Acquired Cooperation in Finite-Horizon Dynamic Games*

IGAL MILCHTAICH[a]

and

AVI WEISS[b]

Bar-Ilan University

April 2002

Abstract

No matter how many times a prisoner's-dilemma-like game is repeated, 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 the late stages of the repeated interaction, when unilateral defection has become unprofitable. In this case, cooperation may be attainable also in the early stages, since any defection in these stages may be effectively punished by all the other players also ceasing to cooperate. 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 in the late periods and, at the same time, punishments are less effective than at the beginning. These conditions are sufficient for cooperation in all periods to be an equilibrium outcome.

---

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

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

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

1 Introduction

There are numerous situations in which people engage in recurring identical symmetric interactions over a limited period of time. Parents of school-aged children are requested to partake in activities that only partially benefit their own offspring. People in residential areas know they are affected by, and in turn affect, their neighbors. College students who have never met are put into the same dormitory room, and must find a way to make it "work." Many of these interactions are carried out in a prisoner's-dilemma-type setting, in which the cooperative outcome is not individually optimal in any single period. In such a setting, if the number of interactions is finite and commonly known, it is well known that cooperation is unsustainable in equilibrium.

Nevertheless, in many such settings cooperation is attained. Thus, parents take turns driving children to school in car pools. They also participate in parents' meetings, and donate time and money to school activities. People cooperate by refraining from activities that annoy their neighbors, even when planning to leave the neighborhood. Roommates find ways to allocate chores and space and to coordinate times even if they will be roommates only for a year.

The explanation we explore in this paper for cooperation in such circumstances is that the players' preferences over the possible outcomes in each period are not constant over time but may be modified by past experience. Specifically, successful cooperation in early stages of a repeated interaction may change each player's perception of the payoff structure of the game in such a manner that defecting from cooperation becomes increasingly less attractive as time progresses (despite the fact that the physical form of the interaction does not change). If this is so, and if cooperation is maintained for a sufficiently large number of periods, then, after a certain point, it becomes an equilibrium behavior in the stage game. And if these changes are foreseen by all players, and are taken into account when deciding which actions to take in each period, then cooperation may also be maintainable in the early stages of the game, since the anticipated benefit from future cooperation is sufficient to prevent players from defecting in the early stages.

Such a benefit, of course, is dependent on the ability to maintain cooperation in all periods, and, in particular, the middle ones. However, because of the nearer horizon, the promise of benefiting from future cooperation is weaker in the middle periods than in the early ones. At the same time, the payoff structure may not yet have changed enough to make cooperation an equilibrium behavior in the stage game. Thus, even if (1) cooperation is an equilibrium behavior in each of the late periods, assuming a perfect history of cooperation, and (2) the one-time benefit from defection in any of the early periods is less than that from cooperation in all the subsequent periods, cooperation may not be attainable because there are one or more periods in the middle in which it is better to defect, and, therefore, players cannot assume that the other players will cooperate in these periods. In such a case, players may have no incentive to cooperate with the others even in the early periods, and, therefore, the history of cooperation needed for making it an equilibrium behavior in the late periods will not materialize. Consequently, there will be no cooperation throughout much or all of the game.

After presenting the basic setup, we concern ourselves with conditions sufficient to guarantee the players' willingness to cooperate in the middle stages of a repeated interaction, when (1) and (2) in the previous paragraph hold (i.e., when it is not profitable to defect from cooperation in either the early or the late stages of the game, assuming that cooperation will be maintained throughout). We observe that one such condition is that, when only one player deviates by ceasing to cooperate at a certain period t, that player's expected payoff is either monotonic or U-shaped (i.e., first decreasing, then increasing) function of t. It is not difficult to see that, if this condition holds, then the assumption that players are better off cooperating with the others than defecting in the early and in the late stages of the interaction automatically implies that it is an equilibrium behavior for all players to cooperate in all periods.

We consider two settings in which the condition just mentioned is satisfied. In the first, cooperation makes people increasingly more altruistic towards one another. Specifically, as the history of cooperation between them increases, players put increasing weight on the utility of the other players at the expense of their own. Thus, if cooperation is maintained in a certain period, each player's payoff function in the next period is a weighted average of his and the others' current payoff functions. As a result, the players progressively internalize the social costs or benefits of their actions. Assuming that cooperation is socially beneficial, and the interaction is sufficiently long, 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 periods), because of the anticipated future gains. As we show, 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 setting we consider, the players' payoff function is fixed, but their information about it is incomplete. Specifically, players are not certain about the consequences—in terms of the effects on their own payoffs—of not cooperating with the other players. However, as time progresses, they receive certain information that makes them believe with increasing confidence that, as long as the other players cooperate, it is best for them to do the same. In this setting, if the number of repetitions is large enough, then from some point on players will prefer cooperating with the others to defecting even in the one-shot game. In addition, the one-time benefit from defection in any of the early periods is outweighed by the benefit from cooperation in all subsequent periods. However, whether or not players will actually want to cooperate in all periods depends on the way their beliefs about the profitability of deviation from cooperation evolve over time.

A simple, concrete scenario in which the players' beliefs may change in the way outlined above is when the duration of the game is stochastic and negatively correlated with the benefit from defection. Specifically, in each period, the probability that the game will continue for at least one more period is higher in the "good" state of the world, in which defection is not profitable, than in the "bad" state, in which it is profitable. As time progresses without the game terminating, the posterior probability that the good state has obtained increases, as does the players' incentive to cooperate. We show that if the continuation probabilities in the two states of the world are (for instance) constant over some time interval, and if the players are willing to cooperate in the first period and from some late period on, then they are automatically also willing to cooperate in all the intermediate periods, and so cooperation in all periods is attainable. We also give a counterexample showing that if the continuation probabilities in the bad state are not constant but decrease over time, then cooperation may fail because there is a single period in the middle in which defection is profitable.

