# Purification in the Infinitely-Repeated Prisoners' Dilemma*

V. Bhaskar[†] George J. Mailath[‡] Stephen Morris[§]

July 13, 2006
revised October 26, 2007

**Abstract**

This paper investigates the Harsanyi (1973)-purifiability of mixed strategies in the repeated prisoners' dilemma with perfect monitoring. We perturb the game so that in each period, a player receives a private payoff shock which is independently and identically distributed across players and periods. We focus on the purifiability of one-period memory mixed strategy equilibria used by Ely and Välimäki (2002) in their study of the repeated prisoners' dilemma with private monitoring. We find that any such strategy profile is not the limit of one-period memory equilibrium strategy profiles of the perturbed game, for almost all noise distributions. However, if we allow infinite memory strategies in the perturbed game, then any completely-mixed equilibrium is purifiable. *Keywords:* Purification, belief-free equilibria, repeated games. *JEL Classification Numbers:* C72, C73.

## 1. Introduction

Harsanyi's (1973) purification theorem is one of the most compelling justifications for the study of mixed equilibria in finite normal form games. Under this justification, the complete-information normal form game is viewed as the limit of a sequence of incomplete-information games, where each player's payoffs are subject to private shocks. Harsanyi proved that every equilibrium (pure or mixed) of the original game is the limit of equilibria of close-by games with incomplete information. Moreover, in the incomplete-information games, players have essentially strict best replies, and so will

*Mailath is grateful for support from the National Science Foundation under grants #SES-0095768 and #SES-0350969. Morris is grateful for support from the John Simon Guggenheim Foundation, the Center for Advanced Studies in the Behavioral Sciences and National Science Foundation Grant #SES-0518929.

[†]University College London
[‡]University of Pennsylvania
[§]Princeton University

not randomize. Consequently, a mixed strategy equilibrium can be viewed as a pure strategy equilibrium of any close-by game of incomplete information. Harsanyi's (1973) argument exploits the regularity (a property stronger than local uniqueness) of equilibria of "almost all" normal form games. As long as payoff shocks generate small changes in the system of equations characterizing equilibrium, the regularity of equilibria ensures that the perturbed game has an equilibrium close to any equilibrium of the unperturbed game.[1]

Very little work has examined purification in dynamic games. Even in finite extensive games, generic local uniqueness of equilibria may be lost when we build in natural economic features into the game, such as imperfect observability of moves and time separability of payoffs. Bhaskar (2000) has shown these features may lead to a failure of local uniqueness and purification: For a generic choice of payoffs, there is a continuum of mixed strategy equilibria, none of which are the limit of the pure strategy equilibria of a game with payoff perturbations.

For infinitely repeated games, the bootstrapping nature of the system of equations describing many of the infinite horizon equilibria is conducive to a failure of local uniqueness of equilibria. We study a class of symmetric one-period memory mixed strategy equilibria used by Ely and Välimäki (2002) in their study of the repeated prisoners' dilemma with private monitoring. This class fails local uniqueness quite dramatically: there is a two dimensional manifold of equilibria.

Our motivation for studying the purifiability of this class of strategies comes from the recent literature on repeated games with private monitoring. Equilibrium incentive constraints in games with private monitoring are difficult to verify because calculating best replies typically requires understanding the nature of players' beliefs about the private histories of other players. Piccione (2002) showed that by introducing just the right amount of mixing *in every period*, a player's best replies can be made independent of his beliefs, and thus beliefs become irrelevant (and so the equilibrium is *belief-free*, see remark 1).[2] This means in particular that these equilibria of the perfect monitoring game trivially extend to the game with private monitoring. Piccione's (2002) strategies depend on the infinite history of play. Ely and Välimäki (2002) showed that it suffices to consider simple strategies which condition only upon one period memory of both players' actions. These strategies again make a player indifferent between his actions regardless of the action taken by the other player, and thus a player's incentives do not change with his beliefs. Kandori and Obara (2006) also use such strategies to obtain

[1] See Govindan, Reny, and Robson (2003) for a modern exposition and generalization of Harsanyi (1973). A brief introduction can also be found in Morris (forthcoming).

[2] This was not the first use of randomization in repeated games with private monitoring. A number of papers construct nontrivial equilibria using initial randomizations to instead generate uncertainty over which the players can then update(Bhaskar and Obara (2002), Bhaskar and van Damme (2002), and Sekiguchi (1997)).

stronger efficiency results via private strategies in repeated games with imperfect public monitoring.

At first glance, the equilibria of Piccione (2002) and Ely and Välimäki (2002) involve unreasonable randomizations: in some cases, a player is required to randomize differently after two histories, even though the player has identical beliefs over the continuation play of the opponent. Moreover, the randomizations involve a delicate intertemporal trade-off. While there are many ways of modeling payoff shocks in a dynamic game, these shocks should not violate the structure of dynamic game. In repeated games, a reasonable constraint is that the payoffs shocks should be independently and identically distributed over time, and moreover, the period $t$ shock should only be realized at the beginning of period $t$. Our question is: Do the delicate intertemporal trade-offs survive these independently and identically distributed shocks?

Our results show that, in the repeated game with perfect monitoring, none of the Ely-Välimäki equilibria can be purified by one-period memory strategies. But they can be purified by infinite horizon strategies, i.e., strategies that are no simpler than those of Piccione (2002). We have not resolved the question of whether they can be purified by strategies with finite memory greater than one.

However, while equilibria of the unperturbed perfect monitoring game are automatically equilibria of the unperturbed private monitoring game, our purification arguments do *not* automatically extend to the private monitoring case. We conjecture—but have not been able to prove—that in the repeated game with *private* monitoring all the Ely-Välimäki equilibria will not be purifiable with finite history strategies but will be purifiable with infinite history strategies.

The paper is organized as follows. The next Section introduces belief-free equilibria and purifiability using a simple example. In Section 3, we review the completely mixed equilibria of the repeated prisoners' dilemma introduced by Ely and Välimäki (2002). The negative purification result for one-period history strategies is in Section 4. In Section 5, we present the positive purification result for infinite history strategies. Finally, in Section 6, we briefly discuss possible extensions and the private monitoring case.

## 2. An Introductory Example

Before discussing the repeated prisoners' dilemma, we present a simpler game (in which only one of the players is long-lived) to introduce belief-free equilibria and the issues underlying their purification.[3,4]

[3] The example is also special in that the incentive of the long-lived player to play $H$ in the stage game is independent of the behavior of the short-lived player (see footnote 6). Our analysis of the prisoners' dilemma does not have this feature.

[4] Mailath and Samuelson (2006, Section 7.6) illustrate equilibrium constructions with long-lived and short-lived players using this example.

|   | $h$ | $\ell$ |
|---|---|---|
| $H$ | $2,3$ | $0,2$ |
| $L$ | $3,0$ | $1,1$ |

Figure 1: The product-choice game. Player 1, the row player, is long-lived, while the other player is short-lived.

### 2.1. The unperturbed game

The stage game is illustrated in Figure 1. We think of the row player (who is long-lived) as a firm choosing between high ($H$) and low ($L$) effort and the column player (who is short-lived) as a customer choosing between a high ($h$) or low ($\ell$) priced product. Since $L$ is strictly dominant in the stage game, $L\ell$ is its unique Nash equilibrium.

In the repeated game, the long-lived player has discount factor $\delta$. In each period there is a new column player who (being short-lived) myopically optimizes. The game has perfect monitoring: in period $t$, both the long-lived player and the the short-lived player (or, more specifically, the short-lived player of period $t$) know the history of play $h^t \in (\{H, L\} \times \{h, \ell\})^t$.

Since the long-lived player has a myopic incentive to play $L$, intertemporal incentives are needed to induce the long-lived player to play $H$. For example, as long as $\delta > 1/2$, the trigger profile using Nash reversion is an equilibrium. In this equilibrium, the long-lived player is induced to play $H$ (with a corresponding choice of $h$ by the short-lived player) by the threat that a deviation to $L$ triggers permanent play of the myopic Nash equilibrium $L\ell$.[5] This specification of behavior is an example of histories coordinating continuation play: The continuation play after any history in which the long-lived player has always played $H$ is described by the trigger profile, while continuation play after every other history is described by permanent myopic Nash.

