



## Cooperation in a modified version of the finitely repeated prisoners' dilemma game

Maarten C.W. Janssen\*, Joeri Gorter, Sjoerd van de Meerendonk

*Department of Economics-Micro, Erasmus University, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands*

Received 20 July 1995; received in revised form 6 March 1996

---

### Abstract

In this paper we consider a series of finitely repeated prisoners' dilemma games in which the pay-off the players receive in a period depends on how they have played the game in the past. We show that this modification of the finitely repeated prisoners' dilemma game makes cooperation a feasible equilibrium configuration in the beginning of play.

*JEL classification:* C7; C72

*Keywords:* Prisoners' dilemma; Finitely repeated games

---

### 1. Introduction

It is widely acknowledged that cooperation is not possible in the finitely repeated prisoners' dilemma (PD) when there is common knowledge of rationality among the players. The reason is that the equilibrium strategies can be obtained by reasoning backwards from the last period. Thus, in equilibrium today's behavior cannot influence future behavior or pay-offs. The result is based on a Backwards Induction argument. In order to account for rational cooperation in finitely repeated PD games the backwards induction logic has to be broken. This paper explores one way of doing so.

Two players play a finite number of PD games. The distinguishing feature of the finite-horizon supergame is that the pay-offs of the PD game they play in a certain period depend on the way they have played the PD game in the past. In particular, future pay-offs decrease if players do not cooperate in the present period. It is clear that this gives

---

\* Corresponding author.

players an incentive to behave cooperatively. In this paper, we show that under certain conditions these incentives are large enough to break the backwards induction logic in the beginning of the game, that is, a subgame-perfect Nash equilibrium exists in which the players cooperate in the first number of periods.

The assumption that future pay-offs negatively depend on today's non-cooperation seems to hold in at least two economically significant cases. The first situation that is amenable to the logic of our model is a common resource (e.g., fish) that is exploited by at least two players. (An early reference on dynamic fishing games is Levhari and Mirman, 1980). In such situations a large exploitation (catch) by one of the players in one period leads to higher exploitation costs (lower pay-offs) in future periods. Recently, Dutta and Sundaram (1993, 1993) have shown that overexploitation need not result in an infinite horizon game. Our results can be interpreted as a confirmation of Dutta and Sundaram's finding for finite horizon games. A second situation to which the model applies is a cartel in which durable goods are produced. A producer who exceeds his agreed upon quota affects future demand and this will result in lower future prices and pay-offs for all producers.

The paper is organized as follows. Section 2 presents a game in which the negative dependence of future pay-offs on today's defection is modelled in a simple way. Section 3 provides the main intuition for the result in a 2-period example and Section 4 gives the main result and its proof.

## 2. The model

Two players play a version of a game for  $T$  consecutive periods. In each period the game is a PD game and the players can choose to cooperate ( $C$ ) or to defect ( $D$ ). The basic version of the game they play is described in the matrix below:

$$\begin{array}{c} C \\ D \end{array} \begin{pmatrix} C & D \\ \pi_{cc}, \pi_{cc} & \pi_{cd}, \pi_{dc} \\ \pi_{dc}, \pi_{dc} & \pi_{dd}, \pi_{dd} \end{pmatrix},$$

where  $\pi_{dc} > \pi_{cc} > \pi_{dd} > \pi_{cd}$ . The periods are indexed by  $t$  and  $t = 1 \dots T$ .

The negative dependence of future pay-offs on today's defection can be modelled in a large variety of ways. What the two examples considered in the Introduction seem to have in common is that (i) the larger the number of players defecting in a particular period, the lower future pay-offs will be and (ii) if defection occurs only once, pay-offs decline immediately in the next period, but gradually restore themselves to their initial values. For example, the stock of fish diminishes if a fisherman defects, but gradually restores itself if no further defections occur. The same holds true for a durable good cartel.

The two aspects that these examples have in common are captured in the following way. First, if one player defects, the next period pay-offs of both players are reduced by an amount given by  $\alpha_{dc}$ . On the other hand, if two players defect, next period pay-offs are reduced by  $\alpha_{dd}$  and  $\alpha_{dd} > \alpha_{dc}$ . Second, the reductions in pay-offs in a period that is  $\tau$  periods away from the period of defection are a fraction  $\beta^{\tau-1}$  of the initial reduction, where  $\beta < 1$ . For example, if both players defect in period  $t$ , then the pay-offs in period

$t + 1$  are reduced by an amount of  $\alpha_{dd}$ , the pay-offs in period  $t + 2$  by  $\beta\alpha_{dd}$ , in period  $t + 3$  by  $\beta^2\alpha_{dd}$ , and so on. For convenience we will also assume that the discount factor equals unity.