The general setup and the two more specific settings 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 similar payoff functions but they also hold identical beliefs about them. If players had private information about their preferences, their behavior could have been directed also by their desire not to let the other players know these preferences. For example, to elicit cooperation, egoists may 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). Our assumptions are chosen to exclude such effects, and to keep our model within the framework of symmetric interactions involving symmetric information.

Second, in games with multiple equilibria, it is not unusual to find that any feasible and individually rational payoff vector can be approximated by the average payoff vector in some subgame perfect equilibrium of the T-times repeated game, provided that T is large enough. 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 interior of the set of all feasible payoff vectors is not empty. 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 desired). The reason for this is that, if the players' 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 each of the 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 for cooperation ever to be attained.

2 The setup

2.1 The one-shot game

A finite number n of rational players ($n \geq 2$) is engaged in a symmetric interaction, whose outcome depends on the action each of them takes. Without loss of generality, the action profile itself may be referred to as the outcome of the interaction. Each player's preferences are expressed by a payoff function, mapping outcomes to utilities. The assumption that the interaction is symmetric entails that the payoff $h(x, y, \ldots, z)$ of a player taking a particular action x, while generally depending on the actions $y, \ldots, z$ of the other players, does not depend on who of the others is doing what. In other words, the payoff function h is invariant to permutations of its second to nth arguments. In addition, the same function h also gives the payoff of each of the other players, when that player's action is put as the first argument. The payoff function h and the players' common set of admissible actions (which may be finite or infinite) together define an n-player symmetric game G. A given outcome is a (pure-strategy Nash) equilibrium in G if no player can increase his payoff by changing his action, when the other players do no change theirs.

Two actions play a special role in our analysis. The first, denoted d, is interpreted as not cooperating with the rest of the players, e.g., stepping out of the interaction. If all the players take that action, their payoffs are zero; i.e., $h(d, d, \ldots, d) = 0$. In this case, a single player cannot gain by taking a different action; i.e., $h(x, d, \ldots, d) \leq 0$ for all actions x. We make no assumptions about the players' payoffs if two or more of them choose actions different from d. The other action playing a special role in the analysis is denoted c. Everyone choosing this action is interpreted as cooperation among the players. The corresponding payoff is assumed to be greater than zero. Without loss of generality, we normalize the payoff function by setting $h(c, c, \ldots, c) = 1$. Thus, cooperation is desirable. However, it is not necessarily an equilibrium behavior. That is, there may be some action x with $h(x, c, \ldots, c) > 1$. In this case, cooperation cannot be attained as the play of rational players in G. By a well-known backward induction argument, it may also be unattainable when the game is repeated any finite, commonly known, number of times. More precisely, cooperation cannot be attained if everyone playing d is the only equilibrium in G. The reason is that, as the end of the repeated game approaches, the lure of an immediate profit from choosing an action x more profitable than c becomes stronger than the promise of future gains from cooperation. The main problem this paper is concerned with is as follows: Under what conditions is cooperation sustainable in a finite dynamic game in which repeated cooperation among the players affects the stage game by making defection less attractive? One possible interpretation of such an effect of past cooperation on players' payoff functions is that defection carries a moral cost, which is greater the longer the history of cooperation. Other interpretations are discussed in subsequent sections.

2.2 The dynamic game

The dynamic (more specifically, stochastic) game Γ that we consider has a finite number of periods. This number may be fixed, and known to all players; or it may be random, in which case only the probability distribution of the number of periods is known. The expected number of periods is denoted e_0. In each period t (t = 1, 2, …), the expected number of remaining periods, denoted e_t, is given by

$$e_t = d_t + d_t d_{t+1} + d_t d_{t+1} d_{t+2} + \ldots, \qquad (1)$$

where d_t is the continuation probability in period t, i.e., the conditional probability that the number of periods is greater than t, given that it is no less than t.[1] Equation (1) also holds for $t = 0$, with d_0 denoting the probability that there is at least one period. (There is clearly little loss of generality in assuming $d_0 = 1$.) The expected number of remaining periods at time t satisfies the recursive formula

$$e_t = d_t\,(1 + e_{t+1}) \qquad (2)$$

($t = 0, 1, \ldots$). If $d_t = 0$, then, as a matter of convention, we set $d_{t+1} = d_{t+2} = \ldots = 0$ and $e_{t+1} = e_{t+2} = \ldots = 0$.

In each period t, the players are engaged in a symmetric stage game, with payoff function h_t. The payoff function is not necessarily constant over time. Specifically, for $t = 2, 3, \ldots$, the period t payoff function is determined by the players' actions in all the preceding periods. These actions are observed by all players. Hence, everyone always knows the payoff function for the next period (but not necessarily the actual payoffs, which also depend on the players' actions in that period). The payoff function after a

[1] Alternatively, the d_t's can be interpreted as (common) discount factors. The discount factor is either the same for all periods t, or it is constant until a certain period T, at which point it drops to zero (i.e., the game ends).

perfect history of cooperation, i.e., when all players chose the action c in all the periods preceding t, is denoted by h_t^c (with $h_1^c = h_1$). The following properties of the payoff function are assumed:

(a) As long as everyone cooperates, the utility from cooperation does not change, i.e., $h_t^c(c, c, \ldots, c) = 1$ for all t.[2]

(b) If period t is not preceded by a perfect history of cooperation, i.e., if in one or more of the periods preceding t some player's action is different from c, then $\max_x h_t(x, d, \ldots, d) = h_t(d, d, \ldots, d) = 0$.