Our interest is in a different class of equilibrium, called *belief-free*. These equilibria are of particular importance in the study of private monitoring games (see remark 1). In this class, the long-lived player randomizes uniformly over $\{H, L\}$ in each period independently of history. This randomization makes the short-lived player indifferent between $h$ and $\ell$ in each period, allowing us to specify behavior for the short-lived player providing intertemporal incentives for the long-lived player. Let $p^{a_1}$ be the probability the short-lived player puts on $h$ after the long-lived player's play of $a_1 \in \{H, L\}$ in the previous period.[6] Let $V_1^p(a_1)$ be the long-lived player's expected value from the action

[5] Note that the short-lived player is always playing a myopic best reply.

[6] If the long-lived player's incentive in the stage game to play $H$ depended on the behavior of the short-lived player, the short-lived player's randomization depends on the previous period's realized actions of both players.

$a_1$, when the short-lived player plays $h$ with probability $p$. Then,

$$V_1^p(H) = (1-\delta)2p + \delta V^{p^H} \tag{1}$$

and

$$V_1^p(L) = (1-\delta)(2p+1) + \delta V^{p^L}, \tag{2}$$

where $V^{p'} = \frac{1}{2}V_1^{p'}(H) + \frac{1}{2}V_1^{p'}(L)$ (recall that the long-lived player is randomizing uniformly over $\{H, L\}$). In order to be willing to randomize, the long-lived player must be indifferent between $H$ and $L$, and so $V_1^p(H) = V_1^p(L) = V^p$ for $p = p^H, p^L$.[7] Equating the right sides of (1) and (2) yields

$$V^{p^H} = V^{p^L} + \frac{(1-\delta)}{\delta}. \tag{3}$$

Setting $p = p^H$ in (1), and $p = p^L$ in (2), and solving gives

$$V^{p^H} = 2p^H \qquad \text{and} \qquad V^{p^L} = 1 + 2p^L, \tag{4}$$

and so (3) and (4) implies

$$p^L = p^H - \frac{1}{2\delta}. \tag{5}$$

For $\delta > 1/2$, we thus have a one-dimensional manifold of one-period memory equilibria indexed by $p^H \in [1/(2\delta),\ 1]$, with $p^L$ determined by (5). There is an additional (minor) indeterminacy arising from the lack of restrictions on the short-lived player behavior in the initial period; without loss of generality we assume that the short-lived player plays $p^H$ in the initial period.

**Remark 1 (Private Monitoring and Belief-Free Profiles Equilibria)** A similar construction yields an equilibrium in games with private monitoring. Suppose actions are private, and each player observes a *noisy* private signal $\hat{a}_i$ of the other player's action $a_i$ at the end of the period. Suppose the signals are independent, conditional on the action profile. We assume that the short-lived player of period $t$ knows the history of private signals observed by earlier short-lived players. The payoffs in Figure 1 are now *ex ante* stage game payoffs, derived from an underlying *ex post* stage game game, where player $i$'s payoff is a function of $i$'s action and private signal realization only (so that $i$'s payoff conveys no additional information beyond that conveyed by $i$'s private signal).

Since there are no public histories to coordinate play and the monitoring is conditionally independent, $H$ choices cannot be supported using trigger strategies. However,

[7] This implies $V_1^p(H) = V_1^p(L) = V^p$ for all $p \in [0, 1]$.

when the signals about the long-lived player's actions are sufficiently accurate, there are one-period memory profiles in which the long-lived player plays $H$ in *every* period with positive probability.

Just as in the game with perfect monitoring, the long-lived player randomizes uniformly over $\{H, L\}$ in each period independently of history. This randomization again makes the short-lived player indifferent between $h$ and $\ell$ in each period, allowing us to specify behavior for the short-lived player providing intertemporal incentives for the long-lived player. Similar calculations to (1) through (5) show that there is again a one-dimensional manifold of equilibria (Mailath and Samuelson, 2006, section 12.5): After observing the signal $\hat{a}_1$, the short-lived player chooses $h$ with probability $p^{\hat{a}_1}$ so that

$$p^{\hat{L}} = p^{\hat{H}} - \frac{1}{2\delta(1-2\eta)},$$

where $\eta > 0$ is the noise in the monitoring,

$$\eta = \Pr(\hat{L} \mid Ha_2) = \Pr(\hat{H} \mid La_2) \quad \text{and} \quad 1-\eta = \Pr(\hat{H} \mid Ha_2) = \Pr(\hat{L} \mid La_2).$$

These randomizations are chosen so that the long-lived player is indifferent between $H$ and $L$, *irrespective* of the private history that 2 had observed. This implies that the optimality of the long-lived player's behavior can be verified without calculating the long-lived player's beliefs about the short-lived player's continuation play. For this reason, such equilibria are called *belief-free* (Ely, Hörner, and Olszewski, 2005). ♦

In belief-free equilibria the short-lived player, while always indifferent over her actions, randomizes differently after different histories (a property shared by the belief-free equilibria described in remark 1). Since the long-lived player is randomizing uniformly over $\{H, L\}$ independently of history, the short-lived player's beliefs over the play of the long-lived player are *independent* of these different histories.

### 2.2. The perturbed game

We now investigate the extent to which the requisite randomizations in the belief-free equilibria can be "justified" as the limit of essentially strict equilibria of close-by incomplete information games. In other words, are such equilibria Harsanyi (1973)-purifiable? In keeping with the spirit of repeated games, we perturb the stage game, as illustrated in Figure 2. Player $i$'s payoff shock $z_t^i$ is private to $i$, realized in period $t$. We assume $z_t^i$ is independent across players and histories, and (for simplicity in this section) uniformly distributed on $[0, 1]$. The infinitely-repeated perfect monitoring game with stage game displayed in Figure 2 is denoted $\Upsilon(\varepsilon_1, \varepsilon_2) \equiv \Upsilon(\varepsilon)$. In particular, the set of $t$-period histories in $\Upsilon(\varepsilon)$ is again given by $(\{H, L\} \times \{h, \ell\})^t$, so that past

|  | $h$ | $\ell$ |
|---|---|---|
| $H$ | $2, 3$ | $0, 2 + \varepsilon_2 z_t^2$ |
| $L$ | $3 + \varepsilon_1 z_t^1, 0$ | $1 + \varepsilon_1 z_t^1, 1 + \varepsilon_2 z_t^2$ |

Figure 2: The $(\varepsilon_1, \varepsilon_2)$-perturbed product-choice game.

actions are perfectly monitored, while payoffs shocks remain private. By construction, in any period, if $\varepsilon_2 > 0$, for almost all realizations of the payoff shock, the short-lived player cannot be indifferent between $h$ and $\ell$ (with a similar comment applying to the long-lived player).

An equilibrium of $\Upsilon(\varepsilon)$ is *$\gamma$-close* to the belief-free equilibrium indexed by $p^H$ if for all $i$ and all histories ending in $a_1 \in \{H, L\}$, the ex ante probability (i.e., taking expectations over the current payoff shock) of the long-lived player playing $H$ is within $\gamma$ of $1/2$, and the ex ante probability of the short-lived player playing $h$ is within $\gamma$ of $p^{a_1}$. A belief-free equilibrium is *purified* if for all $\gamma > 0$ there exist $\varepsilon_1 > 0$ and $\varepsilon_2 > 0$ such that there is a $\gamma$-close equilibrium of $\Upsilon(\varepsilon_1, \varepsilon_2)$.

