Acquired Cooperation in Finite-Horizon Games*

IGAL MILCHTAICH[a]

and

AVI WEISS[b]

Bar-Ilan University

December 2002

Abstract

When a prisoner's-dilemma-like game is repeated any finite number of times, the only equilibrium outcome is the one in which all players defect in all periods. However, if cooperation among the players changes their perception of the game by making defection increasingly less attractive, then players may be willing to cooperate in late periods in which unilateral defection has become unprofitable. In this case, cooperation may also be attainable in the first period, since defection then can effectively be punished by cessation of cooperation by all the other players. In this paper, we explore this possibility and consider conditions guaranteeing the players' willingness to cooperate also in the middle periods, in which defection is more profitable than later on, and at the same time, punishments are less effective than at the beginning. These conditions are sufficient for the existence of an equilibrium in which players cooperate in all periods.

---

* We would like to thank Eliakim Katz, with whom the ideas presented in this paper were originally conceived.

[a] E-mail: milchti@mail.biu.ac.il Personal homepage: http://faculty.biu.ac.il/~milchti

[b] E-mail: weissa1@mail.biu.ac.il Personal homepage: http://faculty.biu.ac.il/~weissa1

1 Introduction

The set of equilibria in the finitely repeated prisoner's dilemma differs sharply from that in the infinitely repeated game. In the former, the equilibrium path and payoffs are unique: at equilibrium, both players defect in all periods. In the latter, the sets of equilibrium paths and payoff vectors are both considerably larger. Specifically, any feasible and individually rational payoff vector is obtained at some equilibrium of the game, and there are equilibria in which both players cooperate in all periods. This "discontinuity at infinity" prompted researches to look for variants, or perturbations, of the finitely repeated prisoner's dilemma that also admit equilibria of the kind just mentioned. In this paper, we present and explore one such variant, in which repeated cooperation may be an equilibrium outcome.

Our basic idea is very simple. As the standard backward-induction argument demonstrates, the fundamental impediment to cooperation in the finitely repeated prisoner's dilemma is that, in the last period, the players' action profile must be an equilibrium in the one-shot game. This entails that, regardless of history, both players must defect in that period. Consequently, defection in any <u>preceding</u> period cannot be deterred by the threat of responding to it in kind in the last period. Such a threat would be hollow, since, in any subgame perfect equilibrium at least, defection in the last period takes place <u>anyway</u>. In fact, even without the perfection requirement, there can be no <u>last</u> period in which any player cooperates with positive probability. Therefore, the probability of cooperation ever occurring is zero. In this paper, we diverge from the classical model by allowing the players' preferences over the possible outcomes in each period to depend on history. Specifically, we assume that a history of cooperation makes defection an increasingly less attractive alternative as time progresses. One possible interpretation for such an effect of past cooperation on the willingness to cooperate is that defection carries a moral cost, which is greater the longer the history of cooperation. Empirical support for the idea that history may affect people's behavior in prisoner's-dilemma-like games can be found in the work of Brown et al. (2002), which is described in Section 2.

If a history of cooperation indeed tends to affect players' preferences in the way just indicated, then, assuming a sufficiently large number of periods, none of the players would want to defect in the last period. Moreover, in periods close to the last one, the incentive to defect would be weak or nonexistent. This may also affect the players' behavior in the early periods of the game, since defection can now be effectively deterred by the threat of switching from cooperation to repeated defection. However, even if (1) the rate at which preferences evolve is fast enough for defection close to the end of the game to be unprofitable, and (2) the number of periods is large enough for early defection to be punishable, cooperation in all the periods might still not be equilibrium behavior. The problem lies is the middle periods. As the end of the game approaches, players have increasingly less to lose from defection (which leads to the other players defecting in all subsequent periods). At the same time, by assumption, the evolution of their preferences is such that they have increasingly less to gain from defection. The resultant of these two opposing forces may point in different directions in different periods. In particular, it is possible that conditions (1) and (2) above—which may be viewed as boundary conditions—hold, and yet cooperation is not sustainable because it is better to defect in one or more of the intermediate periods. This may have a profound effect on the outcome. Since players can no longer assume that the other players will cooperate in the middle periods, they cannot be deterred from defecting in the early ones. Consequently, the history of cooperation required to make it an equilibrium behavior in the last period may not materialize, with the result that there may be defection throughout all or much of the game.

It appears, then, that in order to know whether cooperation in a given interaction is sustainable, it is necessary to check not only the boundary conditions but also those corresponding to each of the intermediate periods. The main purpose of this paper is to show that, under a particular additional assumption spelled out below, this is, in fact, not necessary. Rather, cooperation is sustainable if and only if the boundary conditions hold, and in this case, all the other conditions hold automatically. Roughly, the condition is that the gain from being the first player to stop cooperating in any period t does not exceed a certain weighted average (with weights that may depend on t) of the gain from doing so one period earlier and one period later. The exact formulation of this result is given in Section 5. Before developing the general result, however, we consider in Sections 3 and 4 two special settings in which the above condition can be given a particularly simple form. These serve to exemplify the potentially wide applicability of our results.

In the first scenario we consider, cooperation makes people increasingly more altruistic towards one another. Specifically, the longer the history of cooperation between them, the more weight each player places on the utility of the other players at the expense of his own. Thus, if cooperation is maintained until a certain period t, each player's payoff function in period $t + 1$ is a weighted average of his own and the other players' payoff functions in period t. As a result, the players progressively internalize the social costs or benefits of their actions. Assuming that cooperation is socially beneficial and the number of periods is sufficiently large, cooperation will be attainable in the last period because the payoff from defection will fall below that from cooperation. It will also be attainable in the first period (assuming cooperation in all the intermediate ones), because of anticipated future gains. As we show, in our setup this automatically implies the existence of a subgame perfect equilibrium in which players cooperate in all periods. We illustrate this result with an example concerning the voluntary provision of public goods.

In the second scenario we consider, the players' payoff function is fixed, but their information about it is incomplete. Specifically, players are not certain about the consequences of not cooperating with the other players in terms of the effect on their own payoff. However, as time progresses, they receive certain public information which gives them increasing confidence that, as long as the other players cooperate, it is best for them to do the same. In particular, this may occur if the duration of the game is stochastic and negatively correlated with the benefit from defection. As time progresses without the game terminating, the players' expected gain from defection decreases, and, consequently, their incentive to cooperate increases. The general result previously mentioned implies that, if the players are willing to cooperate in all the periods preceding and following some time interval during which the continuation probabilities are constant (or increasing), then they will automatically also be willing to cooperate in all the intermediate periods. We give a counterexample showing that if the continuation probabilities decrease over time, then cooperation may fail even if players are willing to cooperate in the first and last periods. This failure is caused by a single period in the middle in which defection is profitable.

