+p_0)/2$.

Part 2: Now, we show that in any equilibrium with memory K, the strategies must be independent of the Kth lag of the public signals. Suppose not, and:

$$\begin{aligned} \sigma(y_{t-K}^{t-1}, \theta_{t-K}^{t-1}) &= q \\ \sigma(y_{t-K}^{t-1\prime}, \theta_{t-K}^{t-1\prime}) &= q' \\ (y_{t-K-1}^{t-1}, \theta_{t-K-1}^{t-1}) &= (y_{t-K-1}^{t-1\prime}, \theta_{t-K-1}^{t-1\prime}) \end{aligned}$$

where $(y_s^r, \theta_s^r) = (y_i, \theta_i)_{i=s}^r$. and $1 \geq q' > q \geq 0$. Define:

$$v_1 = E_{\theta_t} v_t(y_{t-K}^{t-1}, \theta_{t-K}^{t-1}, y_t = 1, \theta_t)$$
$$v_0 = E_{\theta_t} v_t(y_{t-K}^{t-1\prime}, \theta_{t-K}^{t-1\prime}, y_t = 0, \theta_t)$$

It follows that at history $(y_{t-K}^{t-1}, \theta_{t-K}^{t-1})$, playing D is weakly preferred to playing C, given that the other player chooses q. The above analysis then implies that:

$$q[(p_2 - p_1)(1 - \delta + \delta v_{1t} - \delta v_{0t})] + (1 - q)[(p_1 - p_0)(1 - \delta + \delta v_{1t} - \delta v_{0t})] \leq c(1 - \delta)$$

Similarly, at history $(y_{t-K}^{t-1\prime}, \theta_{t-K}^{t-1\prime})$, playing C is weakly preferred to playing D, given that the other player chooses q'. The analysis in Part 1 then implies that:

$$q'[(p_2 - p_1)(1 - \delta + \delta v_{1t} - \delta v_{0t})] + (1 - q')[(p_1 - p_0)(1 - \delta + \delta v_{1t} - \delta v_{0t})] \geq c(1 - \delta).$$

Together, these inequalities contradict $p_2 - p_1 > p_1 - p_0$.

The proposition then follows inductively. $\square$

It is worth emphasizing that this proposition holds regardless of the size of δ.

In our setup, the signals enter directly into the players' payoffs. An alternative approach would be to follow Abreu, Milgrom, and Pearce (1991) and disentangle the signals from the payoffs of the stage game. Suppose for example the stage game payoffs are:

	C	D
C	(m, m)	$(-b, m+g)$
D	$(m+g, -b)$	$(0, 0)$

and the signals are as described above. Then, we can show that Γ_∞ is independent of b and p_0. Intuitively, the upper endpoint Γ_∞ is determined by the social cost of the punishments necessary to support (C, C) as an equilibrium. The parameters (b, p_0) have no effect on deterring a player from deviating to D. In contrast, with finite memory, it becomes problematic to support playing (D, D) forever as a punishment. The parameters (b, p_0) have to be chosen so that a player does not find it in his interest to deviate to playing C. Hence, Γ_K depends on (b, p_0) while Γ_∞ does not.

6. Finite Memory: Convergence

We now show that if $p_1 \leq (p_2 + p_0)/2$ and $v_{pun} > v_{\min}$, then there exists K^* such that if $K \geq K^*$, then $\Gamma_K = \Gamma_\infty$.

A. A Convergence Result for Maximal Payoffs

To show this, we first show that for K sufficiently large, we can construct an equilibrium with memory K that has payoff $v_{\max}$.

To do so, consider the following strategy with memory K. The strategy is of the form:

$$\sigma_K^{\max}(h^t) = 0 \text{ if } \theta_{t-k^*+1}(h^t) \geq \pi_K, \text{ where}$$
$$k^* = \min\left\{k \in \{1, ..., K\} : y_{t-k+1}(h^t) = 0\right\} \tag{6}$$
$$\sigma_K^{\max}(h^t) = 1 \text{ otherwise.}$$

Verbally, we can think of two possible phases in this equilibrium: a "cooperate" phase and a "non-cooperate" phase. Players start in the cooperate phase, and stay there until they observe $y = 0$ *and* a sufficiently high realization of θ. Then they start playing a "non-cooperate" phase in which they both play D which lasts until either they have not observed $y = 0$, or in the most recent period in which $y = 0$ the realization of θ is sufficiently low.

