Establishing Cooperation Without Pre-Play Communication

Randall L. Calvert

Working Paper No. 17
July 1998

W. ALLEN
WALLIS
Institute of
POLITICAL
ECONOMY

UNIVERSITY OF
ROCHESTER

Establishing Cooperation without Pre-Play Communication

Establishing Cooperation In a Perfect-Information PD

Focality in a two-player Prisoner's Dilemma. Suppose two players are to engage in a garden-variety symmetric repeated PD game with sufficiently light discounting. Formally, they face a problem of coordination, since, although there are cooperative equilibria, there are many different ones, as well as the noncooperative equilibrium. Given the symmetric positions of the players and the fact that full cooperation is both efficient and supportable in equilibrium, rather modest focal-point reasoning ought to lead the players to full cooperation.

Even if the players do not attribute focal point status to efficiency, a loose sort of forward induction reasoning may help: if Player 1 unilaterally cooperates for a move, Player 2 must suppose that Player 1 expects (or at least *had* expected) some sort of cooperative equilibrium, and Player 2 might conclude that it would be worthwhile after all to cooperate too, even if now his first cooperation is likely to be unilateral.

Of course, Player 1 might have been focused narrowly on a "grim trigger" strategy, and thus intend to defect indefinitely following his initial unrequited cooperation. Nevertheless, when he observes Player 2's unilateral cooperation in response, he must infer that Player 2 intended neither always to defect, perhaps making 1's grim trigger retaliation irrational. Importantly, both must conclude as well that they are not playing the same strategies and that they both see the possibility of a cooperative equilibrium. At some point, intuitively, one finds it unreasonable that a player, having observed unilateral cooperation by the other player, would think it very likely that the other player's most recent cooperation was the last concession he was willing to make, and that no further cooperation would be forthcoming. Short of such ungenerous thinking, even in the absence of any focality for a symmetric, efficient equilibrium, it seems likely that cooperation by one player would elicit cooperation by the other, and that a cooperative equilibrium will be acheived perhaps after some initial blundering.

If a symmetric equilibrium has no focal power, a bigger problem is introduced. Now either player may reasonably expect an equilibrium in which, say, he cooperates only on alternate iterations while the other player cooperates on every iteration (or may suspect the other player of aiming at such a one-sided outcome).

Establishing cooperation in a perfect-information, many-player PD.

Establishing cooperation when timing is an issue. (the Jurisdiction Game)

Establishing Cooperation When it Might Not Be Reciprocated

In the following analysis I ignore the coordination problem temporarily and turn to another, also important, difficulty that is overlooked in the above approaches:

--how cooperation might be established in a setting in which there is some mistrust.

--how cooperation can be established in the course of a single repeated interaction.

I examine one specific type of mistrust--the suspicion that one's partner may have nothing to gain from cooperating--and try to learn the conditions under which cooperation can be established and the patterns of emergence that might be observed.

In doing this, I avoid attaching any arbitrary first stage of communication, and I stick to a setting in which cooperation may be established between two players within a single repeated game.

I begin by introducing an element of incomplete information into the previous PD game, namely, concerning the (cooperate, cooperate) payoff.

		Player 2	
		C	D
Player 1	C	x_1, x_2	$-\beta$, α
	D	α, $-\beta$	0, 0

where $\beta > 0$; $\alpha > 1$; $\alpha - \beta < 2$

Each x_i takes on the value 1 with probability p_i and -1 with probability $1-p_i$; x_i is known only to player i, but p_1 and p_2 are common knowledge. The game is to be infinitely repeated with discount factor $\delta \in (0, 1)$.

If $x_1 = x_2 = 1$, then this is a PD game. However, if $x_i = -1$ then player i has no interest in cooperation, mutual or otherwise.

How can cooperation get started in this game? In order to assume away the coordination problem and concentrate on the problem of establishing "cooperation," I assume it's common knowledge that, if cooperation is established, it will be enforced through a strategy of "Tit for Tat with Apologies." TFTA has the following nice properties:

(1) Subgame perfection
(2) Robustness to unprogrammed defections
(3) It's a sensible way to limit the risk of being the first to cooperate

This sets up a signaling problem--if player i has $x_i = 1$ then he may signal that fact to the other player by cooperating on the initial move. The cost of this signal is the risk due to the possibility that the other player j has $x_j = -1$ and will not respond with cooperation; in that case i just loses β. In addition, there may in principle be a temptation, even when $x_i = 1$, to wait and let the other player bear the risk of signaling. This may be a matter either of playing it safe or of taking advantage of the other player's willingness to begin by cooperating unilaterally.

