

Cooperation cannot be sustained in a Discounted Repeated Prisoners' Dilemma with Patient Short and Long Run Players *

Mustafa Oğuz Afacan[†]
Sabancı University

Mehmet Barlo[‡]
Sabancı University

August, 2008

Abstract

This study presents a modified version of the repeated discounted prisoners' dilemma with long and short-run players. In our setting a short-run player does not observe the history that has occurred before he was born, and survives into next phases of the game with a probability given by the current action profile in the stage game. Thus, even though it is improbable, a short-run player may live and interact with the long-run player for infinitely long amounts of time. In this model we prove that under a mild incentive condition on the stage game payoffs, the cooperative outcome path is not subgame perfect no matter how patient the players are. Moreover with an additional technical assumption aimed to provide a tractable analysis, we also show that payoffs arbitrarily close to that of the cooperative outcome path, cannot be obtained in equilibrium even with patient players.

Journal of Economic Literature Classification Numbers: C72; C73; C79

Keywords: Repeated Prisoners' Dilemma; Short-Run and Long-Run Players; Folk Theorem;

*We thank Ahmet Alkan, Mehmet Baç, Guilherme Carmona, Alpay Fliztekin, and Hakan Orbay for their helpful comments. All remaining errors are ours.

[†]Email moguzafacan@su.sabanciuniv.edu for comments and questions

[‡]Email barlo@sabanciuniv.edu for comments and questions.

1 Introduction

The analysis of the discounted repeated prisoners' dilemma reveals that cooperation can be obtained when players are sufficiently patient. Moreover, the perfect Folk Theorem of Fudenberg and Maskin (1986) shows that this conclusion holds for any strictly individually rational payoff. However, assuming that one of the players is short lived changes these conclusions drastically as is pointed out in Fudenberg, Kreps, and Maskin (1990). In fact, in the framework of repeated prisoners' dilemma this assumption reduces the set of equilibrium outcomes to the repetition of the non-cooperative action.

In this study we consider a repeated prisoners' dilemma with a long-run player (player 1), and countably many short run players. In order to provide an easier reading, we describe the model thinking of a soccer club, where player 1 is the owner of the club and the short run players, the coaches. Time is discrete, and every period a coach is born and we refer to a coach born in period t by player $(2.t)$. In the first period, player 1 faces player (2.1) , and they both choose an action in $\{C, D\}$. Then, in the same period a public signal in $\{0, 1\}$ is realized, where 0 is to be interpreted as *failure* and 1 to be *success*. The probability distribution on the set of public signals is determined by the action profile chosen in that period. In case of failure, they both receive zero payoffs and coach $(2.t)$ gets fired, and in the next period player 1 faces coach (2.2) . On the other hand, in case of success, they receive the prisoners' dilemma payoffs (all strictly positive) and coach (2.1) does not get fired, and is active also in period 2. Thus, in any period t , player 1 may face (with some probabilities) any one of players $(2.\tau)$, $\tau \leq t$. We assume that a coach born in period t dies in that period when he is not employed. A t length history in this setting is given by t actions and t public signal, all of which player 1 observes. But player $(2.\tau)$ $\tau \leq t$, does not observe the history (both public and private) before his birth, but the $t - \tau$ tail of a t length history (we work with the convention that the 0 tail of any history is the empty set). All the players discount future payoffs, but not necessarily with the same discount rate.

We show that when payoffs from short run deviations are sufficiently high, sustaining cooperation turns out to be impossible even with patient players. In particular, this study proves that when short run deviation payoffs are sufficiently high, then the cooperative outcome path (which consists of the cooperative action regardless of failures and successes) is not subgame perfect for any discount factor. Moreover, we also show that payoffs arbitrarily close to that of the cooperative outcome path cannot be obtained in equilibrium even with patient players. The main reason of this observation is that, when cooperation is considered a deviation by player 1 is more difficult to be punished because: (1) a coach born in period t must cooperate on the first day that he is active (note that such players do not observe that player 1 has deviated in the

past), and (2) only coaches who faced a deviation by player 1 and became successful are able to punish him.

While our result can be immediately used to provide a simple explanation for why soccer coaches get fired very frequently¹, an important application includes an infinitely repeated version Kyle's market model, due to Kyle (1989), with a long run and many short run traders, where all are *informed*.