Note that the above is just one way to formalize the negative dependence of future pay-offs on today's defection. The qualitative aspects of the results we will obtain are, however, robust to various alternative formulations. Also note that the Prisoners' Dilemma character of the original game is preserved in all periods no matter how the game has been actually played in the past. This is why we speak of the possibility of rational cooperation in finitely repeated prisoners' dilemma games.

### 3. A two-period example

To understand the intuition for the main result we will derive in the next section it may be helpful to consider a 2-period example. If the game is played twice, it is clear that the players will defect in the second period. If only one of them defects in the first period, their second period pay-offs will be  $\pi_{dd} - \alpha_{cd}$ . On the other hand, if they both cooperate in the first period, their second period pay-offs will be equal to  $\pi_{dd}$ . Accordingly, the cost of an early defection is equal to  $\alpha_{cd}$ . If this cost is larger than the benefits of defection in the first period, more precisely if  $\pi_{dc} - \pi_{cc} < \alpha_{dc}$ , then there exists an equilibrium in which both players cooperate in the first period. Thus, the negative dependence of future pay-offs on today's defection creates an incentive for players to cooperate in the beginning of play.

### 4. Results

In this section we will show that for a class of parameter values a subgame-perfect Nash equilibrium exists in which the players choose to cooperate in the beginning of the game. The strategies that support the cooperative outcome are trigger strategies. An upper bound is determined for the number of periods in which cooperation can be an equilibrium outcome.

*Definition 1:* Let  $\alpha_{it}$  denote the action taken by player  $i$  in period  $t$ . A trigger strategy for player  $i$ , denoted by  $s_i$ , is given by

- (a)  $a_{it} = C$ ;
- (b)  $a_{it} = C$  for  $t = 2 \dots t^* - 1$  if  $a_{jt} = C$ ,  $j = 1, 2$ ,  $\tau = 1 \dots t - 1$ ; otherwise  $a_{it} = D$ .
- (c)  $a_{it} = D$  for  $t = t^* \dots T$ .

In order to prove the existence of a trigger-strategy equilibrium for certain parameter values, two different ways of a possible deviation need to be distinguished. First, a player might choose to play  $D$  before  $t^*$  and, second, a player might choose to play  $C$  after  $t^*$ . For presentational purposes, it is useful to write down the pay-offs in a trigger-strategy equilibrium and also the pay-offs of the two possible deviations given that the other player plays the trigger strategy. This is done in the following three 'facts.'

**Fact 1:** The trigger strategy equilibrium pay-offs, denoted by  $\pi^*$ , are given by

$$(t^* - 1)\pi_{cc} + (T - t^* + 1)\pi_{dd} - \alpha_{dd} \sum_{i=1}^{T-t^*} \frac{1 - \beta^i}{1 - \beta}.$$

To understand this expression note that the defection in period  $t^*$  leads to a reduction in future pay-offs of  $\alpha_{dd} + \beta\alpha_{dd} + \beta^2\alpha_{dd} + \dots + \beta^{T-t^*-1}\alpha_{dd}$ . This can be rewritten as  $\alpha_{dd}1 - \beta^{T-t^*}/1 - \beta$ . The defection in period  $t^* + 1$  results in a similar way in a reduction of future pay-offs of  $\alpha_{dd}1 - \beta^{T-t^*-1}/1 - \beta$ , and so on. Adding all the reductions in future pay-offs of subsequent defections yields the expression  $\alpha_{dd} \sum_{i=1}^{T-t^*} 1 - \beta^i/1 - \beta$ . Similar calculations are at the basis of Facts 2 and 3.