Remark: Note that under this strategy, agents switch from a non-cooperate phase to a cooperate phase if they see $y_{t-j} = 1$ for $j = 1, ...K$, *or* if they see $y_t = 0$ and $\theta_t < \pi_K$. Hence, they switch back to cooperate with a higher probability than if they play (4). This means that the punishment phase is somewhat less severe for a given level of memory K. Nonetheless, (6) is more analytically tractable in proving the limiting result that follows.

Suppose that a strategy of this form delivers a payoff $v_{\max}$. Given a non-cooperate history with cutoff k^*, define X_k to be the continuation payoff, where $k = K - k^* + 1$. Under this definition, X_k is the continuation payoff in a non-cooperate phase if the players are going to switch to a cooperate phase after seeing $y_{t+j} = 1$ for all $j = 1, .., k$. Then, it must be true that:

$$v_{\max} = p_2(1 - \delta + \delta v_{\max}) + (1 - p_2)\delta v'_{pun} - c(1 - \delta)$$

where v'_{pun} is given by

(7) $$v'_{pun} = \pi_K v_{\max} + (1 - \pi_K) X_K$$

and X_k satisfies the recursive relationship:

$$\begin{aligned} X_k &= p_0(1 - \delta + \delta X_{k-1}) + (1 - p_0)\delta v'_{pun} \\ X_0 &= v_{\max} \end{aligned}$$

It follows that:

$$X_k = (p_0(1-\delta) + (1-p_0)\delta v'_{pun}) \sum_{i=0}^{k-1} (p_0\delta)^i + \delta^k p_0^k v_{\max}$$

For this strategy to be a viable equilibrium, we need to verify three things. First, we need to make sure that the players find it weakly optimal to play C in the cooperate phase of the equilibrium. Note, though, from the definition of v_{pun}, that $v'_{pun} = v_{pun}$ and also that:

$$v_{\max} = p_1(1 - \delta + \delta v_{\max}) + (1 - p_1)\delta v_{pun}$$

and so in the cooperate phase, players are indifferent between playing C or not.

Second, we need to verify that players are willing to play D in the non-cooperate phase of the equilibrium. Consider a history $h^{t-1} = (y_{t-1}, ..., y_{t-K})$ in which $y_{t-k^*} = 0$, $\theta_{t-k^*} > \pi$, and $y_{t-k} = 1$ for all $k < k^*$. The equilibrium payoff in this history is equal to X_{K-k^*+1}. Thus, for each k^* if playing D is weakly preferred to C it must be the case that

$$X_{K-k^*+1} \geq p_1(1 - \delta + \delta X_{K-k^*}) + (1 - p_1)\delta v_{pun} - c(1 - \delta),$$

where $X_0 \equiv v_{\max}$. Since X_{K-k^*+1} satisfies the recursive equation:

$$X_{K-k^*+1} = p_0(1 - \delta + \delta X_{K-k^*}) + (1 - p_0)\delta v_{pun},$$

we can conclude that:

$$(p_1 - p_0)(1 - \delta + \delta X_{K-k^*} - \delta v_{pun}) \leq c(1 - \delta).$$

Thus, to make sure that players want to play D in the non-cooperate phase, we must verify the above inequality for all $k^* \in \{1, ..., K\}$.