#### 2.2.1. Perturbing only the short-lived player, $\varepsilon_1 = 0$.

When $\varepsilon_2 > 0$, the short-lived player's best reply to any long-lived player behavior is, for almost all realizations of $z_2^t$, unique. Moreover, the best reply is history independent when the long-lived player's behavior is also history independent. Since the provision of intertemporal incentives to the long-lived player in a belief-free equilibrium requires history dependent play by short-lived players, any equilibrium of $\Upsilon(0, \varepsilon_2)$ close to that belief-free equilibrium must feature history dependent play by the long-lived player. Nonetheless, constructing an equilibrium of $\Upsilon(0, \varepsilon_2)$ close to any belief-free equilibrium is straightforward: Fix a belief-free equilibrium indexed by $p^H$ (with $p^L$ given by (5)). Since the long-lived player's payoffs are not perturbed, his incentives are preserved by setting the ex ante probabilities of $h$ equal to those in the belief-free equilibrium, while the behavior of the long-lived player is chosen to "rationalize" the behavior of the short-lived players. Specifically, let $a_1 \in \{H, L\}$ denote the long-lived player's action in the previous period. The ex ante probability the short-lived player plays $h$ is given by $\Pr(h) = p^{a_1}$. Since

$$\Pr(h) = \Pr(z_t^2 \leq \hat{z}_{a_1}^2) = \hat{z}_{a_1}^2,$$

where $\hat{z}_t^2$ is the short-lived player type indifferent between $h$ and $\ell$, we set $\hat{z}_{a_1}^2 = p^{a_1}$. The long-lived player plays $H$ with probability $\pi_1^{a_1}$ to make $\hat{z}_t^2$ indifferent between $h$ and $\ell$. Solving

$$\pi_1^{a_1} \times 3 = \pi_1^{a_1} \times 2 + (1 - \pi_1^{a_1}) \times 1 + \varepsilon_2 \hat{z}_{a_1}^2$$

gives

$$\pi_1^{a_1} = \frac{1 + \varepsilon_2 \hat{z}_{a_1}^2}{2}. \tag{6}$$

Note that, as $\varepsilon_2 \to 0$, the ex ante probability of $H$ after $a_1$ in the previous period, $\pi_1^{a_1}$, converges to 1/2, as required.

### 2.2.2. Perturbing only the long-lived player, $\varepsilon_2 = 0$.

Turning to the other extreme, we now suppose $\varepsilon_1 > 0$ and $\varepsilon_2 = 0$. Since the short-lived players' payoffs are not perturbed, short-lived players are only willing to play history dependent strategies if the ex ante probability the long-lived player plays $H$ is 1/2, implying the marginal type is 1/2.

Fix a belief-free equilibrium indexed by $p^H$ and let $a_1 \in \{H, L\}$ denote the long-lived player's action in the previous period. Let $W^\varepsilon(a_1)$ be the long-lived player's ex ante expected discounted value from following the strategy of playing $H$ if $z_t^1 \leq 1/2$ and $L$ if $z_t^1 > 1/2$ from this period (since $z_t^1 = 1/2$ with zero probability, the specification for the marginal type is irrelevant),

$$W^\varepsilon(a_1) = \int_0^{1/2} \{(1-\delta)2\pi_2^{a_1} + \delta W^\varepsilon(H)\}\, dz_t^1 + \int_{1/2}^1 \left\{(1-\delta)(2\pi_2^{a_1} + 1 + \varepsilon_1 z_t^1) + \delta W^\varepsilon(L)\right\} dz_t^1,$$

where $\pi_2^{a_1}$ is the probability the short-lived player puts on $h$ (after observing $a_1$ from the previous period). Simplifying.

$$W^\varepsilon(a_1) = (1-\delta)\left\{2\pi_2^{a_1} + \frac{1}{2} + \frac{3\varepsilon_1}{8}\right\} + \frac{\delta}{2}\{W^\varepsilon(H) + W^\varepsilon(L)\},$$

implying

$$W^\varepsilon(H) - W^\varepsilon(L) = 2(1-\delta)(\pi_2^H - \pi_2^L). \tag{7}$$

The expected discounted value to the long-lived player from playing $H$ is

$$(1-\delta)2\pi_2^{a_1} + \delta W^\varepsilon(H),$$

while from $L$, given the payoff realization $z_t^1$, is

$$(1-\delta)(2\pi_2^{a_1} + 1 + \varepsilon_1 z_t^1) + \delta W^\varepsilon(L).$$

Since the type $z_t^1 = 1/2$ must be indifferent between $H$ and $L$,

$$(1-\delta)2\pi_2^{a_1} + \delta W^\varepsilon(H) = (1-\delta)(2\pi_2^{a_1} + 1 + \varepsilon_1/2) + \delta W^\varepsilon(L),$$

that is

$$\delta(W^\varepsilon(H) - W^\varepsilon(L)) = (1-\delta)(1+\varepsilon_1/2). \tag{8}$$

This implies (from (7))

$$\pi_2^L = \pi_2^H - \frac{1+\varepsilon_1/2}{2\delta},$$

which for small $\varepsilon_1$ is close to (5).

#### 2.2.3. Perturbing both players, $\varepsilon_1 > 0$ and $\varepsilon_2 > 0$.

Key to the construction of the purifying equilibrium of $\Upsilon(0, \varepsilon_2)$ is the lack of long-lived player perturbations, so that ex ante probabilities satisfying (5) made the long-lived player indifferent between $H$ and $L$ after any history. Key to the construction of the purifying equilibrium of $\Upsilon(\varepsilon_1, 0)$ is that the unperturbed short-lived player must face the *history-independent* randomization $\frac{1}{2} \circ H + \frac{1}{2} \circ L$ in order to be willing to randomize (in a history dependent manner).

The difficulty with purifying belief-free equilibria arises from the interaction between the problem of providing incentives for two players. Fix $\varepsilon_1 > 0$ and $\varepsilon_2 > 0$, and consider a candidate equilibrium in which in period $t$, the long-lived player puts probability $\pi_1^{a_1}$ on $H$ and the short-lived player puts probability $\pi_2^{a_1}$ on $h$, where $a_1 \in \{H, L\}$ is the long-lived player's action in period $t-1$. As before, $\hat{z}^i_{a_1} = \pi_i^{a_1}$ is the marginal type for each player. The reasoning in Subsection 2.2.1 still applies, and so in order for $\hat{z}^2_{a_1}$ to be indifferent we need, from (6),

$$\pi_1^{a_1} = \frac{1+\varepsilon_2 \hat{z}^2_{a_1}}{2} = \frac{1+\varepsilon_2 \pi_2^{a_1}}{2}.$$

Since $\pi_2^H \neq \pi_2^L$, this implies $\pi_1^H \neq \pi_1^L$. On the other hand, the reasoning in Subsection 2.2.2 still applies, and so in order for $\hat{z}^1_{a_1}$ to be indifferent we need, from (8),

$$\delta(W^\varepsilon(H) - W^\varepsilon(L)) = (1-\delta)(1+\varepsilon_1 \hat{z}^1_{a_1}) = (1-\delta)(1+\varepsilon_1 \pi_1^{a_1}),$$

which cannot be satisfied for distinct values $\pi_1^H \neq \pi_1^L$. Hence, belief-free equilibria cannot be purified by strategies that depend only on the long-lived player's last period action. An analysis similar to that in Section 4 shows that allowing each player's behavior to depend on *both* players' last period action does not change this conclusion.

## 3. The Infinitely Repeated Prisoners' Dilemma

Let $\Gamma(0)$ denote the infinitely-repeated perfect-monitoring prisoners' dilemma with stage game displayed in Figure 3. Each player has a discount rate $\delta$. The class of mixed strategy equilibria Ely and Välimäki (2002) construct can be described as follows: The

|  | $C$ | $D$ |
|---|---|---|
| $C$ | $1, 1$ | $-\ell, 1+g$ |
| $D$ | $1+g, -\ell$ | $0, 0$ |

Figure 3: The unperturbed prisoners' dilemma stage game.

profiles have one-period memory, with players randomizing in each period with player $i$ assigning probability $p^i_{aa'}$ to $C$ after the action profile $a_i a_j = aa'$, where $j = 3 - i$. As in the previous section, the profile is constructed so that after *each* action profile, the player is indifferent between $C$ and $D$. Consequently, a player's best replies are independent of his beliefs about the opponent's history. The requirement that after $a_i a_j = aa'$, player $i$ is indifferent between playing $C$ and $D$ when player $j$ is playing $p^j_{a'a}$ yields the following system (where $W^i_{aa'}$ is the value to player $i$ after $aa'$, and the second equality in each displayed equation comes from the indifference requirement):

$$\begin{aligned} W^i_{aa'} &= (1-\delta)(p^j_{a'a} + (1-p^j_{a'a})(-\ell)) \\ &\qquad + \delta\left\{p^j_{a'a} W^i_{CC} + (1-p^j_{a'a}) W^i_{CD}\right\} \qquad (9) \\ &= (1-\delta)p^j_{a'a}(1+g) + \delta\left\{p^j_{a'a} W^i_{DC} + (1-p_{a'a}) W^i_{DD}\right\}. \qquad (10) \end{aligned}$$

