PURIFICATION IN THE INFINITELY-REPEATED PRISONERS' DILEMMA

By

V. Bhaskar, George J. Mailath and Stephen Morris

July 2006

COWLES FOUNDATION DISCUSSION PAPER NO. 1571

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.econ.yale.edu/

Purification in the Infinitely-Repeated Prisoners' Dilemma[*]

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

July 13, 2006

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 a class 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 all such strategy profiles are 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 not randomize. Consequently, a mixed strategy equilibrium can be viewed as a pure

[*]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

[‡]Yale University and University of Pennsylvania

[§]Princeton University

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 how these features may lead to a failure of local uniqueness and purification: i.e., 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.[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 stronger efficiency results via private strategies in repeated games with imperfect public monitoring.

[1] See Govindan, Reny, and Robson (2003) for a modern exposition and generalization of Harsanyi (1973).

[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)).

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.[3] 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 be not be purifiable with finite history strategies but will be purifiable with infinite history strategies.

The paper is organized as follows. In Section 2, 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 3. In Section 4, we present the positive purification result for infinite history strategies. Finally, in Section 5, we briefly discuss possible extensions and the private monitoring case.

2. Belief-free Equilibria with Perfect Monitoring

Let $\Gamma(0)$ denote the infinitely-repeated perfect-monitoring prisoners' dilemma with stage game displayed in figure 1. Each player has a discount rate δ. The class of symmetric mixed strategy equilibria Ely and Välimäki (2002) construct can be described as follows: The profiles have one-period memory, with players randomizing in each period with probability $p_{aa'}$ on C after the action profile aa'. 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, and in this sense the equilibria are, to use the language introduced by Ely, Hörner, and Olszewski (2005),

[3] Anticipating the notation from the next section, this occurs, for example, when $g = \ell$ (the incentive to play D is independent of the action of the opponent), so that $p_{CC} = p_{DC}$ and $p_{CD} = p_{DD}$.

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

Figure 1: The unperturbed prisoners' dilemma stage game.

"belief-free." The requirement that after aa', player 1 is indifferent between playing C and D, when player 2 is playing $p_{a'a}$ yields the following system (where $W_{aa'}$ is the value to a player after aa', and the second equality in each displayed equation comes from the indifference requirement):

$$\begin{aligned} W^i_{aa'} &= (1-\delta)\left(p^j_{a'a} + \left(1-p^j_{a'a}\right)(-\ell)\right) \\ &\qquad +\delta\left\{p^j_{a'a}W^i_{CC} + \left(1-p^j_{a'a}\right)W^i_{CD}\right\} \qquad (1) \\ &= (1-\delta)\, p^j_{a'a}(1+g) + \delta\left\{p^j_{a'a}W^i_{DC} + (1-p_{a'a})\, W^i_{DD}\right\}. \qquad (2) \end{aligned}$$

Subtracting (2) from (1) gives

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

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:

$$W^i_{CD} - W^i_{DD} = \frac{(1-\delta)\,\ell}{\delta} \qquad (3)$$

and

$$\begin{aligned} W^i_{CC} - W^i_{DC} &= \frac{(1-\delta)(g-\ell)}{\delta} + W^i_{CD} - W^i_{DD} \\ &= \frac{(1-\delta)\,g}{\delta}. \qquad (4) \end{aligned}$$

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 ℓ, while the cost of doing so is that his continuation value falls from W^i_{CD} to W^i_{DD}. Equation (3) 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 (4) 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.[4]

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 (1) and (2) 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 (3) and (4). 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 (1) for $aa' = CC$ gives

$$p^j_{CC} = \frac{(1-\delta)\,\ell + W^i_{CC} - \delta W^i_{CD}}{(1-\delta)\,(1+\ell) + \delta\left(W^i_{CC} - W^i_{CD}\right)}, \tag{5}$$

for $aa' = CD$ gives

$$p^j_{DC} = \frac{(1-\delta)\,\ell + W^i_{CD} - \delta W^i_{CD}}{(1-\delta)\,(1+\ell) + \delta\left(W^i_{CC} - W^i_{CD}\right)}, \tag{6}$$