In a simple repeated version of Kyle's model² with long and short run players, we imagine that the stage game involves 3 players: Two informed traders, first one is the long run informed investor and the second is the short run informed investor, and a market maker.³ Indeed, while the first player can be thought of as an investor with high job security, the second player can be viewed as a fresh graduate from good economics/business programs who was hired as a trader, but his job is not secure and his continued employment critically depends on his performance (in the short run). We, consider the situation where the information rent of the investors depend on their actions which only take 2 values, high (defection) or low (cooperation). Another simplification that needs to be done is to assume that the market maker is myopic (discounts future returns with discount rate of zero) and cannot condition his behavior on the past. He can identify the real value of the asset with a probability which depends on the action choices of the two traders. Indeed, when the market maker identifies the real value of the asset, it can be considered as a *failure* for the two traders and in this case they obtain zero excess returns. Moreover in this contingency the short run informed investor loses his job, and is replaced by another in the next period. On the other hand, if the market maker cannot identify the real value of the asset, then the excess return that the informed traders obtain (the information rent) depends again on their action choices, and is in the form of a standard prisoners' dilemma. Furthermore, in this state of the world, particular short run informed investor continues to be employed in the next period.

Our results then imply that even with patient investors, cooperation can not be obtained when deviations in the short run are sufficiently beneficial.

¹It is appropriate to point out that in the first 10 weeks of the 2007-2008 season of the Turkcell Superlig (the Premiere Turkish Soccer League) 10 coaches have been fired.

²Kyle's Market Model is a one shot financial economics model for asset pricing in which there are four parties involved in the trade of an asset. The real value of the asset is only known to one of these traders, the *informed* trader, while the other uninformed but rational player is trying guess the value from the total demand for this asset in order to maximize his returns. The third party involved is the noise trader (also called hedger) who is not rational, and has to demand some random amount of this asset, and this provides noise into the model. The final player is the market maker, who does not observe the real value of the asset (and hence is uninformed), and tries to get the efficiency of the market by determining the price based on the total demand he observes. In the equilibrium of this model, the informed trader achieves a surplus, which can be thought as a rent for his information.

³It needs to be mentioned that we do not have an uninformed but rational player and the noise traders.

It needs to be emphasized that restricting a player 2 born in period $t > 1$ not to observe histories (actions and public signals) prior to his birth is an important one. Indeed, the formulation with player $(2.t)$ being able to observe the public signals (but not past action profiles) prior to his birth is more appealing. We need to say that it is not known to us whether or not some versions of our result can be extended to such situations. On the other hand, such endeavors are not trivial. Consider player $(2.t)$ being able to observe the public signals that has occurred before his birth, and to see some implications and related complications consider the following strategy: On the first day he is born, he cooperates if the public signals are all consisting of 1's (i.e. the past has been nothing but success); otherwise, he defects. And, in later periods he cooperates only when player 1 cooperated in the periods that are observable to player $(2.t)$; otherwise, he defects. With this strategy a deviation of player 1 from the cooperative path can be punished more effectively, because his deviation would increase the probability of failure, and thus, the probability of his deviation being punished by a player 2 who was born in a period after player 1's deviation. Thus, a general formulation with player $(2.t)$, $t > 1$ observing public signals in periods $\tau = 1, \dots, t-1$, would call for checking whether or not we can adopt techniques presented in Abreu, Pearce, and Stachetti (1990), and analyzing equilibrium payoffs.

The analysis of the repeated prisoners' dilemma has always been one of the cornerstones in the literature. Our analysis differs from the standard (complete information) versions in the following ways: As in Fudenberg, Kreps, and Maskin (1990) and Fudenberg and Levine (2006) our model features short and long-run players with the important distinction that short-run players may survive with some probability into future phases of the game. Moreover, the second important difference is the additional restriction that short-run players do not observe the history prior to their birth. In that regard, our setup shares some similarities with the analysis of limited memory in the context of the repeated prisoners' dilemma with complete information and long-run players. We refer the reader to Aumann (1981), Neyman (1985), Rubinstein (1986), Kalai and Stanford (1988), Sabourian (1998), Barlo, Carmona, and Sabourian (2008), and Barlo and Carmona (2007) for more on the subject. On the other hand, Cole and Kocherlakota (2005) delivers a similar conclusion to ours in the context of a repeated prisoners' dilemma with imperfect monitoring and finite memory: They prove that for some parameter settings the only strongly symmetric public perfect equilibrium consists of the repetition of the non-cooperative action profile regardless of the discount factor. Meanwhile, other studies of the repeated prisoners' dilemma with imperfect monitoring include Bashkar and Obara (2002), Mailath and Morris (2002), Mailath, Obara, and Sekiguchi (2002), and Piccione (2002).