**Fact 2:** The pay-off of defecting by playing  $D$  ('mean' deviation) in period  $t_m \leq t^* - 1$ , given the trigger strategy of the other player, is denoted by  $\pi_m(t_m)$ , and given by

$$(t_m - 1)\pi_{cc} + \pi_{dc} - \alpha_{dc} \frac{1 - \beta^{T-t_m}}{1 - \beta} + (T - t_m)\pi_{dd} - \alpha_{dd} \sum_{i=1}^{T-t_m-1} \frac{1 - \beta^i}{1 - \beta}.$$

**Fact 3.** The pay-off of defecting by playing  $C$  ('nice' deviation) in periods  $t^* \leq t < t_n$ , given the trigger strategy of the other player, is denoted by  $\pi_n(t_n)$  and given by

$$(t^* - 1)\pi_{cc} + (t_n - t^*)\pi_{cd} - \alpha_{cd} \sum_{i=T-t_n+1}^{T-t^*} \frac{1 - \beta^i}{1 - \beta} + (T - t_n + 1)\pi_{dd} - \alpha_{dd} \sum_{i=1}^{T-t_n} \frac{1 - \beta^i}{1 - \beta}.$$

Note that in Fact 3 we have only considered playing  $C$  a number of times immediately after  $t^*$ . It is obvious that if it is optimal at all to play  $C$  after  $t^*$ , it should be so for some periods immediately after  $t^*$ .

As in the infinitely repeated PD, cooperation is not possible for all possible configurations of the pay-off matrix. Cooperation is not possible if the pay-off  $\pi_{dc}$  is 'too large' compared to  $\pi_{cc}$  and  $\pi_{dd}$ . A necessary condition for cooperation to be a possible equilibrium phenomenon is given in the Lemma below.

**Lemma:** A trigger-strategy equilibrium with an initial phase of  $(C,C)$  can exist only if  $\pi_{dc} - \pi_{cc} < \alpha_{dc}/(1 - \beta)$ .

*Proof.* We compare  $\pi^*$  with  $\pi_m(t^* - 1)$ . Using Facts 1 and 2, it is easily seen that  $\pi^* \geq \pi_m(t^* - 1)$  if, and only if,

$$\pi_{cc} - \pi_{dc} + \alpha_{dc} \frac{1 - \beta^{T-t^*+1}}{1 - \beta} \geq 0.$$

This inequality can only be satisfied if  $\pi_{dc} - \pi_{cc} < \alpha_{dc}/(1 - \beta)$ . ■

The condition  $\pi_{dc} - \pi_{cc} < \alpha_{dc}/(1 - \beta)$  has a natural interpretation. The difference  $\pi_{dc} - \pi_{cc}$  is the additional pay-off a player gets on defecting when the other player still cooperates. On the other hand,  $\alpha_{dc}/(1 - \beta)$  is the supremum of the cost of an early

defection (when  $T$  approaches infinity). If the benefit of defecting is larger than the maximal cost, cooperation cannot be sustained in equilibrium. We will restrict ourselves to the case in which the condition formulated in the Lemma holds.

*Proposition:* Suppose  $\pi_{dc} - \pi_{cc} < \alpha_{dc}/(1 - \beta)$  and  $\pi_{dd} - \pi_{cd} > (\alpha_{dd} - \alpha_{dc})/(1 - \beta)$ . A trigger strategy equilibrium with an initial phase of cooperation exists for all  $T$  such that

$$T > \frac{\log \left[ 1 - \frac{1-\beta}{\alpha_{dc}} (\pi_{dc} - \pi_{cc}) \right]}{\log \beta}.$$

Moreover, this equilibrium is subgame-perfect.

*Proof.* The proof proceeds in five steps.

(i) We first show that if one of the player wants to deviate from the equilibrium path before  $t^*$  he will do so in period  $t^* - 1$ . In order to prove this, we will prove that  $\pi_m(t_m - 1) < \pi_m(t_m)$  for all  $t_m \leq t^* - 1$ . Using Fact 2 it is easily seen that  $\pi_m(t_m - 1) < \pi_m(t_m)$  if and only if

$$\pi_{cc} - \pi_{dd} + \alpha_{dc} \frac{\beta^{T-t_m} - \beta^{T-t_m+1}}{1 - \beta} + \alpha_{dd} \frac{1 - \beta^{T-t_m}}{1 - \beta} > 0.$$

As  $\pi_{cc} > \pi_{dd}$  this inequality is always satisfied. The intuition behind this step is that the ‘cost’ of defection increases with the number of periods that is still ahead, while the ‘gain’ of defection remains constant.