A (pure) strategy in the dynamic game Γ specifies the action to be taken in each period as a function of the players' actions in all the preceding periods. The players' strategy profile unambiguously determines each player's expected payoff in Γ. This payoff can be written as $d_0 h_1 + d_0 d_1 h_2 + d_0 d_1 d_2 h_3 + \ldots$, where, for each period t, the arguments in h_t are the players' actions in that period (as specified by their strategies), with the player's own action listed first. The coefficient $d_0 d_1 \cdots d_{t-1}$ is the probability that the game will have at least t periods. The problem described in the previous subsection can now be given a more precise formulation, namely: Finding conditions on the probability distribution of the duration of the game (equivalently, the continuation probabilities) and on how the payoff function evolves guaranteeing the existence of some subgame perfect 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 the dynamic game Γ).

2.3 Conditions for cooperation

Consider the following strategy in the dynamic game Γ described above:

(C) Play c until someone has deviated, and shift to d from that point on.

If all the players use the strategy C, each player's expected payoff in Γ is equal to e_0 ($= d_0 + d_0 d_1 + \ldots$). For C to be a symmetric equilibrium strategy it is necessary and sufficient that the expected payoff of a player who deviates from C does not exceed e_0.

---

[2] It is, of course, conceivable that the utility from cooperation does change over time. However, to simplify our analysis, we assume that it does not.

For a player who deviates by cooperating until a certain period t but then taking a different action x in that period, the maximum expected payoff, denoted by $M(x, t)$, is given by

$$M(x, t) = d_0 + d_0 d_1 + \ldots + d_0 d_1 \cdots d_{t-2} + d_0 d_1 \cdots d_{t-1}\, h_t^{c}(x, c, \ldots, c). \tag{3}$$

(By assumption (b) above, the player's maximum payoff in periods $t+1$, $t+2$, … is zero, since his deviation is punished by all the other players shifting to d.) Thus, a necessary and sufficient condition for C to be a symmetric equilibrium strategy is that, for all $t \geq 1$,

$$\max_x M(x, t) \leq e_0. \tag{4}$$

By (1) and (3), a condition equivalent to (4) is that $d_{t-1} = 0$ or

$$\max_x h_t^{c}(x, c, \ldots, c) \leq 1 + e_t. \tag{5}$$

If one of the above equivalent conditions holds for all t, then it follows from assumption (b) above that everyone employing strategy C is in fact a subgame perfect symmetric equilibrium in Γ. Thus, the players' assertion of their commitment to play according to C is credible.

3 Altruism: Learning to care

One conceivable outcome of repeated cooperation among the players is that, as time goes by, they become progressively less selfish, for instance, because they get to know the other players personally and develop empathy for 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. In addition, to stay within the framework of symmetric interactions, the players' degree of selfishness, which is the weight s they put on their own utility, is assumed to be the same for all players. Each player's payoff is the convex combination $s\, h + (1 - s)\, \bar{h}$ of his own utility, which is given by a payoff function h satisfying the conditions in Section 2.1, and the average utility $\bar{h}(x, y, \ldots, z) = (1/n)\,[h(x, y, \ldots, z) + h(y, x, \ldots, z) + \ldots + h(z, y, \ldots, x)]$. (Note that the latter also incorporates the player's own utility.) The extreme case $s = 1$ represents complete selfishness: the players' payoffs are equal to their own utilities. The other extreme case, $s = 0$, represents complete unselfishness: all payoffs are given by the average utility.

In the first period, $t = 1$, the players' payoff function h_1 is equal to h. If, in this period, the cooperative outcome obtains, the next period payoff function is given by $h_2 = h_2^c = s\ h + (1 - s)\ \bar{h}$, where $0 < s < 1$ is the fixed, exogenously given degree of selfishness. If a different outcome obtains, the payoff function does not change, and remains the original one, h, in all periods. (Note that, in both cases, the average period 2 payoff function $\bar{h}_2$ is equal to $\bar{h}$.) If the cooperative outcome again obtains in the second period, consistency demands that the next period payoff function be given by $h_3 = h_3^c = s\ h_2^c + (1 - s)\ \bar{h}_2$. A quick calculation shows this to be equal to $s^2\ h + (1 - s^2)\ \bar{h}$. Repeated application of this consistency principle yields the following rule: If cooperation was obtained in all the preceding periods, the period $t + 1$ payoff function is given by

$$\begin{aligned} h_{t+1} = h_{t+1}^c &= s\ h_t^c + (1 - s)\ \bar{h} \\ &= s^t\ h + (1 - s^t)\ \bar{h}. \end{aligned} \tag{6}$$

Otherwise, $h_{t+1} = h$. This rule clearly satisfies assumptions (a) and (b) above.

Suppose that cooperation maximizes social welfare, in the sense that the average utility $\bar{h}$ is at its maximum when everyone cooperates. Since it follows from (6) that, as t tends to infinity, h_t^c tends to $\bar{h}$, it is at least plausible that, if cooperation persists for a sufficiently long time, then, from some period T and onwards, it becomes an equilibrium behavior in the stage game. Suppose, in addition, that the expected number of periods is sufficiently large so that the cooperation condition (4) holds in the first period. Under which circumstances is it possible to conclude that *C*, the strategy of cooperating as long as everyone else does so, is a symmetric equilibrium strategy in the dynamic game Γ described above? More generally, if there are two periods in which (4) holds, when does it automatically follow that it also holds in all the intermediate periods? The following assertion is obvious.