Subtracting (10) from (9) gives

$$\begin{aligned} 0 = p^j_{a'a}\left\{(1-\delta)(-g+\ell) + \delta\left[(W^i_{CC} - W^i_{DC}) - (W^i_{CD} - W^i_{DD})\right]\right\} \\ - (1-\delta)\ell + \delta(W^i_{CD} - W^i_{DD}). \end{aligned}$$

Since at least two of the probabilities differ (if not, $p^j_{aa'} = 0$ for all $aa'$), the coefficient of $p^j_{aa'}$ and the constant term are both zero:

$$W^i_{CD} - W^i_{DD} = \frac{(1-\delta)\ell}{\delta} \qquad (11)$$

and

$$\begin{aligned} W^i_{CC} - W^i_{DC} &= \frac{(1-\delta)(g-\ell)}{\delta} + W^i_{CD} - W^i_{DD} \\ &= \frac{(1-\delta)g}{\delta}. \qquad (12) \end{aligned}$$

These two equations succinctly capture the tradeoffs facing potentially randomizing players. Suppose a player knew his partner was going to play $D$ this period. The myopic incentive to also play $D$ is $\ell$, while the cost of doing so is that his continuation value falls from $W^i_{CD}$ to $W^i_{DD}$. Equation (11) says that these two should exactly balance. Suppose instead the player knew his partner was going to play $C$ this period. The myopic incentive to playing $D$ is now $g$, while the cost of playing $D$ is now that his continuation value falls from $W^i_{CC}$ to $W^i_{DC}$. This time it is equation (12) that says that these two should exactly balance. Notice that these two equations imply that a player's best replies are independent of the current realized behavior of the opponent.[8]

A profile described by the four probabilities $(p^i_{aa'} : aa' \in \{C,D\}^2)$ for each player $i \in \{1,2\}$ is an equilibrium when (9) and (10) are satisfied for the four action profiles $aa' \in \{C,D\}^2$, and for $i = 1,2$. Since the value functions are determined by the probabilities, the four probabilities are free parameters, subject only to (11) and (12). This redundancy implies a two-dimensional indeterminacy in the solutions for each of the players, and it is convenient to parameterize the solutions by $W^i_{CC}$ and $W^i_{CD}$.

Solving (9) for $aa' = CC$ gives

$$p^j_{CC} = \frac{(1-\delta)\ell + W^i_{CC} - \delta W^i_{CD}}{(1-\delta)(1+\ell) + \delta(W^i_{CC} - W^i_{CD})}, \tag{13}$$

for $aa' = CD$ gives

$$p^j_{DC} = \frac{(1-\delta)\ell + W^i_{CD} - \delta W^i_{CD}}{(1-\delta)(1+\ell) + \delta(W^i_{CC} - W^i_{CD})}, \tag{14}$$

for $aa' = DC$ (using (12)) gives

$$p^j_{CD} = \frac{(1-\delta)(\ell - g/\delta) + W^i_{CC} - \delta W^i_{CD}}{(1-\delta)(1+\ell) + \delta(W^i_{CC} - W^i_{CD})}, \tag{15}$$

and, finally, for $aa' = DD$ (using (11))gives

$$p^j_{DD} = \frac{(1-\delta)\ell(1-1/\delta) + W^i_{CD} - \delta W^i_{CD}}{(1-\delta)(1+\ell) + \delta(W^i_{CC} - W^i_{CD})}. \tag{16}$$

As in the introductory example (see footnote 6), if the myopic incentive to play the dominant action is independent of the opponent's play, i.e., $g = \ell$, there is a particularly simple belief-free profile. By setting $W^i_{CC} = W^i_{CD}$, we obtain a profile where $j$'s randomization is independent of $j$'s previous action, $p^j_{Ca} = p^j_{Da}$ for all $a \in \{C,D\}$. In such an equilibrium, after any two histories ending in $CC$ and in $CD$, player 1 randomizes differently even though player 2's continuation play is the same.

We have described an equilibrium if the expressions in (13)-(16) are probabilities.

[8] This is the starting point of Ely and Välimäki (2002), who work directly with the values to a player of having his opponent play $C$ and $D$ *this* period.

**Theorem 1** *There is a four-dimensional manifold of mixed equilibria of the infinitely-repeated perfect monitoring prisoners' dilemma: Suppose* $0 < W^i_{CD} \leq W^i_{CC} \leq 1$ *satisfy the inequalities*

$$(1-\delta)g/\delta + \delta W^i_{CD} \leq (1-\delta)\ell + W^i_{CC} \tag{17}$$

$$\text{and} \qquad (1-\delta)\ell \leq \delta W^i_{CD}. \tag{18}$$

*The profile in which player* $i$ *plays* $C$ *with probability* $p^i_{aa'}$ *after* $a_i a_j = aa'$ *in the previous period (and plays* $p^i_{\tilde{a}\tilde{a}'}$ *in the first period, for any* $\tilde{a}, \tilde{a}' \in \{C, D\}$*), where* $p^i_{aa'}$ *for all* $a, a' \in \{C, D\}^2$ *are given by (13)-(16), is an equilibrium. Moreover, (17) and (18) are satisfied for any* $0 < W^i_{CD} < W^i_{CC} \leq 1$*, for* $\delta$ *sufficiently close to* $1$*.*

**Proof.** We need only verify that (17) and (18) imply that the quantities described by (13)-(16) are probabilities. It is immediate that $p^j_{CD} < p^j_{CC}$ and $p^j_{DD} < p^j_{DC} \leq p^j_{CC}$, so the only inequalities we need to verify are $0 \leq p^j_{CD}, p^j_{DD}$ and $p^j_{CC} \leq 1$. Observe first that the common denominator in (13)-(16) is strictly positive from $W^i_{CD} \leq W^i_{CC}$.

Now, $p^j_{CC} \leq 1$, since $W^i_{CC} \leq 1$.

We also have $p^j_{CD} \geq 0$, since

$$\begin{aligned}
&(1-\delta)(\ell - g/\delta) + W^i_{CC} - \delta W^i_{CD} \geq 0 \\
&\qquad\qquad \iff (1-\delta)\ell + W^i_{CC} \geq (1-\delta)g/\delta + \delta W^i_{CD},
\end{aligned}$$

which is (17).

Finally, $p^j_{DD} \geq 0$ is equivalent to (18). ∎

For each specification of behavior in the first period, there is a four-dimensional manifold of equilibria. Our analysis applies to all of these manifolds, and for simplicity, we focus on the profiles where players play $p^i_{CC}$ in the first period.

## 4. One Period Memory Purification

We now argue that it is impossible to purify equilibria of the type described in Section 3 for generic distributions of the payoff shocks using equilibria of the perturbed game with one period history dependence.

Let $\Gamma(\varepsilon)$ denote the infinitely-repeated perfect-monitoring prisoners' dilemma with stage game displayed in Figure 4. The payoff shock $z^i_t$ is private to player $i$, realized in period $t$, independently and identically distributed across players, and histories, according to the distribution function $F(.)$.[9] The distribution function has support $[0, 1]$, and

[9] Harsanyi (1973) allowed the payoff perturbations to depend on both players' actions. For tractability, we have restricted attention to the case where the perturbations do not depend on the opponent's actions. We believe our results carry over to the general case, but have not verified this.

| | $C$ | $D$ |
|---|---|---|
| $C$ | $1+\varepsilon z_t^1, 1+\varepsilon z_t^2$ | $-\ell+\varepsilon z_t^1, 1+g$ |
| $D$ | $1+g, -\ell+\varepsilon z_t^2$ | $0, 0$ |

Figure 4: The perturbed prisoners' dilemma stage game.

a density bounded away from zero. Let $\mathcal{F}$ be the collection such distribution functions endowed with the weak topology. A property is *generic* if the set of distribution functions for which it holds is open and dense in $\mathcal{F}$. We explain the critical implication of genericity after the proof of Theorem 2, when we explain why the uniform distribution must be ruled out.