2 The Model

We consider a similar game to the partnership game of Radner, Myerson, and Maskin (1986). Every period, a long-run (infinitely lived) player, henceforth to be referred to as the first player, is interacting with one of countably many short-run players. Every period a short-run player is born, and the one born in period t will be referred to as player $(2.t)$.

The period interaction is a modified version of the standard prisoners' dilemma: each action profile in $A \equiv \{C, D\} \times \{C, D\}$ is followed by a public signal θ in $\{0, 1\}$, where $\theta = 1$ signals the "success" of the interaction between player 1 and 2 in period t . When $\theta = 0$, both agents do not obtain any payoffs from their interaction. Moreover, player $(2.t')$, $t' \leq t$ that player 1 has played against in period t , is "fired" and player 1 will interact with player $(2.(t+1))$ in the next period. On the other hand, in case of success the players obtain the following payoffs:

$$\begin{array}{cc}
 & C & D \\
 C & (1, 1) & (c, b) \\
 D & (b, c) & (d, d)
 \end{array} \tag{1}$$

where $b > 1 > d > c > 0$ and $\frac{b+c}{2} < 1$. Moreover, then $(2.t')$, $t' \leq t$, is not to be fired and will be the player 2 that player 1 will play against in period $t+1$.

We let $\Pr(\theta = 0|CC) = p_1$, $\Pr(\theta = 0|CD) = \Pr(\theta = 0|DC) = p_2$ and $\Pr(\theta = 0|DD) = p_3$, with $0 < p_1 < p_2 < p_3 < 1$. It is worthwhile to point out that the probability of success decreases with defection.

Consequently, in period t players obtain the following (short-run) returns:

$$\begin{array}{ccc}
 Pl1/Pl(2.t') & C & D \\
 C & (1-p_1), (1-p_1) & (1-p_2)c, (1-p_2)b \\
 D & (1-p_2)b, (1-p_2)c & (1-p_3)d, (1-p_3)d
 \end{array}$$

for $t' \leq t$, where t' is the period in which the player 2 that player 1 faces was born. In order to ensure that the short-run payoffs is given by a prisoners' dilemma, we have the following assumption:

Assumption 1 $(1-p_2)b > (1-p_1)$, $(1-p_3)d > (1-p_2)c$, $(1-p_1) > (1-p_3)d$, $b > 1 > d > c > 0$ and $\frac{b+c}{2} < 1$.

We denote the set of histories by H , any $t-1$ length history consists of signals and action played up to t period,

$$h_t = ((a_1, \theta_1), (a_2, \theta_2), \dots, (a_{t-1}, \theta_{t-1})).$$

Let f_1 be the pure strategy of player 1 such that $f_1(h_t) \in \{C, D\}$ for each period t . We denote the set of all pure strategies of first player by F_1 .

The following assumption will play a critical role in our analysis:

Assumption 2 *Assume that player (2.t), a second player born in period t , is restricted to use pure strategies that do not depend on the history (both public and private) that has happened before he was born.*

Consequently, for any $t' \geq t \geq 1$ and for any $t' - 1$ length history h , let $T^{t'-t}(h)$ be the $t' - t$ tail of h . That is, given that $h = ((a_\tau, \theta_\tau)_{\tau=1}^{t'-1})$, $T^{t'-t}(h) \equiv ((a_\tau, \theta_\tau)_{\tau=t}^{t'-1})$. Obviously, if $t' = t$, $T^0(h) = e$ for all $t' - 1$ length history h . We let $f_{(2.t)}$ be the pure strategy of player (2.t) so that for any $t' \geq t$ and $h_{t'}$ any $t' - 1$ length history, $f_{(2.t)}(h_{t'}) : T^{t'-t}(h_{t'}) \rightarrow \{C, D\}$, thus, $f_{(2.t)} = \{f_{(2.t)}(h_{t'})\}_{t' \geq t}$. Denote the set of all pure strategies for second players by F_2 .⁴