for $aa' = DC$ (using (4)) gives

$$p^j_{CD} = \frac{(1-\delta)\,(\ell - g/\delta) + W^i_{CC} - \delta W^i_{CD}}{(1-\delta)\,(1+\ell) + \delta\left(W^i_{CC} - W^i_{CD}\right)}, \tag{7}$$

and, finally, for $aa' = DD$ (using (3))gives

$$p^j_{DD} = \frac{(1-\delta)\,\ell\,(1-1/\delta) + W^i_{CD} - \delta W^i_{CD}}{(1-\delta)\,(1+\ell) + \delta\left(W^i_{CC} - W^i_{CD}\right)}. \tag{8}$$

We have described an equilibrium if the expressions in (5)-(8) are probabilities.

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

$$(1-\delta)\,g/\delta + \delta W^i_{CD} \le (1-\delta)\,\ell + W^i_{CC}, \tag{9}$$

$$(1-\delta)\,\ell \le \delta W^i_{CD}. \tag{10}$$

[4] 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.

The profile in which player 1 plays C with probability $p^1_{aa'}$ and player 2 plays C with probability $p^2_{a'a}$ after aa' in the previous period (and both players play $p^i_{\tilde{a}\tilde{a}'}$ in the first period, for any $\tilde{a}, \tilde{a}' \in \{C, D\}$), where $p^i_{aa'}$ are given by (5)-(8), is an equilibrium. Moreover, (9) and (10) are satisfied for any $0 < W^i_{CD} < W^i_{CC} \leq 1$, for δ sufficiently close to 1.

Proof. We need only verify that (9) and (10) imply that the quantities described by (5)-(8) are probabilities. It is immediate that $p^j_{CD} < p^j_{CC}$ and $p^j_{DD} < p^j_{DC}$, so the only inequalities we need to verify are $0 \leq p^j_{CD}, p^j_{DD}$ and $p^j_{CC}, p^j_{DC} \leq 1$. Observe first that the common denominator in (5)-(8) is strictly positive from $W^i_{CD} \leq W^i_{CC}$.

Now, $p^j_{CC} \leq 1$, since $W^i_{CC} \leq 1$. The quantity p^j_{DC} is no larger than 1 since

$$\begin{aligned} (1-\delta)\,\ell + W^i_{CD} - \delta W^i_{CD} &\leq (1-\delta)\,(1+\ell) + \delta\left(W^i_{CC} - W^i_{CD}\right) \\ \iff W^i_{CD} - \delta W^i_{CC} &\leq 1-\delta, \end{aligned}$$

which is implied by $W^i_{CD} \leq W^i_{CC} \leq 1$.

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

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

which is (9).

Finally, $p^j_{DD} \geq 0$ is equivalent to (10). ■

Indeed, 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.

3. One period memory purification

We now argue that it is impossible to purify equilibria of the type described in Section 2 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 2. 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 $F(.)$. The distribution function has support $[0, 1]$, and 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 $\mathcal{F}$.

	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 2: The perturbed prisoners' dilemma stage game.

An equilibrium of $\Gamma(\varepsilon)$ is γ-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 γ 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 γ-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 γ distance of p.*

Proof. Fix $\gamma' = \min\{p^i_{aa'}\}/2$. Any profile within γ' of p is completely mixed. We first explore the implications of a one period memory equilibrium in $\Gamma(\varepsilon)$ that is γ' close to p.

Fix a one period memory equilibrium of $\Gamma(\varepsilon)$, and denote the ex ante probability of player i playing C after observing the action profile aa' by $\pi^{i\varepsilon}_{aa'}$. In any one period memory equilibrium, 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. Let j denote $3-i$, and note that the action profile from the point of player j is $a'a$. The ex post payoff from C after aa' and given the payoff realization z^i_t, is

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

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

$$V^{i\varepsilon}_{aa'}\left(z^i_t; D\right) = (1-\delta)\,\pi^{j\varepsilon}_{a'a}(1+g) + \delta\left\{\pi^{j\varepsilon}_{a'a} W^{i\varepsilon}_{DC} + \left(1-\pi^{j\varepsilon}_{a'a}\right) W^{i\varepsilon}_{DD}\right\}. \quad (11)$$