Conditions for cooperation--whether maintainable. It should be obvious that no player will want to signal cooperation if it would be impossible to maintain cooperation in the repeated game, even when $x_1 = x_2 = 1$. If TFTA is going to be the retaliation scheme that enforced cooperation, then we'll have to meet the usual conditions on the discount factor and payoffs:

$$\delta \;\geq\; \max\left[\frac{\alpha-1}{\beta+1}\;,\;\frac{1}{\beta+1}\right].$$

Conditions for cooperation--responding to a signal. Suppose $x_1 = 1$ but player 1 defects in iteration 1, while 2 cooperated in iteration 1. Under no circumstances is it rational for 2 to cooperate unless $x_2 = 1$, so player 1 now knows that cooperation is possible. It can be accomplished, however, only if player 1 now "apologizes": since it is now common knowledge that Player 1 knows that $x_2 = 1$: if Player 1 continues to defect, Player 2 will only assume that $x_1 = -1$ and will therefore continue to defect also. Player 1 stands to gain from cooperation, and delaying the apology only delays the onset of cooperation; thus if Player 1 is willling to initiate cooperation at all, he will prefer to initiate it immediately. To see the conditions under which this will be rational for player 1, we set the DPV of 1's payoff from period 2 onward from cooperating and proceeding to play TFT greater than or equal to 1's payoff from defecting now and forever:

$$-\beta \;+\; \frac{\delta}{1-\delta} \;\geq\; 0\;,$$

that is,

$$\delta \;\geq\; \frac{1}{\beta+1}$$

Which is identical to one of the lower bounds for TFTA to be an equilibrium, both of which must also hold anyway if player 1 is to respond.

Conditions for cooperation--unilateral signaling. Suppose next that player 2 is not expected to signal (as we will see, this will happen if p_1 is too small), but that the conditions for TFTA are met. Under what conditions will player 1 signal cooperation if $x_1 = 1$? Not signaling will certainly yield zero forever. Signaling may be rewarded with an "apology" response, followed by cooperation, or by no response, depending on player 2's type. The condition for signaling is thus

$$-\beta \;+\; \delta p_2\left(\alpha \;+\; \frac{\delta}{1-\delta}\right) \;\geq\; 0\;,$$

or

$$p_2 \;\geq\; \frac{(1-\delta)\,\beta}{\delta^2 \;+\; \alpha\delta(1-\delta)} \;\equiv\; P\;.$$

Conditions for cooperation--bilateral signaling. Alternatively, suppose that 2 can be expected to signal if $x_2 = 1$. Given that $x_1 = 1$, should 1 also signal, or wait to see if 2 signals first (and then apologize and begin cooperation)? The condition for 1 to signal immediately is:

$$p_2\,\frac{1}{1-\delta} \;-\; (1-p_2)\,\beta \;\geq\; p_2\,(\alpha \;-\; \delta\beta \;+\; \frac{\delta^2}{1-\delta})\;,$$

or

$$p_2 \;\geq\; \frac{\beta}{(1+\delta)(1+\beta)\;-\;\alpha} \;\equiv\; Q\ .$$

Equilibria. Comparing the two thresholds shows that, depending on the values of α, β, and δ, we could have either P > Q or P < Q , even if both are strictly between 0 and 1. Examples:

δ	α	β	P	Q
.8	2	3	.625	.577
.9	2	3	.303	.536
.9	2	8	.808	.530
.9	6	8	.593	.721

[This possibility exists because 1's decision of whether to signal is quite different under the two different conjectures about 2's intentions. Most obviously, the parameter α cuts in opposite directions in the two cases: if 2 will signal, a large α predisposes 1 not to signal, whereas if 2 will not signal, a large α predisposes 1 to signal.]

The graphs show the players' equilibrium signalling strategies in either case, as functions of p_1 and p_2. These are straightforward except for what happens in the "middle" box of the graph.

In the most straightforward case, P > Q (Graph 1): that is, player 1 is willing to signal when 2's probability of cooperation is at least P provided that 2 will signal too, but a higher likelihood of cooperation by 2 is required for 1 to signal unilaterally. In that case,

- (1) both players will signal whenever either p_i exceeds the higher threshold;
- (2) if both p_i < P and one is less than Q, then neither player will signal;
- (3) if p_1 > P but p_2 < Q, then only player 2 will signal, and vice versa;
- (4) finally, if both p_i are between Q and P, then there are multiple equilbria--one in which both signal, one in which neither signals, and perhaps mixed-strategy equilibria as well.