An outcome path $\pi = \{\pi_t\}_{t \in \mathbb{N}}$ where for any t , $\pi_t = ((a_{1,t}, a_{2,t}), \theta_t)$. We denote the set of outcome paths by Π . Moreover, a strategy pair $f = (f_1, f_2)$ induces the set of possible outcome paths $\pi(f) \in A^\infty \equiv A \times A \times \dots$ as follows: $\pi_1(f) = ((f_1(e), f_{(2.1)}(e)), \theta_1)$. Note that if $\theta_1 = 1$, the player 2 that player 1 faces in period 2 is (2.1). Thus, $\pi_2(f) = (f_1(a_1, \theta_1), f_{(2.1)}(a_1, \theta_1), \theta_2)$. But, if $\theta_1 = 0$, then player 1 faces player (2.2), thus, $\pi_2(f) = (f_1(a_1, \theta_1), f_{(2.2)}(e), \theta_2)$. An inductive argument can be used to formalize $\pi_t(f)$ as above. To elaborate on this process formally, we need to define the following function: For any $t - 1$ length history h_t , let $\theta^t = (\theta_1, \dots, \theta_{t-1})$, and define $\iota : \{0, 1\}^{t-1} \rightarrow \{1, \dots, t\}$ by

$$\iota(h_t) = \begin{cases} \tau & \text{if there exists } t' \leq t - 1 \text{ with } \theta_{t'} = 0, \\ & \text{and } \tau = \arg \max_{\{t'' \leq t\}} \{t'' : \theta_{t''-1} = 0\} \\ 1 & \text{otherwise.} \end{cases}$$

In words, $\iota(h_t) = \tau \in \{2, \dots, t\}$ if $\theta^{\tau-1}$ is the last 0 entry in θ^t , and $\iota(h_t) = 1$ if θ^t does not contain any 0 entry. Thus, for any $t - 1$ length history h_t , $(2.\iota(h_t))$ with $\iota(h_t) \leq t$ is the player whom player 1 faces in period t . Thus, for any $t \in \mathbb{N}$,

$$\pi_t(f) = \left(\left(f_1(\{\pi_\tau(f)\}_{\tau=1}^{t-1}), f_{(2.\iota(\{\pi_\tau(f)\}_{\tau=1}^{t-1}))}(T^{t-\iota(\{\pi_\tau(f)\}_{\tau=1}^{t-1})}) \right), \theta_t \right).$$

Players discount future payoffs with the discount factor $\delta_i \in [0, 1)$, $i \in \{1\} \cup \{(2.t) : t \in \mathbb{N}\}$. In

⁴Note that since in period t' the exact identity of the second player is one of $\cup_{\tau=1}^{t'}(2.\tau)$ depending on previous signals that are already included in the $t' - 1$ length history $h_{t'}$. Thus, the birth-period of the particular second player that player 1 faces in period t' , is already given by the history. Hence, the behavior of a second player born in period t , for a $t' - 1$ length history $h_{t'}$, with $t' \geq t$, is described by $f_{(2.t)}(h_{t'}) : T^{t'-t}(h) \rightarrow \{C, D\}$.

particular, for every given outcome path π the payoffs are given by:

$$\begin{aligned}
U_1(\pi) &= (1 - \delta_1) \sum_{t=1}^{\infty} \delta_1^{t-1} u_1(\pi_t), \\
U_{(2,t)}(\pi) &= \begin{cases} (1 - \delta_{(2,1)})(u_2(\pi_1)) & \text{if } t = 1; \\ + \sum_{\tau=2}^{\infty} \delta_{(2,1)}^{\tau-1} \prod_{k=1}^{\tau-1} \Pr(\theta_k = 1 | a_k) u_2(\pi_\tau) & \\ (1 - \delta_{(2,t)})(u_2(\pi_t)) & \text{if } t > 1 \\ + \sum_{\tau=t+1}^{\infty} \delta_{(2,t)}^{\tau-t} \prod_{k=t}^{\tau-1} \Pr(\theta_k = 1 | a_k) u_2(\pi_\tau) & \text{and } \theta_{t-1} = 0; \\ 0 & \text{otherwise.} \end{cases}
\end{aligned} \tag{2}$$

where $u_i(\pi_\tau) = \Pr(\theta_\tau = 1 | a_\tau) u_i(a_\tau)$, and $u_i(a_\tau)$ is as given in the prisoners' dilemma described by equation 1.