The general setup and the two more specific scenarios considered here are constructed so as to differentiate our results from previously suggested solutions to the cooperation problem. First, in the setting in which information is incomplete, there is no private information. Players not only have identical payoff functions but they also hold identical beliefs about them. If the players' preferences differed and they had private information about them, their behavior might also have been directed by their desire not to let the other players know their preferences. For example, to elicit cooperation, egoists might want to behave (at least part of the time) as if they were altruists. The fact that informational asymmetries can generate cooperative behavior in finitely repeated games is well known (Kreps et al., 1982; Fudenberg and Maskin, 1986). Our assumptions are chosen to exclude such effects, and to keep our model within the framework of symmetric interactions involving symmetric information.

Second, cooperation in the finitely repeated prisoner's dilemma (and other games) may be possible if there is uncertainty about the duration of the game. Indeed, it suffices that the duration is not <u>commonly known</u> among the players. Neyman (1999) shows how to construct an uncertainty structure, representing a small departure from the common knowledge assumption on the number of repetitions (which, in particular, is almost certainly equal to some fixed integer T), and corresponding strategies for the two players such that, with very high probability, cooperation is attained in an arbitrarily high percentage of the periods. A necessary feature of such a construction is a positive probability (possibly very small) that the game goes on for an arbitrarily long time, i.e., an unbounded number of repetitions. In our model, this is not required. Indeed, even though we allow for uncertain duration of the game, this is not an essential feature of the model; our results also hold in the special case of a fixed, commonly known number of repetitions.

Finally, in games with multiple equilibria, it is not unusual to find that <u>any</u> feasible and individually rational payoff vector can be approximated by the average payoff vector in some pure-strategy subgame perfect equilibrium in the T-times repeated game, for large enough T. Indeed, as Benoit and Krishna (1985) show, a sufficient condition for this is that, for each player i, there are two pure-strategy equilibria in the one-shot game with different payoffs for i. (An additional condition is that the set of all feasible payoff vectors has a nonempty interior. However, if there are only two players, this condition can be dispensed with.) This "limit" folk theorem is not, however, applicable to games in which the equilibrium payoffs are unique. In fact, if the equilibrium payoffs coincide with the players' individual rationality (or minimax) levels, they are also the unique equilibrium payoffs in the repeated game, regardless of the number of repetitions (and whether or not subgame perfection is required). This is because, if the players' equilibrium payoffs are equal to their individual rationality levels, then, by definition, there is no way the other players can punish a defector by lowering his payoff. Therefore, the only self-enforcing strategy profiles in the repeated game are those inducing the equilibrium payoffs in all repetitions. In this paper, we do not assume that the one-shot game has more than one equilibrium. In fact, the first-period game may (but need not) be the prisoner's dilemma. If this were also the stage game in all the subsequent periods, then repeated defection would be the only equilibrium behavior. Thus, our assumption that experience can modify incentives is crucial if cooperation is ever to be attained.

2 The setup

A finite number n ($n \geq 2$) of players is engaged in a symmetric stochastic game Γ. The number of periods in Γ (which is always finite) is determined by the state of the world ω, which is a random element of some set Ω.[1] All the players share a common (finite or infinite) action set, which is the same in all states of the world and in all periods. In each period t (t = 1, 2, …), each player chooses one of these actions. The history in period t is the list h_t of all the players' actions in all the preceding periods. (Thus, $h_1 = \emptyset$.) We assume that, in each period t, all the players know the history h_t, but do not know anything about the state of the world. This assumption entails, in particular, complete information symmetry: players do not have any private information. We also assume symmetry in payoffs: each players' payoff in period t is given by the same payoff function $v_t(x_t, y_t, \ldots, z_t; h_t, \omega)$, whose arguments are the player's own action x_t in that period, the other players' actions $y_t, \ldots, z_t$ (listed in an arbitrary order), the history h_t, and the state of the world ω. Symmetry entails that the payoff function is invariant to permutations of its second to nth arguments; thus, the payoffs do not depend on the identity of any of the players. The payoff of each player in the stochastic game Γ is the sum of his payoffs in all the periods[2],

[1] The number of periods may be the same in all states of the world, in which case there is no uncertainty about it.

[2] Note that this does not preclude discounting: discount factors can always be embedded in the payoff functions.

$$v_1(x_1, y_1, \ldots, z_1; h_1, \omega) + v_2(x_2, y_2, \ldots, z_2; h_2, \omega) + \cdots . \quad (1)$$

(Even though the number of periods is finite, and depends on the state of the world ω, we may view (1) as an infinite sum by adopting the convention that $v_t(\,\cdot\,;\,\cdot\,, \omega) = 0$ for any t that exceeds the length of the game.) Since a player's payoff in each period t depends, among other things, on the (random) state of the world, it is itself a random variable. The expectation of this random variable, which is denoted by

$$w_t(x_t, y_t, \ldots, z_t; h_t),$$

is the player's expected payoff in period t. Obviously, this can be expressed as a function of the next-period history h_{t+1}. Therefore, the expected payoff of any player in Γ can be expressed as a function of the histories h_2, h_3, ….

A (pure) strategy in Γ is a mapping that assigns to each period $t \geq 1$ and corresponding history h_t a particular action x_t. A strategy profile is an assignment of a strategy to each of the n players in Γ. Such an assignment uniquely determines the history h_t in all periods t, and therefore each player's expected payoff in Γ. A strategy profile is an equilibrium in Γ if there is no player who can increase his expected payoff by changing his strategy, if the other players do no change theirs.

Two actions, which play a special role in our analysis, are assumed to exist. The first action, denoted c, is interpreted as cooperation with the rest of the players. The second, denoted d, is interpreted as defection (e.g., stepping out of the interaction). The history in which all the players cooperated in all the periods preceding t is denoted by h_t^c. The main problem with which this paper is concerned is finding sufficient conditions for the existence of a symmetric equilibrium such that, on the equilibrium path, all the players play c in all periods. If such an equilibrium exists, then we will say that cooperation is sustainable (in Γ). Since the results we seek are positive in nature, there is no need to consider the whole class of mappings from the set of all possible histories h_t to actions, which rapidly increases in size as t increases. Instead, we restrict attention to the simplest kind of strategies, namely, trigger strategies that prescribe cooperation until, but not after, the first period in which some players fail to cooperate. Thus, we look for conditions guaranteeing that the following is a symmetric equilibrium strategy in Γ:

(C) Play c until someone has deviated, and shift to d in all subsequent periods.