OBSERVATION. *For a given action x, suppose that there is some* $T_1 \geq 1$ *and some* $T_2 \geq T_1 + 2$ *such that* $M(x, T_1) \leq e_0$ *and* $M(x, T_2) \leq e_0$*. A sufficient condition for the inequality* $M(x, t) \leq e_0$ *to hold for all* $T_1 \leq t \leq T_2$ *is that* $M(x, t)$ *is either monotonic or "U-shaped" as a function of t in this time interval, i.e., for all* $T_1 < t < T_2$,

$$\textit{if } M(x, t) - M(x, t - 1) \geq 0, \textit{ then } M(x, t + 1) - M(x, t) \geq 0.$$

In particular, if we know that it is not profitable for individual players to deviate from C in the first period, as well as in any of the late periods, then a sufficient condition for such a deviation not be to profitable in <u>any</u> period is that, in the intermediate periods, the defecting individual's maximum payoff is either monotonic or U-shaped as a function of the time of defection. As the proof of the following proposition shows, a sufficient condition for <u>this</u> is that, in the intermediate periods, the continuation probabilities are weakly increasing.

PROPOSITION 1. *Suppose that the payoff function h has the property that a* <u>*single*</u> *player cannot increase social welfare by not cooperating with the others*; *i.e.*,

$$\bar{h}(x, c, \ldots, c) \leq 1 \text{ for all actions } x. \tag{7}$$

Suppose also that there is some period T such that the cooperation condition (4) *holds for* $t = 1$ *and for all* $t \geq T$.[3] *Then, a sufficient condition for cooperation to be sustainable is that* $\boldsymbol{d}_1 \leq \boldsymbol{d}_2 \leq \cdots \leq \boldsymbol{d}_{T-1}$.

Proof. We will show that, if the cooperation condition (4) does <u>not</u> hold for some $1 < t < T$, then there is at least one such t for which $\boldsymbol{d}_{t-1} > \boldsymbol{d}_t$. Suppose there is an action x such that, for some $1 < t < T$, $M(x, t) > e_0$. Then, there are some T_1 and T_2, with $1 \leq T_1 \leq T_1 + 2 \leq T_2 \leq T$, such that $M(x, T_1) \leq e_0$ and $M(x, T_2) \leq e_0$ but $M(x, t) > e_0$ for all $T_1 < t < T_2$. It follows, by the Observation, that there is some $T_1 < t < T_2$ such that $M(x, t) \geq M(x, t-1)$ but $M(x, t+1) < M(x, t)$. Since (by (3) and (6))

$$M(x, t+1) - M(x, t) = \boldsymbol{d}_0\boldsymbol{d}_1 \cdots \boldsymbol{d}_{t-1} \left(1 + \boldsymbol{d}_t\, h^{c}_{t+1}(x, c, \ldots, c) - h^{c}_{t}(x, c, \ldots, c)\right)$$

$$= \boldsymbol{d}_0\boldsymbol{d}_1 \cdots \boldsymbol{d}_{t-1} \left[1 - (1/s - \boldsymbol{d}_t)\, h^{c}_{t+1}(x, c, \ldots, c) + (1/s - 1)\, \bar{h}(x, c, \ldots, c)\right],$$

this implies that the expression in square brackets is negative, but a similar expression in which the index t is replaced by $t - 1$ is either positive or zero. Therefore, $(1/s - \boldsymbol{d}_t)\, h^{c}_{t+1}(x, c, \ldots, c) \geq (1/s - \boldsymbol{d}_{t-1})\, h^{c}_{t}(x, c, \ldots, c)$. The assumption $M(x, t) > e_0$ and (7) imply that $h^{c}_{t}(x, c, \ldots, c) > 1 \geq \bar{h}(x, c, \ldots, c)$. Therefore, by (6), $h^{c}_{t}(x, c, \ldots, c) > \max\{h^{c}_{t+1}(x, c, \ldots, c), 0\}$. It follows that $1/s - \boldsymbol{d}_t > 1/s - \boldsymbol{d}_{t-1}$, or equivalently $\boldsymbol{d}_{t-1} > \boldsymbol{d}_t$, as we wanted to show. ■

---

[3] Equivalently, $\boldsymbol{d}_0 \max_x h(x, c, \ldots, c) \leq e_0$ and, for all $t \geq T$ with $\boldsymbol{d}_{t-1} > 0$ and for all $x \neq c$, $s^{t-1}\,[h(x, c, \ldots, c) - \bar{h}(x, c, \ldots, c)] \leq 1 - \bar{h}(x, c, \ldots, c) + e_t$.

Example: Voluntary provision of public goods

As an example for Proposition 1, consider the following scenario. There is a fixed number of periods, denoted T, with $T \geq 2$. In each period, each player i contributes either zero or one unit of 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 $\bar{x} = (x_1 + x_2 + \ldots + x_n)/n$. Specifically, the quantity produced is given by $f(\bar{x})$, where $f(x)$ is a differentiable function satisfying $f(0) = 0$, $f(1) = 2$, and $1 < df/dx \leq n$.