If both players have x_i = 1, then in cases (1) and (3) they will definitely end up cooperating, although in case 3 there will be a rough beginning as first one and then the other cooperates unilaterally. In case (2), the gains from cooperation will be lost. In case (4), a focus on efficient equilibria might predispose the players to one equilibrium or the other.

The situation is different if P > Q (Graph 2). In that case, there is a range of p_2 values (between Q and P) over which player 1 will signal if he thinks 2 will not, but will NOT signal if he thinks 2 will. In that case,

- (1) both players will signal only if both p_i exceed the higher threshold;
- (2) neither player will signal if both p_i are below the lower threshold;
- (3) in all other cases, either 1 or 2 will signal but not both, depending on the values of p_1 and p_2;
- (4) in particular, if both p_i values are between the two thresholds, then there are multiple equilibria--one in which 1 signals but not 2, one in which 2 signals but not 1, and perhaps mixed strategy equilibria as well.

Graph 1. **Thresholds for signaling only if opponent signals are lower than thresholds for signaling alone.**

<u>Graph 2.</u> **Thresholds for signaling only if opponent signals are <u>higher</u> than thresholds for signaling alone.**

The first three cases occur at different parts of the graph than before, but with the same range of observed outcomes. In case (4), now, it's likely that the two pure-strategy equilibria are not Pareto-comparable.

Thus we may observe a variety of possible patterns of learning to cooperate even in this simple setting. Players may forego gains from cooperation; may venture cooperation and get burned; may unilaterally establish trust; or may simply cooperate from the outset.

Establishing Cooperation in Multiple Settings

One feature of social order (especially, and really of institutions generally) that gets obscured when we talk in terms of repeated games is that it may constrain behavior in all sorts of different settings. Consequently, one trick in establishing cooperation would involve deciding whether to connect cooperation and enforcement in one setting with that in another. If two settings are different, the kind of uncertainties modeled above crop up separately in both. This combination yields some further interesting patterns and factors in establishing cooperation (& hence establishing social order).

As simply as possible: let G' and G'' be two stage games identical to that above. Their asymmetric-info payoffs are x_i' and x_i'' respectively. In each period t = 1, 2, 3,..., at most one of the two games is played for a single iteration. Specifically, in each period the next iteration of game G' occurs with probability p' and the next iteration of G'' occurs with probability p''(1-p'). With probability (1-p'')(1-p'), neither game is played in a given period. The players discount future payoffs by a factor of δ per period. Thus the effective discount factors, δ' and δ'', for G' and G'' respectively, are calculated from δ along with p' and p'', resp. Specifically, we can prove the following

Lemma. Suppose a stage game G is played repeatedly across periods 1, 2, 3, . . ., such that in each period t an iteration of the game occurs with probability p, but with probability (1-p) no iteration occurs. Suppose that players discount expected payoffs accruing in the next *period* (not necessarily the next iteration) by a factor δ. Then the effective per-iteration discount factor for the repeated game is given by $\delta_G = \dfrac{\delta p}{1 - \delta(1-p)}$.

Proof. The proof uses induction to calculate the expected discounted present value (DPV) of iterations of the game, ignoring the current period; it shows that this is equal to the sum over *t* from 1 to infinity of $*_G{}^t\, y$, where *y* is the payoff in each period. Thus $*_G$ is the effective discount factor.

First, calculate the expected DPV of the payoff *y* from the next single iteration. This will be the sum from periods 1 through infinity following the present period of the probability of having the next iteration in that period, times the discounted payoff if that happens:

$$\sum_{t=1}^{\infty} (1-p)^{t-1} p\delta^t y \quad = \quad y \, \frac{\delta p}{1 \; - \; \delta(1-p)} \quad = \quad \delta_G \, y \;\; .$$

Now consider the expected DPV of the payoff from the (T+1)-th iteration after the current period, assuming that the expected DPV payoff from the T-th iteration is $*_G{}^T y$. The probability of the first iteration taking place t periods after the present is (1-p)$^{t\text{-}1}$ p; once that happens in period t, the expected DPV (discounting from the present period, t = 0) of the T-th iteration thereafter is $\delta^t *_G{}^T y$. Summing over t gives the expected DPV of the (T+1)-th iteration from the present:

$$\sum_{t=1}^{\infty} (1-p)^{t-1} p\delta^t \, \delta_G{}^T y \quad = \quad \delta_G{}^T y \, \delta p \sum_{t=1}^{\infty} (1-p)^{t-1} \, \delta^{t-1} \quad = \quad \delta_G{}^T y \, \frac{\delta p}{1-\delta(1-p)} \quad = \quad \delta_G{}^{T+1} y \, .$$