In the next two sections, we present and analyze two special cases of the general model presented above. We return to the general model in Section 5.

3 Altruism: Learning to care

One conceivable outcome of repeated cooperation among the players is that, in time, they become progressively less selfish. This might occur, for instance, if players get to know the other players personally and empathy develops among them. Thus, players put increasing weight on the other players' well-being, at the expense of their own. In keeping with the above setup, this will be assumed to represent a systematic shift in preferences, rather than a strategic choice made by individuals. The specific assumptions we make concerning the evolution of the players' preferences are spelled out below.

The utility function $u(x, y, \dots , z)$ gives each player's utility in each period as a function of the player's own action x in that period and the other players' actions y, $\dots$, z, listed in arbitrary order. The average utility, $\bar{u}(x, y, \dots , z)$, is given by $(1/n)\,[u(x, y, \dots , z) + u(y, x, \dots , z) + \cdots + u(z, y, \dots , x)]$. (Note that $\bar{u}(x, x, \dots , x) = u(x, x, \dots , x)$ for every x.) These functions are assumed to satisfy the following:

$$\max_x u(x, d, \dots , d) = u(d, d, \dots , d) \tag{2}$$

$$\max_x \bar{u}(x, c, \dots , c) = \bar{u}(c, c, \dots , c) \tag{3}$$

and

$$u(c, c, \dots , c) \ge u(d, d, \dots , d). \tag{4}$$

The first assumption states that when all the players defect, no single player can increase his utility by choosing an alternative action. The second assumption states that when all the players cooperate, no single player can increase the average utility by choosing an alternative action. The third assumption states that the utility (or, equivalently, the average utility) in the second case is at least as great as in the first. If cooperation is socially optimal, in the sense that the aggregate utility is maximal when everyone cooperates, then clearly both (3) and (4) hold.

In each period t, each player's payoff is a particular convex combination of his own and the average utility in that period. (Note that the latter also incorporates the player's own utility.) In the first period, the payoff coincides with the utility. Thus, at $t = 1$, players are completely selfish: they are only concerned with their own utility, and assign zero weight to the average utility. The same is also true for any period $t > 1$ which is not preceded by a perfect history of cooperation (i.e., $h_t \neq h_t^c$). In other words, the players revert to complete selfishness as soon as one or more of them fails to play c. On the other hand, if there is cooperation in each period, then the players become progressively less selfish over time. Specifically, each player's payoff in period t (t = 1, 2, …) is given, in this case, by a function $u_t(x, y, \dots , z)$ satisfying the recursive equation

$$u_{t+1} = s\, u_t + (1 - s)\, \bar{u}_t, \tag{5}$$

where $\bar{u}_t(x, y, \dots , z) = (1/n)\, [u_t(x, y, \dots , z) + u_t(y, x, \dots , z) + \cdots + u_t(z, y, \dots , x)]$ and $0 \leq s \leq 1$ is a fixed parameter. This exogenously given parameter, which is the same in all periods and for all players, is called the players' degree of selfishness. Since $u_1 = u$ by assumption, the extreme value $s = 1$ represents a case in which cooperation does not make players any less selfish. The other extreme value, $s = 0$, represents a case in which cooperation in all the previous periods makes players completely selfless: their payoffs equal the average utility. The recursive equation (5) is easy to solve. The explicit expression for the payoff, following a history of cooperation, is:

$$u_t = s^{t-1}\, u + (1 - s^{t-1})\, \bar{u}. \tag{6}$$

Following any other history, the payoff is given by u.

By assumption, the players' utilities and payoffs are not affected by the state of the world. However, the length of the game may be affected by it, i.e., it may be random. It is convenient to express the distribution of the number of periods in terms of the continuation probabilities. The continuation probability in period t, denoted by δ_t, is defined as the conditional probability that the number of periods is at least $t + 1$, given that it is not less than t.[3] If $\delta_t = 0$, then, as a matter of convention, we will set $\delta_{t+1} = \delta_{t+2} = \cdots = 0$. The (unconditional) probability that the number of periods is at

[3] Alternatively, δ_t may be interpreted as the players' (common) discount factor in period $t + 1$. This interpretation makes sense if δ_t is the same in all periods t, or constant until a certain period T and zero from the next period (at which the game ends) onwards.

least t is equal to $\delta_1\delta_2 \cdots \delta_{t-1}$. (For $t = 1$, this product is defined as 1.) If the number of periods is less than t, each player's payoff in period t is defined as zero. Therefore, the expected payoff in period t for a player planning to choose action x in that period, if the other players' actions are $y, \dots, z$ and the history is h_t, is given by

$$w_t(x, y, \dots, z; h_t) = \delta_1\delta_2 \cdots \delta_{t-1} [\varepsilon\, u(x, y, \dots, z) + (1 - \varepsilon)\, \bar{u}(x, y, \dots, z)], \tag{7}$$

where $\varepsilon = s^{t-1}$ if $t \geq 2$ and $h_t = h_t^c$, and $\varepsilon = 1$ otherwise.

If all the players use strategy C, as defined in the previous section, then, by (7), their expected payoff in Γ is equal to $(1 + \delta_1 + \delta_1\delta_2 + \cdots)\, u(c, c, \dots, c)$. If a single player deviates from C and, in some period t, chooses an action different from c, the other players respond by shifting to d from period $t + 1$ onwards. It then follows from (2) that the deviating player would maximize his payoff in these periods by also defecting, and getting the payoff $u(d, d, \dots, d)$ in each period. It follows, by (7), that a necessary and sufficient condition for C to be a symmetric equilibrium strategy in Γ is that, for all t with $\delta_{t-1} > 0$,

$$s^{t-1} (u(x, c, \dots, c) - \bar{u}(x, c, \dots, c)) \leq (\bar{u}(c, c, \dots, c) - \bar{u}(x, c, \dots, c))$$
$$+ (\delta_t + \delta_t\, \delta_{t+1} + \cdots)\, (u(c, c, \dots, c) - u(d, d, \dots, d)) \text{ for all } x \neq c. \tag{8}$$

For $t = 1$, the cooperation condition can be written more simply as

$$\max_{x \neq c} [u(x, c, \dots, c) - u(d, d, \dots, d)]$$
$$\leq (1 + \delta_1 + \delta_1\delta_2 + \cdots)\, (u(c, c, \dots, c) - u(d, d, \dots, d)). \tag{9}$$

Note that $1 + \delta_1 + \delta_1\delta_2 + \cdots$ is the expected number of periods in Γ.

