Volume 112
September 2002

An impartial reasoning solution to the prisoner's dilemma: comment

By ODED STARK

Reprinted from

PUBLIC CHOICE

Public Choice **112:** 401–403, 2002.

An impartial reasoning solution to the prisoner's dilemma: comment

ODED STARK
Department of Economics, University of Oslo, P.O. Box 1095 Blindern, N-0137 Oslo, Norway and University of Vienna, Austria

Accepted 4 January 2001

In a carefully reasoned and elegantly articulated article, Frohlich (1992) provides an explanation why the equilibrium outcome of a one-shot prisoner's dilemma game is mutual cooperation. The explanation runs as follows:

> "Suppose that the game is played from an impartial point of view. Imagine the two players to be meeting and attempting to reach an (unenforceable) agreement on which strategy each will choose. Imagine that each player is still charged with choosing a strategy but is told that s(he) could either be Row or Column, and will randomly be assigned to one or another of those roles *after the strategic choices have been made*. In other words, assume that each player is ignorant regarding which player s(he) is and must therefore reason impartially.
>
> Consider the choices facing the players and imagine them considering their strategic options. What could they expect to get if they each agreed to cooperate? Were they to carry through and choose those strategies then they would each get 2 whether the final assignment was to his/her role or to the other person's role. Is this choice stable? Would either of them have an incentive to renege on the commitment to choose C? The answer is "No". If one player changed his/her choice to D then the payoff would either be 3 or 0 depending on the final assignment of roles. If the player is a maximiner (i.e. chooses to maximize their minimum guaranteed security level) then that prospect would be inferior to a guaranteed 2. If the player is an expected value maximizer, the expected value of (3 + 0)/2 is (by hypothesis) less than 2 and so is again inferior. Were the other party to defect concurrently then each would get 1 which is again inferior to 2. A quick consideration of joint strategies of CD, DC, and DD, reveal that they too yield inferior results. Thus, the mutually cooperative strategic choices C would be chosen from an impartial point of view . . . and it would be in equilibrium".

		Column C	Column D
Row	C	2,2	0,3
	D	3,0	1,1

It seems to me that this explanation is flawed. While in the setting laid out by Frohlich an agent cannot tell *ex ante* whether he will be a row player or a column player, the agent does know that he will not be both. In addition, the agent knows that if he ends up being the row player, the other agent will be the column player, and that if he ends up being the column player, the other agent will be the row player. Finally, the agent knows that one of the events just described will occur with probability 1.[1]

Referring to the two agents as F and G, let us follow the reasoning of agent F. (Given the symmetry of the game, the reasoning of agent G parallels the reasoning of agent F.) Agent F knows that if he becomes the row player, the other agent G must become the column player. Regardless of whether G sticks to C or not, F is strictly better off choosing D. Alternatively, if F becomes the column player, the other agent G becomes the row player. Regardless of whether G sticks to C or not, F is strictly better off choosing D. When not knowing which of several possible events will occur, but when in each possible event D is the preferred strategy, D will be chosen in any event. Thus, (C,C) will not be the outcome of the game.

An error in Frohlich's reasoning lies with the fact that in the standard single-shot prisoner's dilemma game, a given agent will never be in a position to receive the average of the payoffs (0,3), nor will he be in a position to choose between 0 and 3.

Frohlich is correct in sensing that subject to additional reasoning and complementary assumptions, the single-shot prisoner's dilemma game *can* result in the (C,C) outcome. (Examples of such procedures and this outcome are in Stark (1989, 1999).) It appears, however, that the specific line of reasoning that Frohlich alludes to cannot yield the mutual cooperation outcome.

Note

1. It seems that a formal description of the game is as follows: First, there is a chance move that determines with equal probabilities who is the row player and who is the column

player. Next, without the players knowing what choice chance made, and without a player knowing the strategy chosen by the other player, each player chooses C or D. Finally, the players receive their payoffs. The table portrayed in the text is the normal form of this game.

References

Frohlich, N. (1992). An impartial reasoning solution to the prisoner's dilemma. *Public Choice* 74: 447–460.

Stark, O. (1989). Altruism and the quality of life. *American Economic Review* 79: 86–90.

Stark, O. (1999). Siblings, strangers, and the surge of altruism. *Economics Letters* 65: 135–142.