3 A Non-Cooperation Result

Define the *cooperative outcome path* to be an outcome path π such that $\pi^t = ((C, C), \theta_t)$ for any $\theta_t \in \{0, 1\}$ and $t \in \mathbb{N}$. Notice that the cooperative outcome path calls for cooperation independently of the signals.

In the following, we show that under Assumptions 1 and 2 and a condition on stage game payoffs, the cooperative outcome path, cannot be supported as a *Nash equilibrium* for all $\delta_i \in [0, 1)$, $i \in \{1\} \cup \{(2.t) : t \in \mathbb{N}\}$.

Proposition 1 *Suppose that*

$$(1 - p_1)(1 - p_3)d + p_1(1 - p_2)b \geq (1 - p_1). \tag{3}$$

Then the cooperative outcome path π is not Nash Equilibrium path for any $\delta_i \in [0, 1)$, $i \in \{1\} \cup \{(2.t) : t \in \mathbb{N}\}$.

Proof. Under the cooperative path π , $U_1(\pi)$ equals $(1 - p_1)$, because, $U_1(\pi) = (1 - \delta_1) \sum_{t=1}^{\infty} \delta_1^{t-1} u_1(\pi^t)$, and $u_1(\pi^t) = \Pr(\theta_t = 1 | a_t) = (1 - p_1)$, for every $t \in \mathbb{N}$.

Next notice that any strategy of player $(2, t)$, $t \in \mathbb{N}$, inducing the cooperative outcome path π , must require player $(2, t)$ to choose C if either $t = 1$, or $t > 1$ and $\theta_{t-1} = 0$. Let $(f_{(2,t)_{t \in \mathbb{N}}})$ be such a pure strategy profile. Define f'_1 be the strategy for the first player such that $f'_1(h^t) = D$ for all h^t and t . Then, below we show that

$$U_1(f'_1, f_2) > (1 - p_1)(1 - p_3)d + p_1(1 - p_2)b.$$

This is because the payoff of player 1 in period $t > 1$ is greater or equal to $(1 - p_1)u_1(D, D) + p_1u_1(D, C) = (1 - p_1)(1 - p_3)d + p_1(1 - p_2)b$. Moreover, this relation holds with strict inequality

when $t = 1$. This is because, $\Pr(\theta_{t-1} = 0 \mid a_{t-1}) \geq p_1$ for all $a_{t-1} \in A$, thus, $\Pr(\theta_{t-1} = 1 \mid a_{t-1}) \leq (1 - p_1)$ for all $t > 1$; and, in the first period player 2 chooses C while player 1 goes for D . Hence, $U_1(f'_1, f_2) > (1 - \delta_1) \sum_{t=1}^{\infty} \delta_1^{t-1} (1 - p_1) u_1(\{D, D\}) + p_1 u_1(\{D, C\})$, showing that $U_1(f'_1, f_2) > (1 - p_1)(1 - p_3)d + p_1(1 - p_2)b$.

Because that

$$U_1(f'_1, f_2) > (1 - p_1)(1 - p_3)d + p_1(1 - p_2)b \geq (1 - p_1) = U_1(f_1, f_2),$$

f'_1 defined as above is a profitable deviation for player 1 from the cooperative outcome path, thus, the result follows. ■

The intuition behind this result is as follows: Due to Assumption 2, the strategies of the short run players do not depend on the past histories (both public and private) before their birth. Therefore, when the long run player, player 1, deviates from the cooperative outcome path, he can be punished by only the short run player who experienced that deviation, and not by short run players whom player 1 may face in the later phases of the game. Hence, condition 3 guarantees that no matter what the value of the discount factors are, player 1 cannot be punished (because of his deviation from the cooperative path) effectively.

In the rest of this section, we analyze whether or not payoffs arbitrarily close to that of the cooperative outcome path can be obtained in subgame perfection.