If $0 < s < 1$, then the left-hand side in (8) tends to zero as t tends to infinity. Since, by (3) and (4), the right-hand side is nonnegative, it is at least plausible (although, strictly speaking, it does not follow from our assumptions) that the inequality in (8) holds for sufficiently large t. Similarly, (9) (which refers to $t = 1$) may be expected to hold if the expected number of periods in Γ is sufficiently large. This leaves the question of which conditions would imply that (8) also holds for all the intermediate periods. The following proposition shows that a sufficient condition for this is that the continuation probabilities weakly increase (e.g., are constant) during these periods. If they satisfy this condition and (8) holds in the initial period and from some later period T onwards, then it automatically also holds in all the intermediate periods.

PROPOSITION 1. *Suppose that there is some period T such that the cooperation condition* (8) *holds for* $t = 1$ *and for all* $t \geq T$. *Then, a sufficient condition for cooperation to be sustainable is that the continuation probabilities satisfy* $\delta_1 \leq \delta_2 \leq \cdots \leq \delta_{T-1}$.

The proof of Proposition 1, which uses the general lemma in Section 5, is deferred to that section.

A recent study by Brown et al. (2002) lends empirical support to the model presented in this section. These authors conducted an experimental study of a labor market in which the profit received by each producer depends on the amount of effort exerted by its worker. The market consisted of seven firms and 10 workers, and each firm could employ at most one worker. Each round of the experiment consisted of two stages. In the first stage, each of the producers offered a worker a wage, and specified the level of effort he desired him to exert. This offer could be made either to a specific worker or to the market as a whole. If a worker accepted the offer, the experiment entered the second stage, in which the worker chose his effort level. The greater the effort level the greater the firm's payoff, but the lower the worker's payoff. Fifteen rounds of the experiment were run with each subject group, with the subjects retaining the same roles throughout. In the treatment relevant to our theory, firms could identify workers by their ID numbers, and could choose to reemploy the same worker multiple times. In practice, workers who exerted "sufficient" effort tended to be rehired, while those who did not, found themselves unemployed in the next period. Thus, workers had an incentive to exert effort in each of the first 14 periods. However, in period 15, workers had no incentive to exert any effort, as there was no longer an opportunity for the firm to punish them. Thus, we expect the minimum effort level to be exerted in the final period. Surprisingly, this is not what happened. Rather, the effort levels chosen by workers were significantly and positively related to the duration of the relationship with the employer. The authors denote this as a "loyalty effect," stating that "the trust that arises from the repeated experience of being reemployed induces workers to become more loyal to their firms, i.e., they are more willing to take their firm's interests into account when choosing the effort level." This acquired loyalty is what our model attempts to capture. Note, however, that the experiment described, in contrast with our model, involves a highly asymmetric interaction. One kind of symmetric repeated interaction in which altruistic attitudes may be expected to arise is described in the following example.

Example: Voluntary provision of public goods

There is a fixed number T of periods, with $T \geq 2$. (Thus, $\delta_1 = \delta_2 = \cdots = \delta_{T-1} = 1$ and $\delta_T = 0$.) In each period, each of n players can contribute either zero or one unit of a private good as an input to the production of some public good; player i's contribution is denoted by x_i ($\in \{0, 1\}$). The quantity of the public good produced in each period is determined by the players' total contribution in that period, or, equivalently, by the average contribution $\bar{x} = (x_1 + x_2 + \cdots + x_n)/n$. It can therefore be expressed as $f(n\bar{x})$. The production function $f(x)$ is assumed to be differentiable and to satisfy $f(0) = 0$, $f(n) = 2$, and $1/n < df/dx \leq 1$. Each player's gain from each unit of the public good is assumed equal to his loss of utility from contributing one unit of the private good. Thus, the utility of each player i in each period is the difference between the quantity of public good produced in that period and i's contribution x_i:

$$u = f(n\bar{x}) - x_i.$$

Since the marginal product of the public good is assumed to be less than or equal to unity, a zero contribution is a dominant strategy in the one-shot game in which each player's payoff equals his utility. Therefore, the outcome $x_1 = x_2 = \cdots = 0$ is a symmetric equilibrium in that game, with equilibrium payoff zero. On the other hand, since $df/dx > 1/n$, the average utility, which is given by

$$\bar{u} = f(n\bar{x}) - \bar{x},$$

achieves its maximal value of one if and only if $x_1 = x_2 = \cdots = 1$ (i.e., everyone contributes). This shows that assumptions (2), (3), and (4) hold, with $c = 1$ and $d = 0$. We will now consider the effect of altruism, that is, a degree of selfishness $0 < s < 1$.

In the case under consideration, condition (9) takes the form $f(n - 1) \leq T$. Since $f(n - 1) < f(n) = 2$, this condition is satisfied. For $t = T$, condition (8) is equivalent to

$$s^{T-1} \leq \frac{(f(n) - f(n-1)) - \frac{1}{n}}{1 - \frac{1}{n}}. \tag{10}$$

Since $1/n < df/dx \leq 1$, the right-hand side of (10) is greater than 0 but less than or equal to 1. This inequality can therefore be interpreted as requiring the number of periods T to be sufficiently large or the degree of selfishness s to be sufficiently small (or both). In any case, by Proposition 1, (10) is both a necessary and sufficient condition for strategy C, whereby a player contributes to the production of the public good as long as everyone else also does so, to be a symmetric equilibrium strategy in Γ.

In the game Γ, a history of contribution to the production of the public good makes players increasingly more altruistic towards one another. Since 'altruism' in the present context means internalizing the benefits to the other players from the public good, the fact that it might result in players contributing to the production of the public good is not, in itself, very surprising. Our contribution here is in analyzing the case in which altruistic attitudes develop only gradually. Whether or not contributing to the production of the public good in all periods is an equilibrium behavior depends, in this case, on how fast attitudes change and how far the horizon is. Condition (10) is the exact expression of this dependence.

4 Incomplete information: Learning that you care

In the previous section, we showed how cooperation among the players may be maintained throughout, with altruism gradually replacing the prospect of future gains as the motivating power behind the players' willingness to cooperate. Thus, the players' cooperative behavior induces a systematic shift in their preferences, which, in turn, reinforces this behavior. In this section, we explore the possibility that people may learn their preferences, rather than acquire new ones, over the course of time. Specifically, participants in the game, who are initially uncertain about the consequences of unilaterally deviating from cooperation, may receive certain signals suggesting that such a deviation is unprofitable (or, conversely, profitable) to them. Players may, for example, simply learn that they like (or dislike) one another. If players expect many repetitions, and hence significant benefit from future cooperation, they may be willing to cooperate in the early periods, when they are still uncertain about the desirability of defection. The difficulty in maintaining cooperation throughout the game again lies in the middle periods. As the expected number of remaining periods is smaller in the middle than at the beginning of the game, cooperation requires that the players' expected gains from defection decline at a sufficiently fast rate to match the declining gains from cooperation through the end of the game. Thus, whether or not cooperation is sustainable may depend on the precise manner in which the players' beliefs about their payoff functions evolve over time.