We verify this inequality as follows. We know that:

$$p_2(1-\delta+\delta v_{\max})+(1-p_2)\delta v_{pun}-c(1-\delta)=p_1(1-\delta+\delta v_{\max})+(1-p_1)\delta v_{pun},$$

or, equivalently:

$$\delta(v_{\max}-v_{pun})=c(1-\delta)/(p_2-p_1)-1$$

It is trivial to see that X_k is decreasing in k. Hence, it follows that for any k^*:

$$\begin{aligned} & (1-\delta+\delta X_{K-k^*}-\delta v_{pun}) \\ \leq\ & (1-\delta+\delta X_0-\delta v_{pun}) \\ =\ & (1-\delta+\delta v_{\max}-\delta v_{pun}) \text{ (by definition of } X_0) \\ =\ & c(1-\delta)/(p_2-p_1) \\ \leq\ & c(1-\delta)/(p_1-p_0) \end{aligned}$$

Thus, because $(p_2-p_1)\geq(p_1-p_0)$, players are willing to play D throughout the punishment phase.

Finally, we need to find K so that $0\leq\pi_K$. Again, X_K is decreasing in K and hence we can conclude that

$$\pi_K=\frac{v_{pun}-X_K}{v_{\max}-X_K}$$

is increasing in K. Furthermore, note that

$$\begin{aligned} lim_{K\to\infty}X_K &= \left[p_0+(1-p_0)\delta v_{pun}\right]/(1-p_0\delta) \\ &= \frac{(1-\delta)}{(1-p_0\delta)}p_0+\frac{(1-p_0)\delta}{(1-p_0\delta)}v_{pun}, \end{aligned}$$

and so $\lim_{K\to\infty}X_K$ is a convex combination of p_0 and v_{pun}. This implies that if $v_{pun}>p_0$, then there exists K^*, such that for all $K\geq K^*$, $X_K<v_{pun}$.

This analysis verifies the following proposition.

PROPOSITION 2. *If $(p_2+p_0)/2\geq p_1$, and $v_{pun}>v_{\min}$, then there exists K^* such that for all $K\geq K^*$, the maximal element of Γ_K is $v_{\max}$.*

The crux of this proposition is that a K-period non-cooperate phase, if $K \geq K^*$, is sufficiently harsh to induce cooperation. Crucially, as long as $(p_2 + p_0)/2 \geq p_1$, players are willing to play non-cooperate.

B. A Convergence Result for the Equilibrium Payoff Set

We have seen that under the conditions of Proposition 2, the maximal element of Γ_K is $v_{\max}$ for sufficiently large K. The minimal element of Γ_K is $v_{\min}$ for any K. But is Γ_K connected or are there holes in Γ_K? When $K = \infty$, we can use the initial draw of the public randomization device to create any payoff between the endpoints of Γ_∞. But with finite memory, this permanent randomization between equilibria is no longer possible. In this subsection, we show that when the conditions of proposition 2 are satisfied, $\Gamma_K = \Gamma_\infty$ for K sufficiently large.

Suppose that the conditions of proposition 2 are satisfied, and $K > K^*$. Let $\gamma \in \Gamma_\infty$, and consider the following specification of strategies. Let $\tau \in \{0, 1, 2, ..., \infty\}$ be such that

$$p_0(1 - \delta^\tau) + \delta^\tau v_{\max} \leq \gamma < p_0(1 - \delta^{\tau+1}) + \delta^{\tau+1} v_{\max}.$$

Let π_τ such that

$$\pi_\tau \left[p_0(1 - \delta^\tau) + \delta^\tau v_{\max}\right] + (1 - \pi_\tau)\left[p_0(1 - \delta^{\tau+1}) + \delta^{\tau+1} v_{\max}\right] = \gamma.$$

Denote the strategy that supports payoff $v_{\max}$ by $\sigma_{K^*}^{\max}$. Then we can define the strategy σ_γ that supports payoff γ as follows.

$$\sigma_\gamma(y^{t-1}, \theta^{t-1}) = 0 \text{ for all } t < \tau.$$

$$\sigma_\gamma(y^{\tau-1}, \theta^{\tau-1}) = 1 \text{ if } \theta_{\tau-1} \leq \pi_\tau$$

$$\sigma_\gamma(y^{\tau-1}, \theta^{\tau-1}) = 0 \text{ if } \theta_{\tau-1} > \pi_\tau$$

$$For\ t \geq \tau,$$

$$\sigma_\gamma(y^t, \theta^t) = \sigma_{K^*}^{\max}(y_\tau^t, \theta_\tau^t) \text{ if } \theta_{\tau-1} \leq \pi_\tau$$

$$\sigma_\gamma(y^t, \theta^t) = \sigma_{K^*}^{\max}(y_{\tau+1}^t, \theta_{\tau+1}^t) \text{ if } \theta_{\tau-1} > \pi_\tau$$