In the first period, the payoff of each player i is given by

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

Since $df/dx \leq n$, the marginal product does not exceed unity. Therefore, the outcome $x_1 = x_2 = \ldots = 0$ (i.e., no one contributes) is a symmetric equilibrium in the one-shot game, with the equilibrium payoff 0. On the other hand, since $df/dx > 1$, the average utility,

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

is maximal, and equals 1, if and only if $x_1 = x_2 = \ldots = 1$ (i.e., everyone contributes). In particular, (7) holds, with $c = 1$. Since $f(1 - 1/n) < f(1) = 2 \leq e_0$, the first condition in footnote 3 holds. The second condition is equivalent to

$$s^{T-1} \leq \frac{\dfrac{f(1) - f(1 - 1/n)}{1/n} - 1}{n - 1}. \tag{8}$$

Note that, since $1 < df/dx \leq n$, the right-hand side of (8) is greater than zero but less than or equal to 1. Since $d_1 = d_2 = \cdots = d_{T-1} = 1$, Proposition 1 tells us that the inequality (8) is both a necessary and sufficient condition for the 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 the dynamic game defined by (6). Since, by assumption, $0 < s < 1$, (8) can be interpreted as a requirement that T is sufficiently large and/or that s is sufficiently close to zero.

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 payoff functions, rather than acquire new ones, over the course of time. Specifically, participants in the game, who are initially uncertain about the consequences of a unilateral deviation 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) each other. Positive signals, or even the absence of negative ones, may indicate that a unilateral deviation from cooperation is unprofitable, and so cooperation may be attainable in the late periods of the game. If players expect a high number of repetitions, and hence significant benefit from future cooperation, they may be willing to cooperate also in the early periods, when they are still unsure 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' belief in the desirability of defection is weaker in the middle periods than at the beginning. This raises the question of which conditions concerning the evolution of players' beliefs about their payoff function are sufficient to guarantee cooperation in all periods.

There is obviously more than one way the scenario outlined above can be modeled. In particular, there is more than one possible mechanism by which 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 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 payoff function. 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 a premature termination of the interaction.

The model that we will now describe allows for such a correlation. It involves just two possible states of the world, a "good" state and a "bad" state. The (prior) probability that the good state obtains and there is at least one period is denoted g_0. The corresponding probability in the bad state is $\boldsymbol{b}_0$. The continuation probabilities in the good and bad states of the world are denoted g_1, g_2, … and $\boldsymbol{b}_1$, $\boldsymbol{b}_2$, …, respectively. The payoff function in the good state is denoted g, and in the bad state b. These payoff functions satisfy $\max_x g(x, d, \ldots, d) = g(d, d, \ldots, d) = 0$, $\max_x b(x, d, \ldots, d) = b(d, d, \ldots, d) = 0$, and $g(c, c, \ldots, c) = b(c, c, \ldots, c) = 1$. The last assumption, entailing that, if the players cooperate, their payoffs are the same in both states, is meant to exclude the possibility that players may learn the state of the world simply by observing their own payoffs. This case is trivial, since all uncertainly about the state of the world vanishes after the first period. Thus, this assumption implies that, if all the players employ strategy C (as defined in Section 2), the only information they have in period t that they did not have in the initial period is that the game has at least t periods. If, a priori, this is more likely to happen in one state of the world than in the other, the posterior probability that the first state has obtained is higher than the corresponding prior probability. This may affect the players' assessment of the desirability of defection. A more detailed examination of this effect follows.

In each period t, the probability that the payoffs are given by g is equal to p_t, the (posterior) probability, at time t, that the good state has obtained. This probability is given by

$$p_t = \frac{g_0 g_1 \cdots g_{t-1}}{g_0 g_2 \cdots g_{t-1} + \boldsymbol{b}_0 \boldsymbol{b}_1 \cdots \boldsymbol{b}_{t-1}} \tag{9}$$

(provided that the denominator is not zero; otherwise, p_t may be defined arbitrarily.) The probability that the payoffs are given by b equals $1 - p_t$. Therefore, the expected payoffs h_t are given by

$$\begin{aligned} h_t &= p_t\, g + (1 - p_t)\, b \\ &= b + p_t\, (g - b). \end{aligned} \tag{10}$$

Note that the payoff functions do not depend on the players' past behavior (in particular, $h_t^{\mathrm{c}} = h_t$), and that they satisfy assumptions (a) and (b) above. If everyone plays according to the strategy C, the players' expected payoffs in the dynamic game Γ are equal to $\boldsymbol{d}_0 + \boldsymbol{d}_0\boldsymbol{d}_1 + \ldots$ $(= e_0)$, where $\boldsymbol{d}_0 = g_0 + \boldsymbol{b}_0$ is the probability that the game has at least one period and, for $t \geq 1$, $d_t = p_t\, g_t + (1 - p_t)\, b_t$ is the period t continuation probability (unconditional on the state). If all but one player employs strategy C, and that player deviates by cooperating until a certain period t but using a different action x in that period, his maximum expected payoff, $M(x, t)$, is given by (3). Therefore, a necessary and sufficient condition for C to be a symmetric equilibrium strategy is that (4) holds for all $t \geq 1$.

Suppose that, in the good state of the world, the payoff function g is such that all players cooperating is an equilibrium in the one-shot game. (This is the sense in which the state is 'good.') Suppose also that, for $t \geq 1$, the continuation probability in the good state of the world, g_t, is greater than in the bad state, b_t. Then, the (posterior) probability p_t that the good state has obtained increases over time. Under these assumptions, it is at least plausible that, at some period T, this probability becomes large enough for cooperation to be an equilibrium behavior in the stage game. Maintaining cooperation from that period on is not a problem, since the inequality (4) clearly holds for all $t \geq T$. The question we now address is similar to the one addressed in the previous section: If the cooperation condition (4) holds in the first period $t = 1$ and in some later period T, under what circumstances does is automatically follow that it also holds in all the intermediate periods? By our Observation, this is the case if the maximum expected payoff of a deviating player is either monotonic or U-shaped as a function of the time t of deviation in the interval $1 \leq t \leq T$. As the proof of the following proposition shows, a sufficient condition for this is that, in the time interval under consideration, the continuation probabilities weakly increase in both states of the world, but at a (weakly) faster rate in the bad state.