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

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

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

$$\begin{aligned}F^{-1}(1-\pi^{i\varepsilon}_{aa'}) = &\frac{1}{(1-\delta)\varepsilon}\left\{(1-\delta)\ell + \delta W^{i\varepsilon}_{DD} - W^{i\varepsilon}_{CD}\right\}\\ &+ \frac{\pi^{j\varepsilon}_{a'a}}{(1-\delta)\varepsilon}\left\{(1-\delta)(g-\ell) + \delta(W^{i\varepsilon}_{DC} + W^{i\varepsilon}_{CD} - W^{i\varepsilon}_{CC} - W^{i\varepsilon}_{DD}\right\}. \quad (12)\end{aligned}$$

Note that the right hand side of (12) is linear 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 (12) as

$$F^{-1}(1-\pi^{i\varepsilon}_{aa'}) = \alpha^{i\varepsilon} + \beta^{i\varepsilon}\pi^{j\varepsilon}_{a'a}. \quad (13)$$

Equation (13) must hold for all $a, a' \in \{C, D\}$. In other words, the points in the set

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

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

If the points in the set

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

are not collinear, then for γ sufficiently small,[5] 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 (13) 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.

Consider now the case when p has only two distinct values of $p^i_{aa'}$ for all i. From (6) and (8), $p^i_{DC} > p^i_{DD}$, while from (5) and (8), 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

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

[5] Note that the bound on γ, while depending on F, is independent of ε.

	h	ℓ
H	$2+\varepsilon_1 z_t^1,\ 3+\varepsilon_2 z_t^2$	$\varepsilon_1 z_t^1,\ 2$
L	$3,\ \varepsilon_2 z_t^2$	$1,\ 1$

Figure 3: The perturbed product-choice stage game.

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

Some insight into nature of the failure of one period purification can be obtained from considering a simpler example of a belief free equilibrium when player 2 is short-lived. The stage game is displayed in figure 3. As usual, player 1 is long-lived, discounting flow payoffs at rate δ, while player 2 is short-lived. In the unperturbed game ($\varepsilon_1 = \varepsilon_2 = 0$), there is a one dimensional manifold of one period memory belief free equilibria. In this manifold, player 1 randomizes with probability $\frac{1}{2}$ on H independent of history, while player 2 plays h with probability p_H after H in the previous period, and with probability $p_L = p_H - 1/(2\delta)$ after L in the previous period (Mailath and Samuelson, 2006, sections 7.6.1 and 12.5).

Observe that in this equilibrium, player 2 randomizes differently after H and L, even though the current randomization of player 1 is independent of his own past play. As for the prisoners' dilemma this raises the possibility that the equilibrium is not purifiable. Consider first only purifying the behavior of the problematic player 2. i.e., setting $\varepsilon_1 = 0$ and taking $\varepsilon_2 > 0$ small. It is straightforward to verify that by appropriately adjusting the randomization of player 1 after H and L (since $\varepsilon_1 = 0$, there is no difficulty in doing so), the incentives for player 2 can be appropriately preserved.[6] On the other hand, when we attempt to *simultaneously* purify player 1 and 2, we obtain an equation of the form (13), and a failure of one period purification. The difficulty of extending this argument to arbitrary finite number of periods is the same as for the prisoners' dilemma.

Remark 1 The theorem asserts that for any fixed mixed strategy equilibrium p, there does not exist a one period purification for generic shock 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.

Remark 2 In an earlier version of this work, Bhaskar, Mailath, and Morris (2004), 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 all others were not.

[6] It is trivial to purify the behavior of player 1 when $\varepsilon_2 = 0$ and taking $\varepsilon_1 > 0$ small.

Remark 3 We assumed that each player received payoff shocks from the same distribution. Clearly, the same argument would go through with asymmetric payoff distributions.

Remark 4 In the earlier version (Bhaskar, Mailath, and Morris, 2004), we asserted that the type of argument reported here would extend to finite memory strategy profiles of any length. But the argument we gave was invalid. While the assertion might be true, we do not have a proof.

Remark 5 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 (cf. remark 2). More importantly, the induction argument extending the negative one period result to arbitrary finite memory strategies is not valid in the simultaneous move case.

