

Essay on 'Cooperation Under the Shadow of the Future' by Pedro Dal Bó

*Eric Sjöberg*¹

Introduction

This essay is divided into two parts. The first part is a summary where I thoroughly try to describe the paper subject to this essay. I summarize the question, the experimental design and also the results. The second part is a subjective evaluation of the paper where I by points give my thoughts on the paper and its contents that give rise to further questions or inspire further discussions.

Question and motivation

In this paper, Dal Bó tries to evaluate the theoretical implications of infinitely repeated games. He finds this to be an area where there exists an abundance of theoretical literature with quite clear implications but that the empirical field is both underresearched and inconclusive in its findings. Dal Bó wants to improve on earlier empirical research by designing an experiment which not only consists of simulations of repeated games by using probabilities that the game ends after each round played but also by including finite repeated games as controls. The more specific question Dal Bó answers is how the possibility of future interactions alters behaviour in present interaction. Further it is also examined how an alteration of the payoffs affects cooperation levels to answer the question if players also react to higher incentives to cooperate.

All the questions examined in this paper are questions which are quite clearly answered using existing theory. Under the shadow of the future, cooperation should occur more often the higher the probability the game continues is. It is also quite clear that having a higher payoff from cooperation in comparison to playing the Nash equilibrium should facilitate cooperation. So the main achievement of this paper is to find a stronger empirical underpinning of these notions.

Experimental design

Dal Bó starts by summarizing earlier experimental literature and he argues that the inconclusiveness of earlier results stems from the bad experimental design, the problems ranging from participants having to play against the conductors of the experiments to the payoffs to the participants not representing the actual payoffs represented in the game.

¹Eric Sjöberg - Stockholm University, eric.sjoberg@ne.su.se. Essay for Introduction to Game Theory

In Dal Bó's own experiment he uses a prisoner's dilemma game as a stage game where the players are not able to communicate with each other and monetary payments were made according to the outcome. The game is then repeated with a commonly known probability which will represent then the possibility to interact again in the future (the shadow of the future). Dal Bó assigns one of three different probabilities $\delta \{0, 1/2, 3/4\}$ to each game sequence where the number represents the (unconditional) probability that at least one more round will be played. The hypothesis is then that the higher the probability the higher will the level of cooperation be. Dal Bó argues that even if the results of the experiment are in line with this theory it might not be due to the shadow of the future but rather the expected number of times the agents will interact. The main improvement that this paper brings to the field Dal Bó argues is a control for this effect. Therefore Dal Bó also conducts control experiments where each player with certainty know the number of rounds that will be played $\{0, 2, 4\}$ where the numbers coincide with the expected number of rounds to be played in the infinitely repeated games. The hypothesis will then be that the cooperation level in the first round in the type of game with probabilities will be higher than the games with certain number of rounds. Dal Bó argues that if this is true then we truly have a shadow of the future effect.

Another aspect is that Dal Bó separates the prisoner's dilemma game into two types (PD1, PD2) using different payoffs for each type. The payoffs are such that cooperation is an equilibrium when the continuation probability is equal to $3/4$ in both but only in one of them when the probability is equal to $1/2$. If the occurrence of outcomes agrees with these predictions it is an indication that players also are sensitive to the value of cooperation when deciding their strategies.

Results

All the results are in line with the theoretical predictions in order of direction. The magnitude is much harder to interpret. The higher the probability of another round taking place the higher was the rate of cooperation. Also it was clear that in the first round cooperation is higher in the modeled infinitely repeated game compared to corresponding game with a certain final round. Regarding altering the payoffs, the results here are also in line with existing theory. While the rate of cooperation did not differ much when the probability of continuing was high and cooperation was an equilibrium in both, it differed substantially for the lesser continuation probability.

All results were thus in line with existing theory in order of direction. The significance is confirmed by some statistical analysis including p-values but the best intuition is given by just looking at the very clear and instructive tables and graphs in the paper.

Discussion

I want to start by discussing another hypothesis present in the paper but that is given a bit less weight. The games are also formed in a way such that the strategies play the pair cooperate/deviate, deviate/cooperate every second round is an equilibrium in both types when $\delta = 3/4$ but only in PD1 when $\delta =$

1/2. This is according to theory but I would want to see a discussion on how Dal Bó thinks that the player would be able to coordinate on this equilibrium. The results regarding this hypothesis are inconclusive and he interprets them as weak evidence for the hypothesis since the alterations between the two outcomes are slightly more common in PD1 compared to PD2 when $\delta = 1/2$. Dal Bó admits that it is hard to interpret the results for this hypothesis especially since failed attempts to coordinate would also generate the same outcome. In comparison to the other hypothesis this is much vaguer and it does not contribute much more to this paper than to fill an additional page. Especially since even if the evidence would have been in favour of the hypothesis it would be hard to argue that the players actually followed this elaborate scheme and that the result was not only due to accidents and miscoordination. Especially since the expected number of rounds is only 2 I find it improbable that the players would be able to implicitly agree on this strategy.

Regarding strategy choices and identifying strategies, Dal Bó correctly writes *'looking at individual data does not enable one to identify the strategies used by the subjects'*. This is true but in order to examine the willingness to cooperate or if the data actually represents the strategies it puzzles me that he did not made the participants fill in a questionnaire on how they reasoned when they played the game or why he did not made a deep interview with a few randomly selected or even better a combination of both. Given that he had all subjects gathered he could have conditioned the collecting of the payoffs on filling in the form. Anyhow, the extra cost associated with the gathering of this data must be negligible in comparison to the possible benefits. Formation of strategies is a very interesting area and given the multiple of equilibria it would be interesting to know if the participants actually bothered to figure out a strategy and also followed it or rather just acted in the spur of the moment

There are also a number of more or less puzzling facts that do not get attention in the text. One is that cooperation occurs in the one shot game arranged with a frequency of approximately 10 % (which gives an indication that the own monetary outcome is not equal to the actual utility of a given outcome). But I find one relation that would be more interesting to discuss. Dal Bó argue that a reason for higher cooperation when δ is higher might be the number of periods being ahead. This is way he also runs the control experiments. But looking at the data, it seems like there exists one effect like this as well. Cooperation levels in the first round are 35 % when the certain number of rounds is 4 compared to 13 % when the certain number of rounds are 2. It seems like there is an effect of increasing the number of rounds as well. One can therefore ask oneself how much of the increase in cooperation levels when δ is raised can be attributed to interacting in the future and how much that can be attributed to the increase in the expected number of rounds played. By comparison of first round cooperation levels it seems very likely that raising δ has an effect but why cannot both effects be present? It would be interesting to try to disentangle the two effects.

In all it is an interesting paper. The experiment is well executed and the twist of controlling the infinitely repeated games with finite games is a very clever

way to examine and purify the effect of raising the probability of future interaction. The question itself is interesting and on a more philosophical note it would be interesting to try and extend the reasoning into interpreting the continuation probabilities as discounting factors. The usual disclaimers of behavioral economics apply. A special aspect that could explain the result would be to know if the participants thought of a higher δ as an exhortation to cooperate which intuitively does not seem too improbable. Given that the participants did know the different probabilities beforehand but not which one their game would be assigned they might think of the high δ as also a signal to cooperate when they received information of their δ . A better design might have been not to tell participants the different δ s but just the one assigned to their game.