PROPOSITION 2. *Suppose that defection from cooperation is unprofitable in the good state of the world, i.e.,*

$$g(x, c, \ldots, c) \leq 1 \text{ for all actions } x. \tag{11}$$

Suppose also that there is some period T such that the cooperation condition (4) *holds for* $t = 1$ *and for all* $t \geq T$.[4] *Then, a sufficient condition for cooperation to be sustainable is that* $g_1 \leq g_2 \leq \cdots \leq g_{T-1}$ *and* $b_1 - g_1 \leq b_2 - g_2 \leq \cdots \leq b_{T-1} - g_{T-1} \leq 0$.

[4] Equivalently, $\max_x\, [g_0\, g(x, c, \ldots, c) + b_0\, b(x, c, \ldots, c)] \leq e_0$ and, for all $t \geq T$ with $b_{t-1} > 0$ and for all $x \neq c$,

$$b(x, c, \ldots, c) - 1 - b_t - b_t\, b_{t+1} - \ldots \leq \frac{g_0}{b_0}\frac{g_1}{b_1} \cdots \frac{g_{t-1}}{b_{t-1}} (1 - g(x, c, \ldots, c) + g_t + g_t\, g_{t+1} + \ldots).$$

Proof. We show that, if the condition in the proposition holds, then the cooperation condition (4) holds for all $1 < t < T$. Suppose, per contradiction, that the condition in the proposition holds but there is some action x such that, for some $1 < t < T$, $M(x, t) > e_0$. Then, there exist some T_1 and T_2, with $1 \le T_1 \le T_1 + 2 \le T_2 \le T$, such that $M(x, T_1) \le e_0$ and $M(x, T_2) \le e_0$ but $M(x, t) > e_0$ for all $T_1 < t < T_2$. It follows, by the Observation, that there is some $T_1 < t < T_2$ such that $M(x, t) \ge M(x, t-1)$ but $M(x, t+1) < M(x, t)$. Since (by (3), (10), and the identity $\boldsymbol{d}_t\, p_{t+1} = g_t\, p_t$, which follows from $\boldsymbol{d}_t = p_t\, g_t + (1 - p_t)\, \boldsymbol{b}_t$ by means of (9))

$$M(x, t+1) - M(x, t) = \boldsymbol{d}_0 \boldsymbol{d}_2 \cdots \boldsymbol{d}_{t-1} \left(1 + \boldsymbol{d}_t\, h_{t+1}(x, c, \ldots, c) - h_t(x, c, \ldots, c)\right)$$

$$= \boldsymbol{d}_0 \boldsymbol{d}_2 \cdots \boldsymbol{d}_{t-1} \left(1 + \boldsymbol{d}_t\, b(x, c, \ldots, c) + g_t\, p_t\, (g(x, c, \ldots, c) - b(x, c, \ldots, c)) - h_t(x, c, \ldots, c)\right)$$

$$= \boldsymbol{d}_1 \boldsymbol{d}_2 \cdots \boldsymbol{d}_{t-1} \left[1 + (\boldsymbol{d}_t - g_t)\, b(x, c, \ldots, c) - (1 - g_t)\, h_t(x, c, \ldots, c)\right],$$

this implies that the expression in square brackets is negative, but a similar expression in which the index t is replaced by $t - 1$ is either positive or zero. Therefore, $(\boldsymbol{d}_t - g_t)\, b(x, c, \ldots, c) < (\boldsymbol{d}_{t-1} - g_{t-1})\, b(x, c, \ldots, c)$ or $(1 - g_t)\, h_t(x, c, \ldots, c) > (1 - g_{t-1})\, h_{t-1}(x, c, \ldots, c)$. The assumption $M(x, t) > e_0$ and (11) imply that $h_t(x, c, \ldots, c) > 1 \ge g(x, c, \ldots, c)$. The assumption $\boldsymbol{b}_{t-1} - g_{t-1} \le 0$ implies $p_{t-1} \le p_t$. Therefore, by (10), $b(x, c, \ldots, c) > 1$ and $h_{t-1}(x, c, \ldots, c) \ge \max\{h_t(x, c, \ldots, c), 0\}$. It follows that $\boldsymbol{d}_t - g_t < \boldsymbol{d}_{t-1} - g_{t-1}$ or $1 - g_t > 1 - g_{t-1}$. However, the latter inequality contradicts the assumption $g_{t-1} \le g_t$, and the former inequality, which is equivalent to $(1 - p_t)\, (\boldsymbol{b}_t - g_t) < (1 - p_{t-1})\, (\boldsymbol{b}_{t-1} - g_{t-1})$, contradicts the assumptions $\boldsymbol{b}_{t-1} - g_{t-1} \le \boldsymbol{b}_t - g_t$ and $\boldsymbol{b}_{t-1} - g_{t-1} \le 0$ (the latter of which implies $p_{t-1} \le p_t$). These contradictions prove that the condition in the proposition indeed implies that (4) holds for all $1 < t < T$. ■

Example: Constant continuation probabilities

A simple scenario satisfying the condition in Proposition 2 is that of constant continuation probabilities. More specifically, suppose that, over some time interval $1 \le t \le T$, the continuation probabilities are constant, and are equal to g in the good state of the world and b in the bad state, with $b \le g \le 1$. Then, the condition in Proposition 2 holds trivially. To fix ideas, suppose that in period T itself the continuation probabilities are zero; i.e., the game never continues past T. It then follows from Proposition 2 that a sufficient condition for cooperation to be sustainable is that condition (4) holds in the first period and in last period T, and, in the good state of the world, the deviation from cooperation by a single player is not advantageous to the deviating individual (i.e., (11) holds).