An equilibrium of $\Gamma(\varepsilon)$ is $\gamma$-close to $p$ (an equilibrium of the form described in theorem 1), if for all $i$ and all $a, a' \in \{C, D\}$, for all histories ending in $aa'$, the ex ante probability (i.e., taking expectations over the current payoff shock) of player $i$ playing $C$ is within $\gamma$ of $p^i_{aa'}$. An equilibrium $p$ is *purified* for the distribution $F$ if, for all $\gamma > 0$ there exists $\varepsilon > 0$ such that there is an equilibrium of $\Gamma(\varepsilon)$ is $\gamma$-close to $p$.

**Theorem 2** *Let $p$ be a completely mixed strategy equilibrium of the form described in Theorem 1. Generically in the space of payoff shock distributions, there exists $\gamma > 0$ such that for all $\varepsilon > 0$, there is no equilibrium of $\Gamma(\varepsilon)$ with one period memory within $\gamma$ distance of $p$.*

**Proof.** Fix $\gamma' = \min\{p^i_{aa'}\}/2$. Any profile within $\gamma'$ of $p$ is completely mixed. We show by contradiction that there are no one period memory equilibria in $\Gamma(\varepsilon)$ that are $\gamma$ close to $p$, for $\gamma$ sufficiently small.

Fix a one period memory equilibrium of $\Gamma(\varepsilon)$ that is $\gamma$-close to the equilibrium $p$ of $\Gamma(0)$, where $\gamma < \gamma'$. Denote the ex ante probability of player $i$ playing $C$ after observing the action profile $aa'$ by $\pi^{i\varepsilon}_{aa'}$. Player $i$ will play $C$ in period $t$ if and only if the payoff shock $z^i_t$ is sufficiently large. Then the probability of $C$ is $\pi^{i\varepsilon}_{aa'} = \Pr\left\{z^i_t \geq \hat{z}^i_{aa'}\right\} = 1 - F(\hat{z}^i_{aa'})$ for some marginal type $\hat{z}^i_{aa'}$. If $z^t_i \geq \hat{z}^i_{aa'}$ then $i$ plays $C$, and plays $D$ otherwise. Since $\pi^{i\varepsilon}_{aa'} \in (0,1)$, we have $\hat{z}^i_{aa'} \in (0,1)$ for every action profile $aa'$ and for $i \in \{1,2\}$.

The marginal type $\hat{z}^{i\varepsilon}_{aa'}$ is indifferent between $C$ and $D$ when the action profile played in the last period is $aa'$. Let $W^{i\varepsilon}_{aa'}$ denote the ex ante value function of a player at the action profile $aa'$, before the realization of his payoff shock. The interim payoff from $C$ after $aa'$ and given the payoff realization $z^i_t$ is

$$V^{i\varepsilon}_{aa'}(z^i_t, C) = (1-\delta)\left\{\pi^{j\varepsilon}_{a'a} - (1-\pi^{j\varepsilon}_{a'a})\ell + \varepsilon z^i_t\right\} + \delta\left\{\pi^{j\varepsilon}_{a'a} W^{i\varepsilon}_{CC} + (1-\pi^{j\varepsilon}_{a'a}) W^{i\varepsilon}_{CD}\right\},$$

while the payoff to $i$ from $D$ after the profile $aa'$ is

$$V^{i\varepsilon}_{aa'}(z^i_t, D) = (1-\delta)\left\{\pi^{j\varepsilon}_{a'a}(1+g)\right\} + \delta\left\{\pi^{j\varepsilon}_{a'a}W^{i\varepsilon}_{DC} + (1-\pi^{j\varepsilon}_{a'a})W^{i\varepsilon}_{DD}\right\}. \tag{19}$$

Since $\hat{z}^i_{aa'}$ is indifferent,

$$\begin{aligned}(1-\delta)\left\{\pi^{j\varepsilon}_{a'a} - (1-\pi^{j\varepsilon}_{a'a})\ell + \varepsilon\hat{z}^i_{aa'}\right\} + \delta\left\{\pi^{j\varepsilon}_{a'a}W^{i\varepsilon}_{CC} + (1-\pi^{j\varepsilon}_{a'a})W^{i\varepsilon}_{CD}\right\}& \\ = (1-\delta)\pi^{j\varepsilon}_{a'a}(1+g) + \delta\left\{\pi^{j\varepsilon}_{a'a}W^{i\varepsilon}_{DC} + (1-\pi^{j\varepsilon}_{a'a})W^{i\varepsilon}_{DD}\right\}&.\end{aligned}$$

Using $\hat{z}^i_{aa'} = F^{-1}(1-\pi^{i\varepsilon}_{aa'})$,

$$\begin{aligned}F^{-1}(1-\pi^{i\varepsilon}_{aa'}) = &\frac{1}{(1-\delta)\varepsilon}\left\{(1-\delta)\ell + \delta(W^{i\varepsilon}_{DD} - W^{i\varepsilon}_{CD})\right\} \\ &+ \frac{\pi^{j\varepsilon}_{a'a}}{(1-\delta)\varepsilon}\left\{(1-\delta)(g-\ell) + \delta(W^{i\varepsilon}_{DC} + W^{i\varepsilon}_{CD} - W^{i\varepsilon}_{CC} - W^{i\varepsilon}_{DD})\right\}.\end{aligned} \tag{20}$$

Note that the right hand side of (20) is affine in $\pi^{j\varepsilon}_{a'a}$, player $j$'s mixing probability. Let $\alpha^{i\varepsilon}$ and $\beta^{i\varepsilon}$ denote the intercept and slope of this linear function; these do not depend upon the profile $aa'$. We may therefore re-write (20) as

$$F^{-1}(1-\pi^{i\varepsilon}_{aa'}) = \alpha^{i\varepsilon} + \beta^{i\varepsilon}\pi^{j\varepsilon}_{a'a}. \tag{21}$$

Equation (21) must hold for all $a, a' \in \{C, D\}$. In other words, the points in the set

$$\mathcal{Z}^{\varepsilon}_i \equiv \left\{\left(\pi^{j\varepsilon}_{a'a}, F^{-1}(1-\pi^{i\varepsilon}_{aa'})\right) : a, a' \in \{C, D\}\right\}$$

must be collinear, for $i \in \{1, 2\}$.

If the points in the set

$$\mathcal{Z}^0_i \equiv \left\{\left(p^j_{a'a}, F^{-1}(1-p^i_{aa'})\right) : a, a' \in \{C, D\}\right\}$$

are not collinear, then for $\gamma$ sufficiently small,[10] if $|\pi^{j\varepsilon}_{a'a} - p^j_{a'a}| < \gamma$ for all $j \in \{1, 2\}$ and $a, a' \in \{C, D\}$, the points in $\mathcal{Z}^{\varepsilon}_i$ will also not be collinear. But this would contradict (21) and so the existence of the putative equilibrium.

Consider first the case where, for some player $i$, $p$ specifies three distinct mixing probabilities. In that case, it is clear that for generic $F$, the points in the set $\mathcal{Z}^0_i$ are not collinear and we have the contradiction.

[10] Note that the bound on $\gamma$, while depending on $F$, is independent of $\varepsilon$.

Consider now the case when $p$ has only two distinct values of $p^i_{aa'}$ for all $i$. From (14) and (16), $p^i_{DC} > p^i_{DD}$, while from (13) and (16), we deduce $p^i_{CC} > p^i_{DD}$. Thus the only possibility for only two distinct values is $p^i_{CC} = p^i_{DC} \equiv p^i_C$ and $p^i_{CD} = p^i_{DD} \equiv p^i_D < p^i_C$ for all $i$. But this implies

$$\begin{aligned}\mathcal{Z}^0_i =& \big\{ \big(p^j_C, F^{-1}(1-p^i_C)\big), \big(p^j_C, F^{-1}(1-p^i_D)\big), \big(p^j_D, F^{-1}(1-p^i_C)\big), \big(p^j_D, F^{-1}(1-p^i_D)\big) \big\} \\ =& \big\{p^j_C, p^j_D\big\} \times \big\{F^{-1}(1-p^i_C), F^{-1}(1-p^i_D)\big\}.\end{aligned}$$

The points in $\mathcal{Z}^0_i$ clearly cannot be collinear, and we again have a contradiction. ■

The theorem asserts that for any fixed mixed strategy equilibrium $p$, there does not exist a one period purification for generic shock distributions. While we assume that each player receives payoff shocks from the same distribution, the same argument goes through with asymmetric payoff distributions. The theorem does not rule out the possibility that, for generic shock distributions, there will be some mixed strategy equilibrium $p$ (depending on the shock distribution) that is purified with one-period memory strategies.