4. 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 equilibria of the type described in Section 2. 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 where behavior is identical on elements of the partition. All histories with the same last action profile aa' different from CC are equivalent; denote the associated element of the partition by $(aa',0)$. We write this as $haa' \in (aa',0)$ for all h and $aa' \neq CC$. Two histories ending in CC are equivalent if the most recent action profile different from CC in the two histories is the same, aa' say, and if the same number of occurrences of CC occur in the two histories after the last non -CC action profile, aa'. Denote the associated element of the partition by (aa',k), where k is the number of occurrences of CC after the last non-CC action profile, aa'. Finally, if h is the k-period history in which CC has been played in every period, we write (CC,k) for the singleton element of the partition containing h. 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 strategy in the perturbed game will be 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 will indeed show a *uniform* form of purifiability: the bound on ε 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.[7] This turns out to determine the value function at histories in $(aa', 0)$ for all aa'; we write $W^\varepsilon_{aa'}$ for $W^\varepsilon_{aa'}(0)$. 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 $W^\varepsilon_{aa'}(k+1)$.

Given that a player is to play C with probability π and D with complementary probability, let $G(\pi)$ denote the ex ante expected value of the payoff shock to this player, i.e.

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

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

$$\begin{aligned} W^\varepsilon_{CD} &= (1-\delta)\left\{p_{DC}(1+g) + \varepsilon G(p_{CD})\right\} + \delta\left\{p_{DC} W^\varepsilon_{DC} + (1-p_{DC}) W^\varepsilon_{DD}\right\}, &(14)\\ W^\varepsilon_{DC} &= (1-\delta)\left\{p_{CD}(1+g) + \varepsilon G(p_{DC})\right\} + \delta\left\{p_{CD} W^\varepsilon_{DC} + (1-p_{CD}) W^\varepsilon_{DD}\right\}, &(15)\\ W^\varepsilon_{DD} &= (1-\delta)\left\{p_{DD}(1+g) + \varepsilon G(p_{DD})\right\} + \delta\left\{p_{DD} W^\varepsilon_{DC} + (1-p_{DD}) W^\varepsilon_{DD}\right\}, &(16)\end{aligned}$$

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

and

$$W^\varepsilon_{aa'}(k) = (1-\delta)\left\{\pi^\varepsilon_{a'a}(k)(1+g) + \varepsilon G(\pi^\varepsilon_{aa'}(k))\right\} \\ +\delta\left\{\pi^\varepsilon_{a'a}(k) W^\varepsilon_{DC} + (1-\pi^\varepsilon_{a'a}(k)) W^\varepsilon_{DD}\right\}. \tag{17}$$

As we indicated above, (14), (15), and (16) can be solved for W^ε_{CD}, W^ε_{DC}, and W^ε_{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 $\pi^\varepsilon_{CC}(0) = p_{CC}$).

At the history $h = (a'a, k-1)$, the player with payoff realization $z = F^{-1}[1 - \pi^\varepsilon_{a'a}(k-1)]$ must be indifferent between C and D:

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

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

$$W^\varepsilon_{a'a}(k) = \frac{(1-\delta)(g-\ell)}{\delta} + W^\varepsilon_{DC} + W^\varepsilon_{CD} - W^\varepsilon_{DD} \\ + \frac{(1-\delta)\left\{\ell - \varepsilon F^{-1}[1-\pi^\varepsilon_{a'a}(k-1)]\right\} - \delta[W^\varepsilon_{CD} - W^\varepsilon_{DD}]}{\delta \pi^\varepsilon_{aa'}(k-1)}. \tag{18}$$

This can be re-written (using (3)) as

$$W^\varepsilon_{a'a}(k) = \frac{(1-\delta)(g-\ell)}{\delta} + W^\varepsilon_{DC} + W^\varepsilon_{CD} - W^\varepsilon_{DD} \\ + \frac{\delta[(W_{CD} - W^\varepsilon_{CD}) - (W_{DD} - W^\varepsilon_{DD})] - (1-\delta)\varepsilon F^{-1}[1-\pi^\varepsilon_{a'a}(k-1)]}{\delta \pi^\varepsilon_{aa'}(k-1)}. \tag{19}$$