The basic idea of this strategy is that the players play D through period $(\tau - 1)$ (here, we exploit the fact that players can condition their play on calendar time). Then, in period t, they switch to playing $\sigma_{\max}$ if $\theta_{\tau-1}$ is low. Otherwise, they play D in period t, and switch to playing $\sigma_{\max}$ in period $(t+1)$.

By construction, σ_γ delivers payoff γ. We need to verify that σ_γ is indeed an equilibrium with memory K. Note that in the histories in which the strategy specifies that the players choose D, their actions have no effects on future payoffs. Hence, playing D is weakly optimal. Also, we know that $\sigma_{K^*}^{\max}$ is an equilibrium, so that playing according to $\sigma_{K^*}^{\max}$ is weakly optimal whenever the strategy makes this specification.

We still need to verify that σ_γ is a strategy with memory K. Since $\sigma_{K^*}^{\max}$ is a strategy with memory $K^* < K$, it follows that $\sigma_{K^*}^{\max}(y_\tau^t, \theta_\tau^t) = \sigma_{K^*}^{\max}(y_{\tau+1}^t, \theta_{\tau+1}^t)$ for $t \geq (\tau + K^*)$. Thus, the realization of θ_τ does not affect play after period $(\tau + K^*)$, and σ_γ is a strategy with memory $K > K^*$.

7. Discussion

Throughout, we have assumed that players have perfect recall. This means that, while equilibrium strategies are required to be functions of K lags of past history, players can contemplate deviations from equilibrium play that are arbitrary functions of past history. In contrast, if players' recall is limited, then they can only contemplate using strategies that are functions of the last K periods of history.

Let Γ_K^{br} be the set of pure strategy perfect public equilibrium payoffs when players have recall limit K. Then, we can demonstrate a result analogous to Proposition 1: if $(p_2+p_0)/2 < p_1$, then $\Gamma_K^{br} = \{p_0/(1-\delta)\}$ for all K. The proof is identical to that of Proposition 1. Intuitively, in the proof of Proposition 1, we eliminate the possibility of other equilibria by contemplating the possibility of players' deviating to strategies consistent with bounded recall.

We cannot prove a direct analogy to Proposition 2. However, it is simple to see that $\Gamma_K^{br} \supseteq \Gamma_K$ (because there are fewer possible deviations with recall K). Hence, we know that, under the assumptions of Proposition 2:

$$\lim_{K\to\infty} \Gamma_K^{br} \supseteq \Gamma_\infty$$

Until now, because we want to understand the implications of Mailath and Morris (2002), we have used perfect public equilibrium as the equilibrium concept. In such an equilibrium, players' strategies are a function only of public history. In contrast, in a sequential equilibrium, players' strategies can be arbitrary functions of both public and private history. Let Γ_∞^{se} be the set of pure strategy sequential equilibrium payoffs (with no history restrictions on the strategies). From Abreu, Pearce, and Stacchetti (1990), we know that $\Gamma_\infty^{se} = \Gamma_\infty$.[4] Does this equivalence hold up when we impose a finite memory or bounded recall restriction on the equilibrium? Unfortunately, we have no results for the finite memory case. However, there is a surprisingly strong result for bounded recall.