In an earlier version of this work, Bhaskar, Mailath, and Morris (2004), as well as in the introductory example of Section 2, we studied the (non-generic) case of uniform noise. Uniform noise is special because $F^{-1}$ is linear. In this case, some symmetric strategies were purifiable but others were not. Note first that if the noise is uniform, then the set of points $\{(p, F^{-1}(1-p)) : p \in [0,1]\}$ is trivially collinear, since $(p, F^{-1}(1-p)) = (p,\ (1-p))$. Thus, for any profile satisfying $p^j_{aa'} = p^i_{a'a} \equiv p_{a'a}$ and $p_{CD} = p_{DC}$, i.e., a *strongly symmetric* profile, the set $\mathcal{Z}^0_i$ is collinear, and the profile is purifiable using one period memory strategies.

More generally, the critical property the distribution function must satisfy for the argument precluding purification to go through is the following: For a *given* set of three points $\{(p^\kappa_1, p^\kappa_2) \in [0,1]^2 : \kappa = 1,2,3\}$, the set

$$\big\{(p^\kappa_1,\ F^{-1}(1-p^\kappa_2)) : \kappa = 1,2,3\big\}$$

should not be collinear.

**Remark 2 (Stronger Impossibility Results)** In Bhaskar, Mailath, and Morris (2004), we had asserted that the type of argument reported here would extend to finite memory strategy profiles of any length. Unfortunately, the argument we gave was invalid, and while the assertion might be true, we have been unsuccessful in obtaining a proof.

Stronger impossibility results for the purifiability of belief free strategies can be obtained if the stage game is one of perfect information. Bhaskar (1998) analyzes Samuelson's overlapping generations transfer game and shows that finite memory implies that no transfers can be sustained in any purifiable equilibrium. We conjecture that this result extends to any repeated game, where the stage game is one of perfect information and players are restricted to finite memory strategies. In any purifiable equilibrium, the backwards induction outcome of the stage game must be played in every period. Simultaneous moves, as in the present paper, allow for greater possibilities of purification: some belief free strategies are purifiable via one period memory strategies for non-generic payoff shock distributions. More importantly, the induction argument extending the negative one period result to arbitrary finite memory strategies is not valid in the simultaneous move case.

♦

## 5. Purification with Infinite Memory

We now argue that, when we allow the equilibrium of the perturbed game to have infinite history dependence, then it is possible to purify belief-free equilibria of the type described in Section 3. To simplify notation, we focus on *symmetric* equilibria, so that $p_{aa'}$ is the probability player 1 plays $C$ after the profile $aa'$ (with player 2 playing $C$ with probability $p_{a'a}$). Fix an equilibrium with interior probabilities, $p_{CC}$, $p_{CD}$, $p_{DC}$, and $p_{DD} \in (0,1)$.

We first partition the set of histories, $\mathcal{H}$, into equivalence classes, denoted by $(aa', k)$, $aa' \in \{C, D\}^2$ and $k \geq 0$. The equivalence class $(aa', 0)$ with $aa' \neq CC$ consists of all histories whose last action profile is $aa'$. The equivalence class $(aa', k)$ with $aa' \neq CC$ and $k \geq 1$ consists of all histories whose last action profile is $CC$ and there were $k$ occurrences of $CC$ after the last non-$CC$ action profile $aa'$. Finally, the equivalence class $(CC, k)$ is a singleton, containing the $k$-period history in which $CC$ has been played in every period. Note that the null history is $(CC, 0)$, and that any history is an element of the partition $(aa', k)$, where the history ends in $CC$ if $k \geq 1$.

The purifying strategy in the perturbed game is measurable with respect to the partition on $\mathcal{H}$ just described. Fix $\varepsilon > 0$ and let $\pi^{\varepsilon}_{aa'}(k)$ denote the probability with which $C$ is played when $h \in (aa', k)$, and let $W^{\varepsilon}_{aa'}(k)$ denote the ex ante value function of the player at this history. If $\left\{\pi^{\varepsilon}_{aa'}(k)\right\}$ is a sequence (as $\varepsilon \to 0$) of equilibria purifying $p = (p_{CC}, p_{CD}, p_{DC}, p_{DD})$, then $\pi^{\varepsilon}_{aa'}(k) \to p_{CC}$ for all $k \geq 1$ and all $aa'$, and $p^{\varepsilon}_{aa'}(0) \to p_{aa'}$, as $\varepsilon \to 0$. We show a *uniform* form of purifiability: the bound on $\varepsilon$ required to make $\pi^{\varepsilon}_{aa'}(k)$ close to $p_{CC}$ is *independent* of $k$.

The idea is that in the perturbed game, the payoff after a history ending in $CC$ can always be adjusted to ensure that the appropriate realization of $z$ in the previous period is the marginal type to obtain the desired randomization between $C$ and $D$. We proceed recursively, fixing probabilities after any history in an element of the partition $(aa', 0)$ at their unperturbed levels, i.e., we set $\pi^{\varepsilon}_{aa'}(0) = p_{aa'}$. In particular, players randomize in the first period with probability $p_{CC}$ on $C$, and in the second period after a realized action profile $aa' \neq CC$ with probability $p_{aa'}$ on $C$.[11] This turns out to determine the value function at histories in $(aa', 0)$ for all $aa'$. In the second period after $CC$, $W^{i\varepsilon}_{CC}(1)$ is determined by the requirement that the ex ante probability that a player play $C$ in the first period is given by $\pi^{\varepsilon}_{CC}(0) = p_{CC}$. Given the value $W^{\varepsilon}_{CC}(1)$, the probability $\pi^{\varepsilon}_{CC}(1)$ is then determined by the requirement that $W^{\varepsilon}_{CC}(1)$ be the ex ante value at the history $CC$. More generally, given a history $h \in (aa', k)$ and a further realization of $CC$, $W^{\varepsilon}_{aa'}(k+1)$ is determined by the requirement that the ex ante probability that a player play $C$ in the previous period is given by $\pi^{\varepsilon}_{aa'}(k) = p_{aa'}$, and then $\pi^{\varepsilon}_{aa'}(k+1)$ is then determined by $W^{\varepsilon}_{aa'}(k+1)$.

Denote by $V^{\varepsilon}_{aa'}(k; z_t, \hat{a})$ the interim payoff from $\hat{a}$, given the payoff realization $z_t$, after a history falling into the equivalence class $(aa', k)$ (and assuming continuation play follows the purifying strategy). These payoffs are given by (where $W^{\varepsilon}_{aa'} \equiv W^{\varepsilon}_{aa'}(0)$)

$$\begin{aligned} V^{\varepsilon}_{aa'}(k; z_t, C) = (1-\delta)\left\{\pi^{\varepsilon}_{a'a}(k) + (1-\pi^{\varepsilon}_{a'a}(k))(-\ell) + \varepsilon z_t\right\} \\ + \delta\left\{\pi^{\varepsilon}_{a'a}(k) W^{\varepsilon}_{aa'}(k+1) + (1-\pi^{\varepsilon}_{a'a}(k)) W^{\varepsilon}_{CD}\right\} \end{aligned} \quad (22)$$

and

$$V^{\varepsilon}_{aa'}(k; z_t, D) = (1-\delta)\pi^{\varepsilon}_{a'a}(k)(1+g) + \delta\left\{\pi^{\varepsilon}_{a'a}(k) W^{\varepsilon}_{DC} + (1-\pi^{\varepsilon}_{a'a}(k)) W^{\varepsilon}_{DD}\right\}.$$