There is obviously more than one way the above general scenario can be modeled. In particular, there is more than one possible mechanism whereby players may learn how deviation from cooperation would benefit or harm them. It is, however, noteworthy that players may be able to gain information about the consequences of a deviation without ever receiving any outside signals or cues. Specifically, if there is a correlation between the payoff function and the duration of the game, then the very fact that the game is not yet over may tell players something about their payoffs. Such a correlation may arise if, for example, incompatibilities among the players tend to increase both (i) the profitability of a deviation from cooperation and (ii) the probability of early termination of the interaction.

The model described below allows for such a correlation. It involves two kinds of states of the world: "good" states and "bad" states. In a good state of the word, each player's payoff in each period is given by a function $g(x, y, \dots, z)$ of the player's own action x in that period and the other players' actions $y, \dots, z$. In a bad state, the corresponding function is $b(x, y, \dots, z)$. These functions satisfy the following:

$$\max_x g(x, d, \dots, d) = g(d, d, \dots, d), \quad \max_x b(x, d, \dots, d) = b(d, d, \dots, d), \tag{11}$$

$$\max_x g(x, c, \dots, c) = g(c, c, \dots, c), \tag{12}$$

and

$$g(c, c, \dots, c) - g(d, d, \dots, d) \geq b(c, c, \dots, c) - b(d, d, \dots, d) \geq 0. \tag{13}$$

Thus, in any state of the world, it is to a player's advantage to defect when all the other players defect. In a good state, it is also to his advantage to cooperate when all the others cooperate. (This is one sense in which such a state is 'good'.) In any state, the players' payoffs are at least as high when they all cooperate as when they all defect, and the difference between the two is at least as high in a good as in a bad state. (This is another sense in which the former is 'good'.)

Suppose that all the players use strategy C, as defined in Section 2, and thus cooperate in all periods. Their payoffs in each period are then $g(c, c, \ldots , c)$ if the state of the world is good and $b(c, c, \ldots , c)$ if it is bad. If the payoffs are not equal, they may provide players with information about the state. However, we assume that players do not receive any such information. This assumption entails that either players do not know their own payoffs or, more plausibly, $g(c, c, \ldots , c) = b(c, c, \ldots , c)$. It does not, however, require the players' beliefs about the state of the world to be constant over time. For example, if the number of periods in the game tends to be greater in good than in bad states of the world, the posterior probability that a good state has been attained may increase over time. This may affect the players' assessment of the desirability of defection. A more detailed examination of this possibility follows.

Denote the (prior) probability that the state of the world is good by γ_0, and that it is bad by β_0 (= $1 - \gamma_0$). The continuation probabilities, conditional on the state being good, are denoted by $\gamma_1, \gamma_2, \ldots$, and conditional on being bad, by $\beta_1, \beta_2, \ldots$. In each period t, the (posterior) probability p_t that the state is good is given by

$$p_t = \frac{\gamma_0\gamma_1 \cdots \gamma_{t-1}}{\gamma_0\gamma_1 \cdots \gamma_{t-1} + \beta_0\beta_1 \cdots \beta_{t-1}} \tag{14}$$

(provided that the denominator is not zero; if it is zero, we arbitrarily set $p_t = 1$.) The probability of a bad state is $1 - p_t$. Therefore, conditional on having at least t periods, the expected payoff in period t is given by the function

$$u_t = p_t\, g + (1 - p_t)\, b. \tag{15}$$

The <u>unconditional</u> expected payoff in period t of a player who plans to choose action x in that period, if the other players' actions are $y, \ldots , z$, is given by

$$\begin{aligned} w_t(x, y, \ldots , z; h_t) &= \gamma_0\gamma_1 \cdots \gamma_{t-1}\, g(x, y, \ldots , z) + \beta_0\beta_1 \cdots \beta_{t-1}\, b(x, y, \ldots , z) \\ &= \delta_1\delta_2 \cdots \delta_{t-1}\, u_t(x, y, \ldots , z), \end{aligned} \tag{16}$$

where, for all $t \geq 1$,

$$\delta_t = p_t\, \gamma_t + (1 - p_t)\, \beta_t \tag{17}$$

is the continuation probability in period t, unconditional on the state. Note that the history h_t plays no role in these expressions.

If all the players use strategy C, their expected payoff in Γ equals $(\gamma_0 + \gamma_0\gamma_1 + \cdots)\, g(c, c, \ldots, c) + (\beta_0 + \beta_0\beta_1 + \cdots)\, b(c, c, \ldots, c)$. If a single player deviates from C and, in some period t, chooses an action different from c, the other players respond by shifting to d from period $t + 1$ onwards. It then follows from (11) that the deviating player would maximize his payoff in each of these periods by also choosing d. Therefore, a necessary and sufficient condition for C to be a symmetric equilibrium strategy in Γ is that, for all t with $\beta_{t-1} > 0$,

$$\begin{aligned} b(x, c, \ldots, c) - b(c, c, \ldots, c) \le \frac{\gamma_0}{\beta_0}\frac{\gamma_1}{\beta_1}\cdots\frac{\gamma_{t-1}}{\beta_{t-1}} [(g(c, c, \ldots, c) - g(x, c, \ldots, c)) \\ + (\gamma_t + \gamma_t\gamma_{t+1} + \cdots)(g(c, c, \ldots, c) - g(d, d, \ldots, d))] \\ + (\beta_t + \beta_t\beta_{t+1} + \cdots)(b(c, c, \ldots, c) - b(d, d, \ldots, d)) \text{ for all } x \ne c. \end{aligned} \tag{18}$$

For $t = 1$, the cooperation condition can be written more simply as

$$\begin{aligned} \max_{x \ne c} [\gamma_0 (g(x, c, \ldots, c) - g(d, d, \ldots, d)) + \beta_0 (b(x, c, \ldots, c) - b(d, d, \ldots, d))] \\ \le (\gamma_0 + \gamma_0\gamma_1 + \cdots)(g(c, c, \ldots, c) - g(d, d, \ldots, d)) \\ + (\beta_0 + \beta_0\beta_1 + \cdots)(b(c, c, \ldots, c) - b(d, d, \ldots, d)). \end{aligned} \tag{19}$$

Note that $(\gamma_0 + \gamma_0\gamma_1 + \cdots) + (\beta_0 + \beta_0\beta_1 + \cdots)$ is the expected number of periods in Γ.