Counterexample: Decreasing continuation probabilities

It is instructive to give an example in which continued cooperation yields a higher expected payoff than the one-time payoff from defection in the first period, and cooperation is also an equilibrium behavior in the last period, but nevertheless cooperation is not sustainable because, in one of the middle periods, it is better to defect. This example is similar to the previous one except that in the bad state of the world the continuation probabilities are not constant but decrease over time. Specifically, the good and bad states of the world have equal prior probabilities. In the good state, there are exactly 40 periods. In the bad state, the number of periods is determined by a random variable T_b that has a lognormal distribution with $m = 3$ and $s = 0.3$. If $T_b \geq 40$, the number of periods is 40. Otherwise, the number of periods is the smallest integer greater than T_b. (This implies that, in the bad state, the probability is greater than 0.9 that the number of periods is between 12 and 31, with the expected number about 21.5.) 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, in the good state of the world, $a = 0$, and in the bad state, $a = 38$. Thus, in the bad state, the game is the prisoner's dilemma, with the unique equilibrium (d, d). In the good state, (c, c) is also a equilibrium, and thus (11) holds.

In this example, cooperation is not sustainable. This can be seen by comparing the expected payoff from unilateral defection in period t, $h_t(d, c)$, with the expected payoff from continued cooperation in the continuation game starting at time t, which equals $1 + e_t$ (see Figure 1). In the first and last periods, the latter is greater than the former. 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.

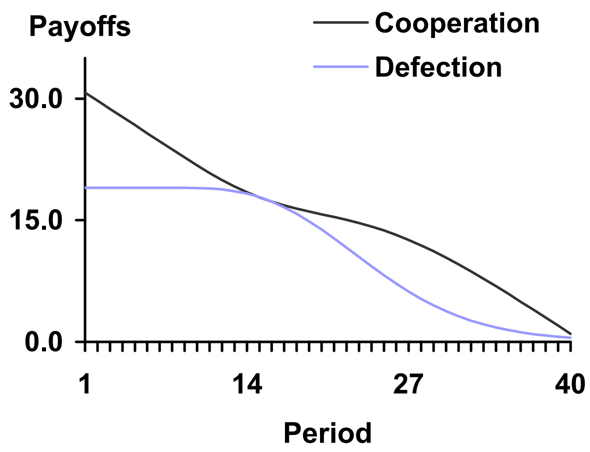

Figure 1. A game with incomplete information in which cooperation is not sustainable. (For details, see text.) In (only) one period, $t = 15$, a player's expected payoff from defection is greater than if cooperation were to continue until the last period.

Summary

In the finitely repeated prisoner's dilemma and similar symmetric games, cooperation is impossible to attain, for any number of repetitions. This paper's starting point is the observation that cooperation may be possible to attain if repeated cooperation among the players changes the game's payoff structure by making defection progressively less attractive, and if the number of repetitions is large enough. The large number of repetitions affects the players' incentives in two ways. First, in the late stages of the game preferences have already changed enough for cooperation to be an equilibrium behavior in the stage game. Second, in the early stages players who defect from cooperation can be effectively punished by the other players, who respond by defecting in all the subsequent periods: The resulting loss for the defector outweighs the one-time benefit from the defection. The problem, however, is that such a punishment becomes increasingly less effective as time progresses. Hence, even if the threat of a punishment deters defection in the early periods, and preferences in the late periods are such that cooperation is an equilibrium behavior in the stage game, cooperation may be unattainable because neither factor effectively discourages defection in one or more of the intermediate periods. The example in Figure 1 demonstrates this possibility.

The main results of this paper concern conditions under which players will be willing to cooperate with the others in all periods. More specifically, these conditions are sufficient for cooperation in all periods to be an equilibrium outcome, in that no player can benefit from deviating by being the first to take a different action in any period. We start with the simple observation that if deviation from cooperation is not beneficial in two given periods, then a sufficient condition for it not to be beneficial in any of the intermediate periods is that, in the time interval under consideration, the expected payoff for a deviating player is either a monotonic or U-shaped function of the time of deviation. We then proceed to examine this condition more closely under two rather general settings.

In the first setting, systematic preference changes are brought about by the players' actions in the preceding periods. Specifically, if everyone cooperates, then all players show increasing empathy towards the others. Consequently, the players' perceived costs and benefits become increasingly close to the social costs and benefits. In particular, if cooperation is socially beneficial, then, from some period T and onwards, it becomes an equilibrium behavior in the stage game. Our first main result (Proposition 1) is that, for cooperation in all periods to be an equilibrium outcome, it suffices that the continuation probabilities in the first $T-1$ periods are weakly increasing. In particular, it suffices that the game has at least T periods with probability 1.