Examining (19), 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 $W^\varepsilon_{aa'}(k)$,[8] (17) can be re-written as

$$(1-\delta)\varepsilon G(\pi^\varepsilon_{aa'}(k)) + b_\varepsilon \pi^\varepsilon_{aa'}(k) + c_\varepsilon(k) = 0, \tag{20}$$

where

$$b_\varepsilon = (1-\delta)(1+g) + \delta\left(W^\varepsilon_{DC} - W^\varepsilon_{DD}\right), \tag{21}$$

and

$$c_\varepsilon(k) = \delta W^\varepsilon_{DD} - W^\varepsilon_{aa'}(k). \tag{22}$$

[8] ¿From (19), while $W^\varepsilon_{aa'}(k)$ is determined by $\pi^\varepsilon_{aa'}(k-1)$, it is independent of $\pi^\varepsilon_{aa'}(k)$.

At $\varepsilon = 0$, equation (20) 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 $pp = (p_{CC}, p_{CD}, p_{DC}, p_{DD})$ 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

$$|W^\varepsilon_{aa'} - W_{aa'}| \leq \xi\varepsilon \tag{23}$$

for all $aa' \neq CC$. This follows from the fact that there is a unique solution to equations (14), (15), and (16) 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 (19), we have

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

Subtracting this equation from (19), we have

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

Since $\delta/(1-\delta) \geq g$ for cooperation to be possible in the prisoners' dilemma, we have

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

Now setting $\varepsilon = 0$ in equation (20), 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 (20), we have

$$(1-\delta)\varepsilon G\left(\pi_{aa'}^{\varepsilon}(k)\right)+b_{\varepsilon}\left(\pi_{aa'}^{\varepsilon}(k)-p_{CC}\right)+\left(b_{\varepsilon}-b_{0}\right)p_{CC}+c_{\varepsilon}\left(k\right)-c_{0}\left(k\right)=0.$$

Now,

$$\begin{aligned}
\left|G\left(\pi_{aa'}^{\varepsilon}(k)\right)\right| &\leq 1,\\
\left|c_{\varepsilon}\left(k\right)-c_{0}\left(k\right)\right| &\leq \delta\left|W_{DD}^{\varepsilon}-W_{DD}\right|+\left|W_{aa'}^{\varepsilon}(k)-W_{CC}\right|, \text{ by (22)},\\
\left|b_{\varepsilon}-b_{0}\right| &\leq 2\delta\varepsilon\xi, \text{ by (21) and (23)},\\
\text{and} \qquad b_{\varepsilon} &\geq (1-\delta)\left(1+g\right)+\delta\left(W_{DC}-W_{DD}\right)-2\varepsilon\delta\xi, \text{ by (21) and (23)}.
\end{aligned}$$

Furthermore, $(1-\delta)\left(1+g\right)+\delta\left(W_{DC}-W_{DD}\right)>0$ since it is equal to the denominator in equation (5), so that $b_{\varepsilon}>0$ for ε sufficiently small. Consequently,

$$\begin{aligned}
\left|\pi_{aa'}^{\varepsilon}(k)-p_{CC}\right| &\leq \frac{1}{b_{\varepsilon}}((1-\delta)\varepsilon+\left|b_{\varepsilon}-b_{0}\right|p_{CC}\\
&\qquad +\delta\left|W_{DD}^{\varepsilon}-W_{DD}\right|+\left|W_{aa'}^{\varepsilon}(k)-W_{CC}\right|)\\
&\leq \frac{(1-\delta)\varepsilon+2\delta\varepsilon\xi p_{CC}+\delta\varepsilon\xi+\varepsilon\left(3\xi+\frac{2\xi+1/g}{p_{CC}-\eta}\right)}{(1-\delta)\left(1+g\right)+\delta\left(W_{DC}-W_{DD}\right)-2\delta\varepsilon\xi},
\end{aligned}$$

The last expression is less than or equal to η if

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

or

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

and the theorem is proved. ■

5. 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 very 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," University of Essex, University of Pennsylvania, and Yale University.

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.

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.