If the continuation probabilities conditional on the state being good are greater than the corresponding probabilities in a bad state, then, for all $t \ge 1$, the ratio γ_t/β_t is greater than unity. Since, by (12) and (13), the right-hand side in (18) is nonnegative, it is then at least plausible (although, strictly speaking, it does not follow logically from our assumptions) that, in this case, the inequality in (18) holds for sufficiently large t. Similarly, (19) (which refers to $t = 1$) may be expected to hold if the expected number of periods in Γ is sufficiently large. The next question is which conditions would imply that (18) all holds for al the intermediate periods. The following proposition asserts that it is sufficient that the continuation probabilities, in both the good and bad states of the world, weakly increase during these periods, but at a weakly faster rate in the bad states. If the continuation probabilities satisfy this condition and (18) holds in the initial period and from some later period T onwards, then it automatically also holds in all the intermediate periods.

PROPOSITION 2. *Suppose that there is some period T such that the cooperation condition* (18) *holds for* $t = 1$ *and for all* $t \geq T$. *Then, a sufficient condition for cooperation to be sustainable is that* $\gamma_1 \leq \gamma_2 \leq \cdots \leq \gamma_{T-1}$ *and* $\gamma_1 - \beta_1 \geq \gamma_2 - \beta_2 \geq \cdots \geq \gamma_{T-1} - \beta_{T-1} \geq 0$.[4]

The proof of Proposition 2, which uses the Lemma given in Section 5, is deferred to that section.

A simple case in which the condition in Proposition 2 is satisfied is that of constant continuation probabilities. Suppose that, in all the periods preceding a certain period *T*, the continuation probabilities conditional on the state being good or bad do not change, and equal γ and β, respectively, with $\beta \leq \gamma \leq 1$. In period *T*, the continuation probabilities are zero, i.e., the game ends. It then follows from the proposition that if the cooperation condition holds in the initial period and in the last period *T*, then cooperation is sustainable.

If the continuation probabilities do not satisfy the condition in Proposition 2, then the two boundary conditions are not sufficient for cooperation to be sustainable. In other words, it is possible that cooperation in the first and last periods is profitable, assuming cooperation also occurs in all the intermediate periods, but nevertheless cooperation is not sustainable owing to one or more periods in the middle in which it is better to defect. In the following example, cooperation fails because of a single such period.

Counterexample: Decreasing continuation probabilities

The state of the world has equal probability of being good or bad. If it is good, then there are exactly 40 periods. If it is bad, the number of periods is determined by a random variable T_b that has a lognormal distribution with $\mu = 3$ and $\sigma = 0.3$. If $T_b \geq 40$, the number of periods is 40. Otherwise, it is the smallest integer greater than T_b. It is possible to show that the continuation probabilities do not satisfy the condition in Proposition 2: conditional on the state of the world being bad, they (i.e., the β_t's) decrease over time. It can also be shown that, if the state is bad, there is a probability greater than 0.9 that the number of periods is between 12 and 31, with the expected number about 21.5.

[4] This clearly implies $\beta_1 \leq \beta_2 \leq \cdots \leq \beta_{T-1}$.

In each period, the two players are engaged in a symmetric 2×2 incomplete-information game with payoff matrix

$$\begin{array}{cc} & \begin{array}{cc} c & d \end{array} \\ \begin{array}{c} c \\ d \end{array} & \begin{pmatrix} 1, 1 & -1, a \\ a, -1 & 0, 0 \end{pmatrix}, \end{array}$$

where $a = 0$ if the state of the world is good and $a = 38$ if it is bad. Thus, in a bad state, the game is the prisoner's dilemma, with the unique equilibrium (d, d). In a good state, (c, c) is also an equilibrium, and thus (12) holds.

In this example, cooperation is not sustainable. This can be seen by comparing the expected payoff from repeated cooperation in the continuation game starting in period t, which equals $1 + \delta_t + \delta_t\, \delta_{t+1} + \cdots$, with the expected payoff from unilateral defection in that period, which is given by $u_t(d, c)$ (see Figure 1). In the first and last periods, the former is greater than the latter. The same is also true in most of the other periods. In fact, there is only one period, namely, $t = 15$, in which defection is (marginally) better than continued cooperation. However, this has an unraveling effect on cooperation in all the preceding periods. Since a player cannot expect the other player to cooperate in any future period in which it is better for him to defect, the existence of such periods has a negative impact on his own incentive to cooperate. It can be shown, more specifically, that in any (pure-strategy) symmetric equilibrium both players cooperate, if at all, only after the 15th period.

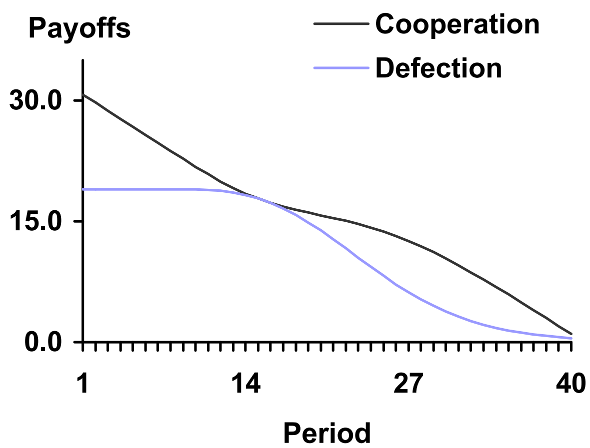

Figure 1. In the game with incomplete information described in the text, there is a single period ($t = 15$) in which each player's expected payoff from defection is greater than from continued cooperation until the last period.

5 General conditions for cooperation

In this section, we return to the general model introduced in Section 2, and derive a sufficient condition for cooperation to be sustainable in this model. We then use this result to prove the two special Propositions 1 and 2.

Our results depend on the following assumption concerning the effect of a history of cooperation on the expected payoffs:

(A) For every $t \geq 2$, there is a constant $L_t \geq 0$ such that, for every $h_t \neq h_t^{c}$,

$$\max_x w_t(x, d, \dots, d; h_t) = w_t(d, d, \dots, d; h_t) = w_t(c, c, \dots, c; h_t^c) - L_t.$$

In other words, in a period in which everyone defects and the history is <u>not</u> one of perfect cooperation, none of the players has an incentive to choose any alternative action. (We make no assumptions concerning the case of a perfect history of cooperation.) In addition, the expected payoff in such a period does not depend on the precise history, and is less than or equal to that corresponding to complete cooperation in the current and all the preceding periods. For the two special models in Sections 3 and 4, assumption A holds. For the former, it is implied by (2) and (4). For the latter, it follows from (11) and (13).