It is important to point out that for any player $(2.t)$, $t \in \mathbb{N}$, the constant outcome paths π_a , $a \in \{C, D\}^2$, defined by $\pi_a^\tau = (a, \theta_\tau)$ for all $\tau \in \mathbb{N}$ and $\theta_\tau \in \{0, 1\}$, deliver returns given by:

$$\begin{aligned} U_{(2.t)}(\pi_{CC}) &= \left(1 - \frac{p_1}{1 - (1 - p_1)\delta_{(2.t)}}\right), \\ U_{(2.t)}(\pi_{DC}) &= \left(1 - \frac{p_2}{1 - (1 - p_2)\delta_{(2.t)}}\right) b, \\ U_{(2.t)}(\pi_{CD}) &= \left(1 - \frac{p_2}{1 - (1 - p_2)\delta_{(2.t)}}\right) c, \\ U_{(2.t)}(\pi_{DD}) &= \left(1 - \frac{p_3}{1 - (1 - p_3)\delta_{(2.t)}}\right) d. \end{aligned}$$

Therefore, even with the normalization by multiplying period returns with $(1 - \delta_{(2.t)})$, the returns of player $(2.t)$ from constant outcome paths depends on his discount factor $\delta_{(2.t)}$. Consequently, the following restriction helps to obtain a trackable analysis.

Assumption 3 *Suppose that $p_1 = p_2 = p_3$.*

When assumption 3 holds, let $p = p_1 = p_2 = p_3$ and for simplicity, we consider

$$\tilde{U}_{(2.t)} = \left(\frac{1 - \delta_{(2.t)} + \delta_{(2.t)}p}{1 - \delta_{(2.t)}}\right) U_{(2.t)},$$

and because $p \in (0, 1)$, $\tilde{U}_{(2,t)}$ is a linear transformation of $U_{(2,t)}$.

Then, $\tilde{U}_{(2,t)}(\pi_{CC}) = (1-p)$, $\tilde{U}_{(2,t)}(\pi_{DC}) = (1-p)b$, $\tilde{U}_{(2,t)}(\pi_{CD}) = (1-p)c$, and $\tilde{U}_{(2,t)}(\pi_{DD}) = (1-p)d$.

The following Proposition proves that payoffs sufficiently close to the one of the cooperative outcome path cannot be obtained in subgame perfection.

Proposition 2 *Suppose that Assumptions 1, 2, 3, and the condition given in inequality 3 hold. Then there exists $\varepsilon > 0$ and $\bar{\delta}_1 < 1$, such that for all $\varepsilon' < \varepsilon$, every payoff $u = (u_1, (u_{(2,t)})_{t \in \mathbb{N}})$ with $\|(u_1, u_{(2,1)}) - ((1-p), (1-p))\| \leq \varepsilon'$, and $\|(u_1, u_{(2,t)}) - ((1-p), p(1-p))\| \leq \varepsilon'$ for all $t \in \mathbb{N}$, cannot be obtained with a Nash equilibrium pure strategy profile for all $\delta_1 \geq \bar{\delta}_1$.*

Proof. Let $\varepsilon > 0$ be sufficiently small so that $\varepsilon < (1 - \frac{b+c}{2})$, a condition needed in order to use Assumption 1 in the following analysis.