Since $\pi^{\varepsilon}_{aa'}(k)$ is the probability of $C$ at a history in $(aa', k)$, the player with any payoff realization $z_t \geq \hat{z} \equiv F^{-1}[1 - \pi^{\varepsilon}_{aa'}(k)]$ chooses $C$, and $D$ otherwise. Moreover, the player with payoff realization $\hat{z}$ must be indifferent between $C$ and $D$, i.e.,

$$V^{\varepsilon}_{aa'}(k; \hat{z}, C) = V^{\varepsilon}_{aa'}(k; D). \quad (23)$$

From (22),

$$V^{\varepsilon}_{aa'}(k; z_t, C) = V^{\varepsilon}_{aa'}(k; \hat{z}, C) + (1-\delta)\varepsilon(z_t - \hat{z}),$$

and since $V^{\varepsilon}_{aa'}(k; z_t, D) \equiv V^{\varepsilon}_{aa'}(k; D)$ is independent of $z_t$, we have (using (23))

$$\begin{aligned} W^{\varepsilon}_{aa'}(k) =& V^{\varepsilon}_{aa'}(k; D) + (1-\delta)\varepsilon \int_{\hat{z}}^{1} (z - \hat{z})\, dF(z) \\ =& V^{\varepsilon}_{aa'}(k; D) + (1-\delta)\varepsilon G(\pi^{\varepsilon}_{aa'}(k)), \end{aligned}$$

where $G(\pi)$ is the ex ante expected incremental value of the payoff shock to this player from playing $C$ with probability $\pi$,

$$G(\pi) = \int_{F^{-1}(1-\pi)}^{1} x - F^{-1}(1-\pi)\, dF(x).$$

[11] More precisely, player 1 randomizes with probability $p_{aa'}$ and player 2 randomizes with probability $p_{a'a}$.

Beginning with histories in $(aa',0)$, we have

$$W^\varepsilon_{CD} = (1-\delta)\left\{p_{DC}(1+g)+\varepsilon G(p_{CD})\right\} + \delta\left\{p_{DC}W^\varepsilon_{DC} + (1-p_{DC})W^\varepsilon_{DD}\right\}, \quad (24)$$

$$W^\varepsilon_{DC} = (1-\delta)\left\{p_{CD}(1+g)+\varepsilon G(p_{DC})\right\} + \delta\left\{p_{CD}W^\varepsilon_{DC} + (1-p_{CD})W^\varepsilon_{DD}\right\}, \quad (25)$$

$$W^\varepsilon_{DD} = (1-\delta)\left\{p_{DD}(1+g)+\varepsilon G(p_{DD})\right\} + \delta\left\{p_{DD}W^\varepsilon_{DC} + (1-p_{DD})W^\varepsilon_{DD}\right\}, \quad (26)$$

and

$$W^\varepsilon_{aa'}(k) = (1-\delta)\left\{\pi^\varepsilon_{a'a}(k)(1+g)+\varepsilon G(\pi^\varepsilon_{aa'}(k))\right\} + \delta\left\{\pi^\varepsilon_{a'a}(k)W^\varepsilon_{DC} + (1-\pi^\varepsilon_{a'a}(k))W^\varepsilon_{DD}\right\}. \quad (27)$$

As we indicated above, (24), (25), and (26) can be solved for $W^\varepsilon_{CD}$, $W^\varepsilon_{DC}$, and $W^\varepsilon_{DD}$. Moreover, these solutions converge to $W_{CD}$, $W_{DC}$, and $W_{DD}$. It remains to determine $W^\varepsilon_{aa'}(k)$ and $\pi^\varepsilon_{aa'}(k)$ for $k \geq 1$ ($W^\varepsilon_{CC}(0)$ is also determined, since $\pi^\varepsilon_{CC}(0) = p_{CC}$).

Solving (23) at a history in $(aa', k-1)$ for $W^\varepsilon_{a'a}(k)$ as a function of $\pi^\varepsilon_{aa'}(k-1)$ and $\pi^\varepsilon_{a'a}(k-1)$ gives

$$W^\varepsilon_{a'a}(k) = \frac{(1-\delta)(g-\ell)}{\delta} + W^\varepsilon_{DC} + W^\varepsilon_{CD} - W^\varepsilon_{DD} + \frac{(1-\delta)\{\ell - \varepsilon F^{-1}[1-\pi^\varepsilon_{a'a}(k-1)]\} - \delta[W^\varepsilon_{CD} - W^\varepsilon_{DD}]}{\delta\pi^\varepsilon_{aa'}(k-1)}. \quad (28)$$

This can be re-written (using (11)) as

$$W^\varepsilon_{a'a}(k) = \frac{(1-\delta)(g-\ell)}{\delta} + W^\varepsilon_{DC} + W^\varepsilon_{CD} - W^\varepsilon_{DD} + \frac{\delta[(W_{CD} - W^\varepsilon_{CD}) - (W_{DD} - W^\varepsilon_{DD})] - (1-\delta)\varepsilon F^{-1}[1-\pi^\varepsilon_{a'a}(k-1)]}{\delta\pi^\varepsilon_{aa'}(k-1)}. \quad (29)$$

Examining (29), we see that the terms in the first line converge to $W_{CC}$ as $\varepsilon \to 0$. Since the numerator of the second line vanishes as $\varepsilon \to 0$, this implies that $W^\varepsilon_{a'a}(k) \to W_{CC}$ provided that $\pi^\varepsilon_{aa'}(k-1)$ is bounded away from zero.

Given a value for $W^\varepsilon_{aa'}(k)$,[12] (27) can be re-written as

$$(1-\delta)\varepsilon G(\pi^\varepsilon_{aa'}(k)) + b_\varepsilon \pi^\varepsilon_{aa'}(k) + c_\varepsilon(k) = 0, \quad (30)$$

where

$$b_\varepsilon = (1-\delta)(1+g) + \delta(W^\varepsilon_{DC} - W^\varepsilon_{DD}), \quad (31)$$

[12] From (29), while $W^\varepsilon_{aa'}(k)$ is determined by $\pi^\varepsilon_{aa'}(k-1)$, it is independent of $\pi^\varepsilon_{aa'}(k)$.

and

$$c_\varepsilon(k) = \delta W^\varepsilon_{DD} - W^\varepsilon_{aa'}(k). \tag{32}$$

At $\varepsilon = 0$, equation (30) admits a solution $\pi^0_{aa'}(k)$ that is independent of $k$ and equals $\frac{-c_0}{b_0} = p_{CC}$. We need to establish that $\pi^\varepsilon_{aa'}(k)$ converges to $p_{CC}$ for all $k \geq 1$, uniformly in $k$.

**Theorem 3** *Let $p = (p_{CC}, p_{CD}, p_{DC}, p_{DD})$ be a symmetric completely mixed one period memory equilibrium of the form described in theorem 1. For all $\eta > 0$, there is exists $\varepsilon(\eta) > 0$ such that for all $\varepsilon < \varepsilon(\eta)$, the equilibrium of the perturbed game $\Gamma(\varepsilon)$ given by the probabilities $\pi^\varepsilon_{aa'}(k)$ described above satisfies*

$$|\pi^\varepsilon_{aa'}(k) - p_{CC}| < \eta \quad \forall k \geq 1.$$

**Proof.** First observe that there exists $\xi > 0$ such that

$$|W^\varepsilon_{aa'} - W_{aa'}| \leq \xi\varepsilon \tag{33}$$

for all $aa' \neq CC$. This follows from the fact that there is a unique solution to equations (24), (25), and (26) when $\varepsilon = 0$.

Now we establish inductively that for any $\eta > 0$ there exists $\varepsilon(\eta)$, not depending on $k$, such that $\varepsilon \leq \varepsilon(\eta)$ and $\left|\pi^\varepsilon_{aa'}(k-1) - p_{CC}\right| \leq \eta$ for all $aa'$ imply $\left|\pi^\varepsilon_{aa'}(k) - p_{CC}\right| \leq \eta$ for all $aa'$. This will prove the theorem.