If all the players use strategy C, their expected payoff in each period t equals $w_t(c, c, \dots, c; h_t^c)$. If a single player deviates from C by choosing an action $x \neq c$ in a period t, that player's expected one-time gain (or, if negative, the loss) in that period equals

$$G_t(x) \stackrel{def}{=} w_t(x, c, \dots, c; h_t^c) - w_t(c, c, \dots, c; h_t^c).$$

Since all the other players use strategy C, they respond to the deviation by shifting to d from period $t + 1$ onwards. It then follows from assumption A that the best the deviating player can do in each period $t' > t$ is to also choose d, and get an expected payoff of $w_{t'}(c, c, \dots, c; h_{t'}^c) - L_{t'}$. Therefore, a necessary and sufficient condition for C to be a symmetric equilibrium strategy in Γ is that, for all $t \geq 1$

$$G_t(x) \leq L_{t+1} + L_{t+2} + \cdots \quad \text{for all } x \neq c. \tag{20}$$

The two cooperation conditions ((8) and (18)) obtained in the preceding two sections are special cases of this one. It follows from assumption A that, if the above condition holds, then the equilibrium is in fact perfect in the sense that the actions it prescribes are also optimal off-equilibrium, i.e., after one or more players have deviated.

As explained in the Introduction, our main concern is the following question: If there are two periods in which the inequality (20) holds, when does it automatically follow that it also holds in all the intermediate periods? The following result identifies a sufficient condition for this. The condition is that, for each intermediate period t in which a player would gain (in expectation) from choosing an action other than c, his gain does not exceed a certain weighted average of his gain from choosing the same action in the periods immediately preceding and following t. (Recall that, for a real number y, $y^+ = \max\{y, 0\}$.)

Lemma. *For a given action x, let $T_1 \geq 1$ and $T_2 \geq T_1 + 2$ be such that the inequality in* (20) *holds for $t = T_1$ and for $t = T_2$. For this inequality to hold also for all $T_1 < t < T_2$, it suffices that, for each such t, there is some $\delta_t \geq 0$ with $\delta_t L_t \leq L_{t+1}$ such that*

$$G_t(x) \leq \left[\frac{G_{t+1}(x) + \delta_t G_{t-1}(x)}{1 + \delta_t}\right]^+. \tag{21}$$

Proof. There exists some $T_1 < t < T_2$ for which the inequality $G_t(x) \leq L_{t+1} + L_{t+2} + \cdots$ does not hold if and only if $\max_{T_1 < t < T_2} (G_t(x) - L_{t+1} - L_{t+2} - \cdots)$ is greater than zero. Suppose this is the case, and consider the last $T_1 < t < T_2$ at which this maximum is attained. Clearly, $G_{t+1}(x) - L_{t+2} - L_{t+3} - \cdots < G_t(x) - L_{t+1} - L_{t+2} - \cdots$ and $G_{t-1}(x) - L_t - L_{t+1} - \cdots \leq G_t(x) - L_{t+1} - L_{t+2} - \cdots$. Therefore, $G_t(x) > L_{t+1} + L_{t+2} + \cdots \geq 0$ and, for any $\delta_t \geq 0$ such that $\delta_t L_t \leq L_{t+1}$,

$$(1 + \delta_t)\, G_t(x) - (G_{t+1}(x) + \delta_t\, G_{t-1}(x))$$

$$= (G_t(x) - G_{t+1}(x)) + \delta_t\, (G_t(x) - G_{t-1}(x)) > L_{t+1} - \delta_t L_t \geq 0,$$

which shows that (21) does not hold, and thus proves the sufficiency of the condition in the lemma. ■

Using the Lemma, we can now prove Propositions 1 and 2.

Proof of Proposition 1. Assume that the continuation probabilities satisfy the condition in the proposition. By (7), $L_t = \delta_1\delta_2 \cdots \delta_{t-1}\, (u(c, c, \ldots, c) - u(d, d, \ldots, d))$ for all $t \geq 2$, and hence $\delta_t L_t = L_{t+1}$. It therefore follows from the Lemma that to prove the assertion of the proposition, it suffices to show that (21) holds for all x and $1 < t < T$. For the rest of the proof, we fix some such pair of action and period.

Using (6) and (7), it is not difficult to check that the following functional identify holds: $w_{t+1}(\,\cdot\,;h^{c}_{t+1}) + \delta_t\, w_{t-1}(\,\cdot\,;h^{c}_{t-1}) - (1+\delta_t)\, w_t(\,\cdot\,;h^{c}_{t}) = \delta_1\delta_2\cdots\delta_{t-2}\,[(\delta_t - s\,\delta_{t-1}\,\delta_t)\, u_{t-1}(\,\cdot\,) - (\delta_{t-1} - s\,\delta_{t-1}\,\delta_t)\, u_t(\,\cdot\,)]$. (If $t = 2$, $\delta_1\delta_2\cdots\delta_{t-2} = 1$ by definition.) Since $u_{t-1}(c, c, \ldots, c) = u_t(c, c, \ldots, c) = u(c, c, \ldots, c)$, this implies

$$G_{t+1}(x) + \delta_t\, G_{t-1}(x) - (1+\delta_t)\, G_t(x) = \delta_1\delta_2\cdots\delta_{t-2}\,[(\delta_t - s\,\delta_{t-1}\,\delta_t)\,(u_{t-1}(x, c, \ldots, c) - u(c, c, \ldots, c)) - (\delta_{t-1} - s\,\delta_{t-1}\,\delta_t)\,(u_t(x, c, \ldots, c) - u(c, c, \ldots, c))]. \quad (22)$$

If $u_{t-1}(x, c, \ldots, c) \geq u_t(x, c, \ldots, c) > u(c, c, \ldots, c)$, then it follows from the assumption $\delta_{t-1} \leq \delta_t$ (≤ 1) that the right-hand side of (22) is nonnegative, and therefore (21) holds. If $u_t(x, c, \ldots, c) \leq u(c, c, \ldots, c)$, then $G_t(x) \leq 0$, and, therefore, the same conclusion holds. Finally, if $u_{t-1}(x, c, \ldots, c) < u_t(x, c, \ldots, c)$, then it follows from (6) and (3) that $u_t(x, c, \ldots, c) \leq \bar{u}(x, c, \ldots, c) \leq u(c, c, \ldots, c)$. As just shown, this implies (21). ■