Let $\Gamma_K^{br,se}$ be the set of pure strategy sequential equilibrium payoffs when players have recall limit K (so that they can only contemplate using strategies that are functions of the last K periods of history). Rather remarkably, $\Gamma_1^{br,se} \supseteq \Gamma_\infty^{se}$ for any parameter settings.[5] Consider any element of Γ_∞^{se}. From results in APS(1990), we know that it can be supported as an equilibrium payoff by a public strategy of the form:

$$\begin{aligned} \sigma(y^t, \theta^t, (a_i)^t) &= 0 \text{ if } \theta_s \geq \pi \text{ and } y_s = 0 \text{ for some } s \leq t \\ &= 1 \text{ otherwise} \end{aligned}$$

This strategy is public because it does not depend on $(a_i)^t$. Then, consider the following strategy:

$$\begin{aligned} \sigma^*(y^t, \theta^t, (a_i)^t) &= 0 \text{ if } a_{it} = D \text{ OR if } y_t = 0 \text{ and } \theta_t \geq \pi \\ &= 1 \text{ otherwise} \end{aligned}$$

Note that σ^* is a (non-public) strategy with memory 1. Our goal is to show that σ^* is a strongly symmetric sequential equilibrium when the players have recall 1.

Why is this true? The key is to make sure that the players play D when they are supposed to. Suppose that in period $(t+1)$, player i sees $a_{it} = D$. There are then two possible scenarios that are consistent with equilibrium play. First, it could be that $y_t = 0$ and $\theta_t \geq \pi$. Then, player i concludes that player j will play D this period, and in every period thereafter. It is optimal for player i to do the same.

[4] This equivalence is not true when we consider mixed strategies (see Kandori and Obara 2000 or Mailath, Mathews, and Sekiguchi 2002).

[5] We thank Stephen Morris for pointing this out to us.

The other possibility is that player i may have seen $y_{t-s} = 0$ and $\theta_{t-s} \geq \pi$ for some $s > 0$, but does not remember doing so. Under this scenario, a_{jt} is also D, and player j will play D in this and all future periods. Again, it is optimal for player i to play D today.

The trick here is that a player's past action serves as a summary statistic that encodes whether game play is in a "cooperate" or "non-cooperate" phase. Implicitly, one lag of private actions encodes the relevant portion of the full public history. Hence, $\Gamma_1^{br,se} \supseteq \Gamma_\infty^{se}$.

As the referee pointed out to us, this reasoning does not carry over to the case of finite memory strategies. In that case, the player knows the full history of past play. Suppose that in period $(t+1)$, player i sees $a_{it} = D$, and he sees $y_{t-s} = 0$ or $\theta_{t-s} < \pi$ for all $s \geq 0$. There is no evidence to indicate that player j has deviated from equilibrium play; player i therefore assumes that player j has never played D in the past, and will not play D in this period. It follows that it is suboptimal for player i to play D today.

References

[1] Abreu, D., Pearce, D., and Stacchetti, E., 1990, Toward a theory of discounted repeated games with imperfect monitoring, *Econometrica* 58, 1041-1063.

[2] Aumann, Robert and Sorin, Sylvain, 1989, Cooperation and Bounded Recall, *Games and Economic Behavior* 1, 5-39.

[3] Bhaskar, V., 1998, Informational constraints and the overlapping generations model: Folk and Anti-Folk theorems, *Review of Economic Studies* 65, 135-149.

[4] Kalai, E., and Stanford, W., Finite rationality and interpersonal complexity in repeated games, *Econometrica* 56, 397-410.

[5] Kandori, M. and Obara, I., 2000, Efficiency in Repeated Games Revisted: The Role of Private Strategies," unpublished.

[6] Lehrer, E., 1988, Repeated games with stationary bounded recall strategies, *Journal of Economic Theory* 46, 130-44.

[7] Mailath, G., and Morris, S., 2002, Repeated games with almost-perfect monitoring, *Journal of Economic Theory* 102, 189-228.

[8] Mailath, G. ,Mathews, S. and Sekiguchi, T., 2002, Private Strategies in Finitely Repeated Games with Imperfect Public Monitoring," *Contributions to Theoretical Economics,* 2(1).

[9] Radner, R., Myerson, R., and Maskin, E., 1986, An example of a repeated partnership game with discounting and uniformly inefficient equilibria, *Review of Economic Studies* 53, 59-70.

[10] Sabourian, H., 1998, Repeated games with M-period bounded memory (pure strategies), *Journal of Mathematical Economics 30,* 1-35.