Thus the expected DPV of all future iterations is

$$\sum_{t=1}^{\infty} \left[\frac{\delta p}{1 \; - \; \delta(1-p)} \right]^t y \quad = \quad \sum_{t=1}^{\infty} \delta_G{}^t y \, ,$$

so $*_G$ is the effective discount factor as required. □

Thus applied to the present two-game setting, the Lemma shows that

$$\delta^{\prime} \quad = \quad \frac{\delta p^{\prime}}{1 \, - \, \delta \, (1-p^{\prime})} \quad \text{and} \quad \delta^{\prime\prime} \quad = \quad \frac{\delta p^{\prime\prime}(1-p^{\prime})}{1 \, - \, \delta \, (1-p^{\prime\prime}+p^{\prime}p^{\prime\prime})} \;\; .$$

Since each period may produce a play of G', a play of G", or neither (but not both), we can also consider the combination of G' and G" as a single repeated game. A play of one of the two games will occur with probability 1−(1−p')(1−p") = p'+p"−p'p", so this combined game will have an implicit discount factor of $\delta^{*} \; = \; \dfrac{\delta \, (p^{\prime}+p^{\prime\prime}-p^{\prime}p^{\prime\prime})}{1 \, - \, \delta \, (1-p^{\prime})(1-p^{\prime\prime})}$. A bit of algebra shows that this is larger than either δ' or δ". If the players both turned out to have x_i = 1 in each of the two games, then they could lump the games together for strategic purposes, with cooperation in each game enforced by the threat of retaliation in both, creating a single game with discount factor δ* > δ' or δ".

Even if the discount factors in G' and G" were too low to support cooperation in either game alone, the combined game's discount factor might nevertheless be sufficient for a cooperative equilibrium. The interesting question now is, by what sequence of play might such cooperation be established? WLOG, suppose G' is drawn first.

Suppose δ' and δ" both insufficient to support cooperation in G' or G" alone, but that δ* would support cooperation in the combined game. That is,

$$\delta^{\prime} \; , \; \delta^{\prime\prime} \quad < \quad \max \left[\frac{\alpha-1}{\beta+1} \; , \; \frac{1}{\beta+1} \right]$$

but

$$\delta^* \geq \max\left[\frac{\alpha - 1}{\beta + 1}\ ,\ \frac{1}{\beta + 1}\right].$$

Cooperation is possible only if both players draw the higher payoff value in both games; this will be known only after both games have been played at least once. In simple equilibria, the players will not cooperate in G' at all until G" occurs; at that point, a cooperation must signal the high payoff in both games, and thus a willingness to cooperate in both G' and G" in all future plays. The observer of G' alone will see several periods of mutual defection, followed by the establishment of cooperation:

	G'	G'	G"	G'	G"	G"	G'	G'	. . .
Player 1	D	D	D	C	C	C	C	C	. . .
Player 2	D	D	C	D	C	C	C	C	. . .

[[At least if players are willing to punish in G" one another's failure to cooperate in initial plays of G', this same situation might generate the opposite pattern--players who cooperate in G' **"on speculation"** that cooperation may eventually be established in G". Then if one player draws the low payoff in G", cooperation becomes impossible, and the cooperation in G' also disappears.]]

Suppose δ' is large enough to support cooperation through TFTA in G' alone. Then they may signal a draw of $x_i' = 1$ as before. Then when G" occurs for the first time they might signal separately and try to establish cooperation in that game--whether or not δ" meets the threshold, since cooperation in G" could now be enforced with retaliation in G'.

Suppose that δ" as well as δ* will support cooperation through TFTA. Signaling generally won't be rational at the first occurrence of G', but will occur with G". After that, if $x_i'' = 1$ it's possible to play G' and G" as a single repeated game as long as $x_i' = 1$ as well--but now this must be signalled separately. The pattern might look like:

	G'	G'	G"	G"	G"	G'	G'	G'	G'	G"	. . .
Player 1	D	D	C	C	C	C	D	C	C	C	. . .
Player 2	D	D	C	C	C	D	C	C	C	C	. . .

and of course other patterns would arise if $x_i' = -1$ for one or both of the players.