Proof of Proposition 2. Assume that the continuation probabilities satisfy the condition in the proposition. By (14), the assumption that $\gamma_t \geq \beta_t$ for $t = 1, 2, \ldots, T-1$ implies that

$$p_1 \leq p_2 \leq \cdots \leq p_T. \quad (23)$$

By (15) and (16), $L_t = \delta_1\delta_2\cdots\delta_{t-1}\,[p_t\,(g(c, c, \ldots, c) - g(d, d, \ldots, d)) + (1-p_t)\,(b(c, c, \ldots, c) - b(d, d, \ldots, d))]$ for all $t \geq 2$. Hence, by (13) and (23), $\delta_t\, L_t \leq L_{t+1}$ for $t = 2, 3, \ldots, T-1$. It therefore follows from the Lemma that, to prove the assertion of the proposition, it suffices to show that (21) holds for all x and $1 < t < T$. For the rest of the proof, we fix some such pair of action and period.

It follows from (14) and (17) that $\gamma_t\, p_t\,(1-p_{t+1}) = p_{t+1}\,(1-p_t)\,\beta_t = p_{t+1}\,(\delta_t - p_t\,\gamma_t)$, and therefore $\gamma_t\, p_t = \delta_t\, p_{t+1}$. This and (15) give

$$\gamma_t\, u_t - \delta_t\, u_{t+1} = (\gamma_t - \delta_t)\, b. \quad (24)$$

Using (16), (24), and a similar equality in which the index t is replaced by $t-1$, it is not difficult to check that the following functional identify holds: $w_{t+1}(\,\cdot\,;h^{c}_{t+1}) + \delta_t\, w_{t-1}(\,\cdot\,;h^{c}_{t-1}) - (1+\delta_t)\, w_t(\,\cdot\,;h^{c}_{t}) = \delta_1\delta_2\cdots\delta_{t-2}\,[\delta_t\,(1-\gamma_{t-1})\, u_{t-1}(\,\cdot\,) - \delta_{t-1}\,(1-\gamma_t)\, u_t(\,\cdot\,) + (\delta_t\,\gamma_{t-1} - \delta_{t-1}\,\gamma_t)\, b(\,\cdot\,)]$. (If $t = 2$, $\delta_1\delta_2\cdots\delta_{t-2} = 1$ by definition.) This implies

$$G_{t+1}(x) + \delta_t\, G_{t-1}(x) - (1+\delta_t)\, G_t(x) = \delta_1\delta_2 \cdots \delta_{t-2}\, [\delta_t\, (1-\gamma_{t-1})\, (u_{t-1}(x, c, \dots, c)$$

$$- u_{t-1}(c, c, \dots, c)) - \delta_{t-1}\, (1-\gamma_t)\, (u_t(x, c, \dots, c) - u_t(c, c, \dots, c)) \qquad (25)$$

$$+ (\delta_t\, \gamma_{t-1} - \delta_{t-1}\, \gamma_t)\, (b(x, c, \dots, c) - b(c, c, \dots, c))].$$

If $u_t(x, c, \dots, c) - u_t(c, c, \dots, c) \le 0$, then $G_t(x) \le 0$, and (21) holds. If $u_t(x, c, \dots, c) - u_t(c, c, \dots, c) > 0$, then, by (15) and (12),

$$(1 - p_t)\, (b(x, c, \dots, c) - b(c, c, \dots, c)) > p_t\, (g(c, c, \dots, c) - g(x, c, \dots, c)) \ge 0 \qquad (26)$$

and, therefore, it follows from (23) that

$$u_{t-1}(x, c, \dots, c) - u_{t-1}(c, c, \dots, c) = (1 - p_{t-1})\, (b(x, c, \dots, c) - b(c, c, \dots, c))$$

$$- p_{t-1}\, (g(c, c, \dots, c) - g(x, c, \dots, c)) \ge u_t(x, c, \dots, c) - u_t(c, c, \dots, c). \qquad (27)$$

We claim that (21) holds in this case, too. To prove this claim, it suffices to show that the right-hand side of (25) is nonnegative. Hence, by (26) and (27), it suffices to show that $\delta_t\, (1 - \gamma_{t-1}) \ge \delta_{t-1}\, (1 - \gamma_t)$ and $\delta_t\, \gamma_{t-1} - \delta_{t-1}\, \gamma_t \ge 0$. The assumption $\gamma_{t-1} - \beta_{t-1} \ge \gamma_t - \beta_t \ge 0$ and (23) together imply $(1 - p_{t-1})\, (\gamma_{t-1} - \beta_{t-1}) \ge (1 - p_t)\, (\gamma_t - \beta_t) \ge 0$. By (17), this is equivalent to $\gamma_{t-1} - \delta_{t-1} \ge \gamma_t - \delta_t \ge 0$, and hence implies that $\delta_t - \delta_{t-1} \ge \gamma_t - \gamma_{t-1} \ge 0$ (where the last equality holds by assumption). Therefore, $\delta_t\, (1 - \gamma_{t-1}) - \delta_{t-1}\, (1 - \gamma_t) = (\delta_t - \delta_{t-1})\, (1 - \gamma_t) + \delta_t\, (\gamma_t - \gamma_{t-1}) \ge 0$ and $\delta_t\, \gamma_{t-1} - \delta_{t-1}\, \gamma_t = (\delta_t - \delta_{t-1})\, (\gamma_t - \delta_t) + \delta_t\, [(\gamma_{t-1} - \delta_{t-1}) - (\gamma_t - \delta_t)] \ge 0$, as had to be shown. ■

References

1. BENOIT, J. P., AND KRISHNA, V. (1985). "Finitely repeated games," *Econometrica* **53**, 905–922.

2. BROWN, M., FALK, A. AND FEHR, E. (2002). "Contractual incompleteness and the nature of market interactions," Institute for Empirical Research in Economics, University of Zurich, Working Paper no. 38.

3. FUDENBERG, D., AND MASKIN, E. (1986). "The folk theorem in repeated games with discounting or with incomplete information," *Econometrica* **54**, 533–554.

4. KREPS, D. M., MILGROM, P., ROBERTS, J., AND WILSON, R. (1982). "Rational cooperation in the finitely repeated prisoners' dilemma," *Journal of Economic Theory* **27**, 245–252.

5. NEYMAN, A. (1999). "Cooperation in repeated games when the number of stages is not commonly known," *Econometrica* **67**, 45–64.