Define r_δ on $\{0, 1\}^\infty$ for $\delta \in [0, 1)$ by $r_\delta(\zeta) = (1-\delta) \sum_{t \in \mathbb{N}} \delta^{t-1} \zeta_t$, for $\zeta \in \{0, 1\}^\infty$. Note that r_δ is continuous on $\{0, 1\}^\infty$, and $\{0, 1\}^\infty$ is compact (due to Tychonoff's Theorem) with the product topology. Moreover, it is clear that $r_\delta(\zeta)$ is continuous for every $\delta \in [0, 1)$. Thus, because that for all $\delta \in [0, 1)$, $r_\delta(\zeta) \in [0, 1]$, we have: For any $\delta_n \rightarrow \delta \in [0, 1)$ and $\zeta_n \rightarrow \zeta$, $\lim_{n \rightarrow \infty} r_{\delta_n}(\zeta_n) = r_\delta(\zeta)$. Moreover, when δ_n tends to 1 and $\zeta_n \rightarrow \zeta$, we have, $\lim_{\delta_n \rightarrow 1} r_{\delta_n}(\zeta_n) = \liminf_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \zeta_t \equiv r(\zeta)$, that is, the fraction of 1's in ζ . This is because, (1) Theorem 7.9 of Rudin (1976) shows that r_{δ_n} converges uniformly on r on $\{0, 1\}^\infty$ since $\lim_n r_{\delta_n}(\zeta) = r(\zeta)$ for every $\zeta \in \{0, 1\}^\infty$ and $\sup_{\zeta \in \{0, 1\}^\infty} |r_{\delta_n}(\zeta) - r(\zeta)|$ converges 0 as n tends to infinity. (2) because r_{δ_n} is a sequence of continuous functions converging uniformly to r on $\{0, 1\}^2$, following the same arguments needed to solve exercise 9 from chapter 7 of Rudin (1976), suffices to establish that for any $\delta_n \rightarrow 1$ and $\zeta_n \rightarrow \zeta$, $\lim_{\delta_n \rightarrow 1} r_{\delta_n}(\zeta_n) = r(\zeta)$.

For any given strategy profile $f_2 = \{f_{(2,t)}\}_{t \in \mathbb{N}}$, let $\zeta(f_2) \in \{0, 1\}^\infty$ be defined by

$$\zeta_t(f_2) = \begin{cases} 1 & \text{if } t = 1 \text{ and } f_{(2,1)}(e) = C, \\ & \text{or } t > 1 \text{ and for any } h \text{ with length } t-1 \text{ and } \theta_{t-1} = 0, f_{(2,t)}(h) = C, \\ 0 & \text{otherwise.} \end{cases}$$

Let $\varepsilon' < \varepsilon$, and consider any payoff $u = (u_1, (u_{(2,t)})_{t \in \mathbb{N}})$ as described in the statement of the Proposition. We restrict attention to strategy profiles f and $\delta \in [0, 1)$ with

$$\|(U_1^\delta(f) - U_{(2,t)}^\delta(f)) - (u_1, u_{(2,t)})\| < \frac{\varepsilon'}{2}, \quad (4)$$

for all $t \in \mathbb{N}$.

Consider a deviation for player 1, f'_1 , in which player 1 chooses D independent of the past, i.e. $f'_1(h) = D$ for all h . Then,

$$\begin{aligned}
U_1^\delta(f'_1, f_2) &\geq (1 - \delta)(1 - p) \left(1_{\zeta_1(f_2)=1} u_1(D, C) + 1_{\zeta_1(f_2)=0} u_1(D, D) \right) \\
&\quad + (1 - \delta)\delta(1 - p) \left((1 - p)u_1(D, D) + p \left(1_{\zeta_2(f_2)=1} u_1(D, C) + 1_{\zeta_2(f_2)=0} u_1(D, D) \right) \right) \\
&\quad + (1 - \delta)\delta^2(1 - p) \left(((1 - p)^2 + p(1 - p)) u_1(D, D) + \right. \\
&\quad \quad \left. (1 - ((1 - p)^2 + p(1 - p))) \left(1_{\zeta_3(f_2)=1} u_1(D, C) + 1_{\zeta_3(f_2)=0} u_1(D, D) \right) \right) \\
&\quad + \dots
\end{aligned} \tag{5}$$

Note that in 5, the player 2 experiencing and surviving after player 1's deviation, is choosing D . If this were not the case, player 1's deviation would be even more profitable. Moreover, because for all $t > 1$, the probability of the opponent of the first player being born before the time period t is given by $(1 - p)^{t-1} + p(1 - p)^{t-2} + p(1 - p)^{t-3} + \dots + p(1 - p)^{t-(t-2)} + p(1 - p)^{t-(t-1)} = (1 - p)$, inequality 5 is reduced to

$$U_1^\delta(f'_1, f_2) \geq (1 - \delta)(1 - p) \left(g_1 + p \sum_{t=2}^{\infty} \delta^{t-1} g_t \right) + \delta(1 - p)^2 u_1(D, D), \tag{6}$$

where $g_t \equiv 1_{\zeta_t(f_2)=1} u_1(D, C) + 1_{\zeta_t(f_2)=0} u_1(D, D)$.