(ii) We next show that if a player wants to deviate from the equilibrium path after  $t^*$ , he will do so in period  $t^*$ . In order to prove this, we will prove that  $\pi_n(t_n) > \pi_n(t_n + 1)$  for all  $t_n \geq t^*$ . Using Fact 3 one can calculate that  $\pi_n(t_n) < \pi_n(t_n + 1)$  if and only if

$$\pi_{dd} - \alpha_{dd} \frac{1 - \beta^{T-t_n}}{1 - \beta} > \pi_{cd} - \alpha_{cd} \frac{1 - \beta^{T-t_n}}{1 - \beta}.$$

is satisfied. As  $\alpha_{dd} > \alpha_{cd}$ , this inequality is satisfied if

$$\pi_{dd} - \alpha_{dd} \frac{1}{1 - \beta} > \pi_{cd} - \alpha_{cd} \frac{1}{1 - \beta}.$$

This is exactly the condition on the pay-off parameters specified in the proposition.

(iii) The third step in the proof is to show that the players do not prefer to deviate from the trigger strategy equilibrium just before  $t^*$ , that is,  $\pi^* \geq \pi_m(t^* - 1)$ . In the proof of the Lemma we have stated that this is the case if

$$\pi_{cc} - \pi_{dc} + \alpha_{dc} \frac{1 - \beta^{T-t^*+1}}{1 - \beta} \geq 0.$$

Rewriting this inequality gives

$$\beta^{T-t^*+1} \leq 1 - \frac{1-\beta}{\alpha_{dc}} (\pi_{dc} - \pi_{cc}).$$

Given the condition of the Lemma and the Proposition, the right hand side of this inequality is positive so that the inequality can be rewritten as

$$T - t^* + 1 \geq \frac{\log \left[ 1 - \frac{1-\beta}{\alpha_{dc}} (\pi_{dc} - \pi_{cc}) \right]}{\log \beta}.$$

As the critical value for  $t^*$  should be larger than 1, this expression gives the lower bound on  $T$  that is stated in the proposition. The intuition behind this step is that  $T$  should be relatively large to make the cost of an early defection larger than the potential gain.

(iv) A fourth step in the proof is to show that  $\pi^* \geq \pi_n(t^*)$ . Using Facts 1 and 3, it is easily shown that this is the case if

$$\pi_{dd} - \alpha_{dd} \frac{1 - \beta^{T-t^*}}{1 - \beta} > \pi_{cd} - \alpha_{cd} \frac{1 - \beta^{T-t^*}}{1 - \beta}.$$

This is a similar condition as the one encountered in Step (ii).

(v) Steps (i)–(iv) have established that the proposed trigger strategy is a Nash equilibrium if  $T$  is large enough. The last step in the proof is to show that this equilibrium is subgame-perfect. For every decision node which is not on the equilibrium path it is true that either at least one of the players has played  $D$  before  $t^*$  or at least one of the players has played  $C$  from  $t$ . In either case, the equilibrium strategy prescribes the playing of  $D$  in all future periods. Deviating from this cannot be optimal as it has been shown in Step (ii) that playing  $C$  from  $t^*$  cannot be optimal when the other player plays  $D$  under the condition specified in the proposition. The same argument, however, holds true for  $t < t^*$ . ■

The proposition demonstrates that rational cooperation is possible if  $\pi_{dd} - \pi_{cd} > (\alpha_{dd} - \alpha_{dc})/(1 - \beta)$  and if the number of periods is large enough. The condition on the pay-offs and the  $\alpha$ 's is needed in order to guarantee that none of the players deviates 'nicely,' that is, cooperates for more periods than specified by the trigger strategy.

The proof of the proposition (see Step (iii)) also gives an upper bound on the number of periods in which cooperation can take place in equilibrium. This upper bound is given by

$$T + 1 - \frac{\log \left[ 1 - \frac{1-\beta}{\alpha_{dc}} (\pi_{dc} - \pi_{cc}) \right]}{\log \beta}.$$

## **References**

- Dutta, P. and R. Sundaram, 1993, How different can strategic models be?, *Journal of Economic Theory*, 60, 42–61.
- Dutta, P. and R. Sundaram, 1993, The tragedy of the commons?, *Economic Theory*, 3, 413–426.
- Levhari, D. and L. Mirman, 1980, The great fish war: An example using a dynamic Cournot–Nash solution, *Bell Journal of Economics*, 11, 322–334.