Suppose that $\left|\pi^\varepsilon_{aa'}(k-1) - p_{CC}\right| \leq \eta$. Observe that setting $\varepsilon = 0$ in (29), we have

$$W_{CC} = \frac{(1-\delta)(g-\ell)}{\delta} + W_{DC} + W_{CD} - W_{DD}.$$

Subtracting this equation from (29), we have

$$\begin{aligned} W^\varepsilon_{a'a}(k) - W_{CC} &= (W^\varepsilon_{DC} - W_{DC}) + (W^\varepsilon_{CD} - W_{CD}) - (W^\varepsilon_{DD} - W_{DD}) \\ &\quad + \frac{\delta[(W_{CD} - W^\varepsilon_{CD}) - (W_{DD} - W^\varepsilon_{DD})] - (1-\delta)\varepsilon F^{-1}[1 - \pi^\varepsilon_{a'a}(k-1)]}{\delta \pi^\varepsilon_{aa'}(k-1)}. \end{aligned}$$

From (17) and (18), $\delta/(1-\delta) \geq g$ and so

$$\begin{aligned} |W^\varepsilon_{a'a}(k) - W_{CC}| &\leq 3\varepsilon\xi + \frac{2\varepsilon\xi + \varepsilon/g}{p_{CC} - \eta} \\ &= \varepsilon(3\xi + \frac{2\xi + 1/g}{p_{CC} - \eta}). \end{aligned}$$

Now setting $\varepsilon = 0$ in equation (30), we have that (recall that $p_{CC} = \pi^0_{aa'}(k)$ for all $k$)

$$b_0 p_{CC} + c_0(k) = 0.$$

Subtracting this equation from equation (30), we have

$$(1-\delta)\varepsilon G(\pi^{\varepsilon}_{aa'}(k)) + b_{\varepsilon}(\pi^{\varepsilon}_{aa'}(k) - p_{CC}) + (b_{\varepsilon} - b_0)p_{CC} + c_{\varepsilon}(k) - c_0(k) = 0.$$

Now,

$$\begin{aligned}
|G(\pi^{\varepsilon}_{aa'}(k))| &\leq 1, \\
|c_{\varepsilon}(k) - c_0(k)| &\leq \delta\left|W^{\varepsilon}_{DD} - W_{DD}\right| + \left|W^{\varepsilon}_{aa'}(k) - W_{CC}\right|, \text{ by (32)}, \\
|b_{\varepsilon} - b_0| &\leq 2\delta\varepsilon\xi, \text{ by (31) and (33)}, \\
\text{and} \qquad b_{\varepsilon} &\geq (1-\delta)(1+g) + \delta(W_{DC} - W_{DD}) - 2\varepsilon\delta\xi, \text{ by (31) and (33)}.
\end{aligned}$$

Furthermore, $(1-\delta)(1+g) + \delta(W_{DC} - W_{DD}) > 0$ since it is equal to the denominator in equation (13), so that $b_{\varepsilon} > 0$ for $\varepsilon$ sufficiently small. Consequently,

$$\begin{aligned}
|\pi^{\varepsilon}_{aa'}(k) - p_{CC}| &\leq \frac{1}{b_{\varepsilon}}((1-\delta)\varepsilon + |b_{\varepsilon} - b_0|\, p_{CC} \\
&\qquad + \delta\left|W^{\varepsilon}_{DD} - W_{DD}\right| + \left|W^{\varepsilon}_{aa'}(k) - W_{CC}\right|) \\
&\leq \frac{(1-\delta)\varepsilon + 2\delta\varepsilon\xi p_{CC} + \delta\varepsilon\xi + \varepsilon(3\xi + \frac{2\xi + 1/g}{p_{CC} - \eta})}{(1-\delta)(1+g) + \delta(W_{DC} - W_{DD}) - 2\delta\varepsilon\xi},
\end{aligned}$$

The last expression is less than or equal to $\eta$ if

$$\begin{aligned}
&(1-\delta)\varepsilon + 2\delta\varepsilon\xi p_{CC} + \delta\varepsilon\xi + \varepsilon(3\xi + \frac{2\xi + 1/g}{p_{CC} - \eta}) \\
&\qquad \leq \eta((1-\delta)(1+g) + \delta(W_{DC} - W_{DD}) - 2\delta\varepsilon\xi)
\end{aligned}$$

or

$$\varepsilon \leq \frac{\eta(1-\delta)(1+g) + \delta(W_{DC} - W_{DD})}{(1-\delta) + 2\delta\xi p_{CC} + \delta\xi + 3\xi + \frac{2\xi + 1/g}{p_{CC} - \eta} + 2\delta\xi\eta} \equiv \varepsilon(\eta),$$

and the theorem is proved. ■

## 6. Discussion

To understand the question of the purifiability of mixed strategy equilibria in infinite horizon games, we work with one elegant class of one-period history strategies. Here we have a striking result: with infinite history strategies, such strategies are purifiable. But if we restrict ourselves to one-period history strategies in the perturbed game, then no such strategy is purifiable (for a generic choice of noise distribution). While we conjecture that this negative result extends to allowing all finite-memory strategy profiles in the perturbed game, we have not been able to solve this case.

As noted in the introduction, much of the interest in the purifiability of mixed strategy equilibria in repeated games comes from the literature on repeated game with private monitoring. The systems of equations for the perfect monitoring case can be straightforwardly extended to allow for private monitoring. Unfortunately, the particular arguments that we report exploit the perfect monitoring structure to reduce the infinite system of equations to simple difference equations, and somewhat different arguments are required to deal with private monitoring.

We conjecture that the infinite horizon purification results would extend using general methods for analyzing infinite systems of equations. Intuitively, private monitoring will make purification by finite history strategies harder, as there will be many different histories that will presumably give rise to different equilibrium beliefs that must lead to identical mixed strategies being played, and this should not typically occur. This argument can be formalized for one period histories, but we have not established the argument for arbitrary finite history strategies. However, we believe that the finite history restriction may place substantial bounds on the set of mixed strategies that can be purified in general repeated games, and we hope to pursue this issue in later work.

## References


Bhaskar, V. (1998): "Informational Constraints and the Overlapping Generations Model: Folk and Anti-Folk Theorems," *Review of Economic Studies*, 65(1), 135–149.

——— (2000): "The Robustness of Repeated Game Equilibria to Incomplete Payoff Information," University of Essex.

Bhaskar, V., G. J. Mailath, and S. Morris (2004): "Purification in the Infinitely-Repeated Prisoners' Dilemma," PIER Working Paper 04-004, Uiversity of Pennsylvania.

Bhaskar, V., and I. Obara (2002): "Belief-Based Equilibria in the Repeated Prisoners' Dilemma with Private Monitoring," *Journal of Economic Theory*, 102(1), 40–69.

Bhaskar, V., and E. van Damme (2002): "Moral Hazard and Private Monitoring," *Journal of Economic Theory*, 102(1), 16–39.

Ely, J. C., J. Hörner, and W. Olszewski (2005): "Belief-Free Equilibria in Repeated Games," *Econometrica*, 73(2), 377–416.

Ely, J. C., and J. Välimäki (2002): "A Robust Folk Theorem for the Prisoner's Dilemma," *Journal of Economic Theory*, 102(1), 84–105.

Govindan, S., P. J. Reny, and A. J. Robson (2003): "A Short Proof of Harsanyi's Purification Theorem," *Games and Economic Behavior*, 45(2), 369–374.

Harsanyi, J. C. (1973): "Games with Randomly Disturbed Payoffs: A New Rationale for Mixed- Strategy Equilibrium Points," *International Journal of Game Theory*, 2(1), 1–23.

Kandori, M., and I. Obara (2006): "Efficiency in Repeated Games Revisited: The Role of Private Strategies," *Econometrica*, 74(2), 499–519.

Mailath, G. J., and L. Samuelson (2006): *Repeated Games and Reputations: Long-Run Relationships*. Oxford University Press, New York, NY.

Morris, S. (forthcoming): "Purification," in *New Palgrave Dictionary of Economics*, ed. by S. Durlauf, and L. Blume. MacMillan.

Piccione, M. (2002): "The Repeated Prisoner's Dilemma with Imperfect Private Monitoring," *Journal of Economic Theory*, 102(1), 70–83.

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