Due to Assumptions 1 and 3, $r_\delta(\zeta(f_2))$ increases to 1 in a continuous manner when $\varepsilon \rightarrow 0$ and $\delta \rightarrow 1$ for f such that f satisfies inequality 4. Thus, the right hand side of inequality 6 tends to $p(1 - p)b + (1 - p)^2 d$ which is strictly greater than $(1 - p)$ due to the condition given in inequality 3.

Thus, when $\varepsilon \rightarrow 0$ and $\delta \rightarrow 1$, for any strategy profile f satisfying 4, $U_1^\delta(f'_1, f_2) > U_1^\delta(f_1, f_2)$. Hence, there exists $\varepsilon > 0$ and $\bar{\delta}_1 < 1$ such that the conclusion of the Proposition holds, an observation finishing the proof. ■

References

- ABREU, D., D. PEARCE, AND E. STACHETTI (1990): "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring," *Econometrica*, 58(5), 1041–1063.
- AUMANN, R. (1981): "Survey of Repeated Games," in *Essays in Game Theory and Mathematical Economics in Honor of Oskar Morgenstern*. Bibliographisches Institut, Mannheim.
- BARLO, M., AND G. CARMONA (2007): "Folk Theorems for the Repeated Prisoners' Dilemma with Limited Memory and Pure Strategies," Sabanci University, Universidade Nova de Lisboa and University of Cambridge.

- BARLO, M., G. CARMONA, AND H. SABOURIAN (2008): “Bounded Memory with Finite Action Spaces,” Sabanci University, Universidade Nova de Lisboa and University of Cambridge.
- BASHKAR, V., AND I. OBARA (2002): “Belief-Based Equilibria in the Repeated Prisoners’ Dilemma with Private Monitoring,” *Journal of Economic Theory*, 102(1), 40–69.
- COLE, H., AND N. KOCHERLAKOTA (2005): “Finite Memory and Imperfect Monitoring,” *Games and Economic Behavior*, 53, 59–72.
- FUDENBERG, D., D. M. KREPS, AND E. MASKIN (1990): “Repeated Games with Long-run and Short-run Players,” *Review of Economic Studies*, 57(4), 555–73.
- FUDENBERG, D., AND D. LEVINE (2006): “A Dual-Self Model of Impulse Control,” *American Economic Review*, 96(5), 1449–1476.
- FUDENBERG, D., AND E. MASKIN (1986): “The Folk Theorem in Repeated Games with Discounting and with Incomplete Information,” *Econometrica*, 54, 533–554.
- KALAI, E., AND W. STANFORD (1988): “Finite Rationality and Interpersonal Complexity in Repeated Games,” *Econometrica*, 56, 397–410.
- KYLE, A. S. (1989): “Informed Speculation with Imperfect Competition,” *The Review of Economic Studies*, 56(3), 317–355.
- MAILATH, G., AND S. MORRIS (2002): “Repeated Games with Almost-Public Monitoring,” *Journal of Economic Theory*, 102, 189–228.
- MAILATH, G., I. OBARA, AND T. SEKIGUCHI (2002): “Maximum Efficient Equilibrium Payoff in the Repeated Prisoners’ Dilemma,” *Journal of Economic Theory*, 102(1), 96–122.
- NEYMAN, A. (1985): “Bounded Complexity Justifies Cooperation in the Finitely Repeated Prisoner’s Dilemma,” *Economic Letters*, 9, 227–229.
- PICCIONE, M. (2002): “The Repeated Prisoner’s Dilemma with Imperfect Private Monitoring,” *Journal of Economic Theory*, 102(1), 70–83.
- RADNER, R., R. MYERSON, AND E. MASKIN (1986): “An Example of a Repeated Partnership Game with Discounting and Uniformly Inefficient Equilibria,” *Review of Economic Studies*, 53, 59–70.

RUBINSTEIN, A. (1986): “Finite Automata Play the Repeated Prisoner’s Dilemma,” *Journal of Economic Theory*, 39, 83–96.

RUDIN, W. (1976): *Principles of Mathematical Analysis*. Mc Graw Hill, 3rd edn.

SABOURIAN, H. (1998): “Repeated Games with M -period Bounded Memory (Pure Strategies),” *Journal of Mathematical Economics*, 30, 1–35.