In the second setting we examine, the payoff structure of the game does not change over time. However, the players' beliefs about their payoffs do change. Specifically, as time progresses, players become increasingly confident that if everyone else cooperates, it is best for them to do the same. More precisely, players assign increasing probability to the "good" state of the world, in which cooperation is an equilibrium behavior in the one-shot game, and decreasing probability to the "bad" state, in which it is not. The reason for these changes in beliefs is that we assume that the number of periods in the bad state tends to be less than in the good state. Consequently, as time goes by without the game terminating, the players' (posterior) belief that the good state has obtained increases. From some period T and onwards, it becomes strong enough for cooperation to be an equilibrium behavior in the stage game. Our second main result (Proposition 2) is that, for cooperation in all periods to be an equilibrium outcome, it suffices that, in the first $T-1$ periods, (1) the continuation probabilities in the good state are weakly increasing; (2) the continuation probabilities in the bad state also increase, at the same or faster rate; but (3) they are less than in the good state. (Part (3) of the condition is the one implying the changes in beliefs mentioned above.) In particular, it suffices that, in the good state, the game has precisely T periods, while in the bad state there is a constant positive probability of termination in each of these periods.

We view these two settings as examples of how our basic observations can be used to obtain concrete conditions for cooperation in repeated interactions in which the players' perception of the payoff structure may change over time. We believe these observations are also applicable in many other settings and examples.

References

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

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.

Bar-Ilan University
Department of Economics
WORKING PAPERS

1-01 **The Optimal Size for a Minority**
Hillel Rapoport and Avi Weiss, January 2001.

2-01 **An Application of a Switching Regimes Regression to the Study of Urban Structure**
Gershon Alperovich and Joseph Deutsch, January 2001.

3-01 **The Kuznets Curve and the Impact of Various Income Sources on the Link Between Inequality and Development**
Joseph Deutsch and Jacques Silber, February 2001.

4-01 **International Asset Allocation: A New Perspective**
Abraham Lioui and Patrice Poncet, February 2001.

5-01 **מודל המועדון והקהילה החרדית**
יעקב רוזנברג, פברואר 2001.

6-01 **Multi-Generation Model of Immigrant Earnings: Theory and Application**
Gil S. Epstein and Tikva Lecker, February 2001.

7-01 **Shattered Rails, Ruined Credit: Financial Fragility and Railroad Operations in the Great Depression**
Daniel A.Schiffman, February 2001.

8-01 **Cooperation and Competition in a Duopoly R&D Market**
Damiano Bruno Silipo and Avi Weiss, March 2001.

9-01 **A Theory of Immigration Amnesties**
Gil S. Epstein and Avi Weiss, April 2001.

10-01 **Dynamic Asset Pricing With Non-Redundant Forwards**
Abraham Lioui and Patrice Poncet, May 2001.

Electronic versions of the papers are available at
http://www.biu.ac.il/soc/ec/working_papers.html

11-01 **Macroeconomic and Labor Market Impact of Russian Immigration in Israel**

Sarit Cohen and Chang-Tai Hsieh, May 2001.

12-01 **Network Topology and the Efficiency of Equilibrium**

Igal Milchtaich, June 2001.

13-01 **General Equilibrium Pricing of Trading Strategy Risk**

Abraham Lioui and Patrice Poncet, July 2001.

14-01 **Social Conformity and Child Labor**

Shirit Katav-Herz, July 2001.

15-01 **Determinants of Railroad Capital Structure, 1830–1885**

Daniel A. Schiffman, July 2001.

16-01 **Political-Legal Institutions and the Railroad Financing Mix, 1885–1929**

Daniel A. Schiffman, September 2001.

17-01 **Macroeconomic Instability, Migration, and the Option Value of Education**

Eliakim Katz and Hillel Rapoport, October 2001.

18-01 **Property Rights, Theft, and Efficiency: The Biblical Waiver of Fines in the Case of Confessed Theft**

Eliakim Katz and Jacob Rosenberg, November 2001.

19-01 **Ethnic Discrimination and the Migration of Skilled Labor**

Frédéric Docquier and Hillel Rapoport, December 2001.

1-02 **Can Vocational Education Improve the Wages of Minorities and Disadvantaged Groups? The Case of Israel**

Shoshana Neuman and Adrian Ziderman, February 2002.

2-02 **What Can the Price Gap between Branded and Private Label Products Tell Us about Markups?**

Robert Barsky, Mark Bergen, Shantanu Dutta, and Daniel Levy, March 2002.

3-02 **Holiday Price Rigidity and Cost of Price Adjustment**

Daniel Levy, Georg Müller, Shantanu Dutta, and Mark Bergen, March 2002.

4-02 **Computation of Completely Mixed Equilibrium Payoffs**

Igal Milchtaich, March 2002.

5-02 **Coordination and Critical Mass in a Network Market – An Experimental Evaluation**

Amir Etziony and Avi Weiss, March 2002.

6-02 **Inviting Competition to Achieve Critical Mass**

Amir Etziony and Avi Weiss, April 2002.

7-02 **Credibility, Pre-Production and Inviting Competition in a Network Market**

Amir Etziony and Avi Weiss, April 2002.

8-02 **Brain Drain and LDCs' Growth: Winners and Losers**

Michel Beine, Fréderic Docquier, and Hillel Rapoport, April 2002.

9-02 **Heterogeneity in Price Rigidity: Evidence from a Case Study Using Micro-Level Data**

Daniel Levy, Shantanu Dutta, and Mark Bergen, April 2002.

10-02 **Price Flexibility in Channels of Distribution: Evidence from Scanner Data**

Shantanu Dutta, Mark Bergen, and Daniel Levy, April 2002.

11-02 **Acquired Cooperation in Finite-Horizon Dynamic Games**

Igal Milchtaich and Avi Weiss, April 2002.