Social Dilemma Games with Confirmed Proposals

Giuseppe Attanasi[1] – *Toulouse School of Economics (LERNA), France*
Aurora García Gallego[2] – *LEE, Universitat Jaume I, Castellón, Spain*
Nikolaos Georgantzís[2] – *GLOBE, Universidad de Granada, Spain*
Aldo Montesano – *Bocconi University, Milan, Italy*

Abstract

We propose a bargaining process as a way of playing and solving non cooperative games. We focus on social dilemma situations and show that sequential proposals which, if confirmed by players, determine real play, may give rise to equilibrium outcomes which differ from the standard non cooperative solution. Specifically, we show that, under standard assumptions, in a prisoners' dilemma with confirmed proposals, there is a unique *confirmed agreement* between players to behave cooperatively. In these games the equilibrium strategy can be unique even though the strategy space of each of the two players and the stages of the game itself are infinite. The experimental evidence obtained on this setup strongly confirms our theoretical predictions, given that almost all the pairs in the lab reach a cooperation agreement. Our specific bargaining structure seems to lead to even more cooperation than does an almost infinitely (indefinitely) repeated game which we studied in the lab for comparison purposes.

JEL classification: C72; C73; C91.

Keywords: Bargaining; Confirmed Proposals; Confirmed Agreement; Almost Infinitely (Indefinitely) Repeated Prisoner's Dilemma.

---

[1] Financial support by the Jean-Jacques Laffont Foundation and by the Chair "Georges Meyer" in Mathematical Economics at Toulouse School of Economics is gratefully acknowledged.

[2] Financial support by the Spanish Ministry of Education (ECO2008-04636/ECON) and Junta de Andalucía (P07-SEJ-03155) is gratefully acknowledged. This research was completed during the time A. García-Gallego and N. Georgantzís were visiting the University of Cyprus whose hospitality is gratefully acknowledged. Both authors want to thank the Spanish Ministry of Education for specific funding (grants PR2007-0152 and PR2007-0153).

1. Introduction

Bargaining plays a central role in situations of interaction among economic agents. Since the seminal contributions by Nash (1950, 1953), bargaining is a central theme for research undertaken in the framework of cooperative and non cooperative game theory. There is a huge literature on rationally justifiable play leading to cooperative outcomes in non-cooperative games.[3] The *confirmed proposal* mechanism introduced in this paper is another non-cooperative process leading to cooperative results. Through the bargaining protocol introduced here, we show how players could switch from a non-cooperative to a cooperative game by simply bargaining with their opponents, without changing the rules of the strategic interaction setup in which they are involved.

While several authors[4] have contributed to our understanding of the consequences of bargaining for the split of wealth among negotiating agents, Rubinstein's (1982) model illustrates an intuitively plausible and theoretically appealing way of reaching an agreement through sequential non cooperative play. While the model has been criticized for a variety of reasons, there is hardly any doubt that it expresses most researchers' point of view on how bargaining should be modeled and on how it actually takes place if the negotiating parties have the right to make proposals as well as to reject those received by others in order to make their own counterproposals until an agreement is finally reached. The consensus on the plausibility of this bargaining protocol is compatible with the fact that bargaining models have been thought as stylized analogues of real world situations in which the negotiators aim at reaching an agreement concerning the distribution of wealth. However, in many occasions, bargaining processes pursue more complex objectives as compared to the split of a pie. The critical reader may observe that such situations are not fundamentally different from bargaining over the split of a pie, as long as they ultimately affect the distribution of wealth. In this paper we illustrate the consequences of applying alternating proposal protocols as a way of playing and as a method of solving non cooperative games, focusing on social dilemmas. From a technical point of view, the most basic difference between our framework and that of bargaining over the split of a pie is that, in our model, two agents bargain about their strategies in a 2x2 game. Apart from the obvious departure from Rubinstein's (1982) model due to the finiteness of the set of possible agreements[5], in our setup, a confirmed agreement between bargaining agents concerns the pair of independent strategies in the *constituent* non cooperative game. This increases by one the degrees of freedom and, thus, the dimension of the outcome space,

[3] Relevant references are Harsanyi (1961), Friedman (1971), Smale (1980) and Cubitt and Sugden (1994).

[4] While an exhaustive list of the relevant references is beyond the scope of this paper, it is worth mentioning Harsanyi (1956, 1962), Sutton (1986) and Binmore (1987).

[5] For a formal treatment of this issue see the insightful analysis by Muthoo (1991).

allowing us to use bargaining with alternating proposals as a method of solving non cooperative games.

Before proceeding with formal details on our framework, we provide an intuitive overview. Suppose that two players bargain over the strategy profile to play, given that each player knows the opponent's set of possible strategies. Then, there is a constituent game whose execution leads to the two players' final payoffs and a dynamic game whose actions in each bargaining period are *proposals* of strategies for the constituent game.

Although we deal with constituent games with finite strategy spaces, the bargaining supergame built on them involves infinite number of strategies and the supergame itself has potentially infinite stages. Nonetheless, we show that, under mild assumptions, the equilibrium outcome of the bargaining process can be unique. We call *confirmed agreement* the corresponding equilibrium contract.

We have tested our theory in the lab. Our experimental results provide strong support for the prediction of cooperation in social dilemma games with confirmed proposals. This contrasts with the moderate fit obtained from other sequential bargaining experiments testing perturbations of the standard alternating proposals framework.[6]

The remaining part of the paper is structured as follows. Section 2 describes a theoretical framework for the study of bargaining as a solution of non cooperative games. Section 3 describes the experimental design implemented to test the predictions of our theory and discusses the results. Section 4 concludes.

2. Bargaining Games with Confirmed Proposals

2.1 The rules of the bargaining game

Games with confirmed proposals are interactive strategic situations in which at least one player, in order to give official acceptance of a contract, must confirm his proposed action in a game in combination to the action chosen by his opponent. The bargaining game can be a game with perfect or imperfect information and/or with complete or incomplete information. When information is incomplete, players can exploit the bargaining process to extract information on their opponent's type through their proposals.

[6] A prominent example is Binmore et al. (2007) including Rubinstein's reaction to the experimental test of his model, surprisingly arguing that his model should not necessarily be taken as a testable theory, but rather as a fable on human behavior. Earlier bargaining experiments include Sutton (1987), Neelin et al. (1988), Ochs and Roth (1989), Weg and Zwick (1999),

Throughout the paper, we will assume that only two players are involved in the bargaining game. We concentrate on two different families of games with confirmed proposals (GCP henceforth): those with *Conditional* and those with *Unconditional* proposals.

2.1.1 Games with Confirmed Conditional Proposals

In this subsection, we assume that both players have the same power of confirmation, i.e. the game is with alternating proposals. Therefore the rules of the game are symmetric, apart from the random selection of the proposer in the first bargaining stage.

Let us denote by S_i the (finite) strategy space for player $i = 1,2$ in the *constituent game*. Let us describe this super-game as a sequence of stages, each one composed by 3 steps. Suppose that player 1 starts the bargaining *super-game*.

Stage 1.1 He communicates to player 2 his will to follow a certain strategy $s_1^1 \in S_1$. It means that 1 would follow s_1^1 if (and only if) the bargaining process would come to a so-called "confirmed agreement" and s_1^1 would be part of this agreement.

Stage 1.2 Player 2 answers to 1's proposal by communicating her will to follow strategy $s_2^1 \in S_2$ if (and only if) 1 will confirm his previous strategy s_1^1.

Then, player 1 has two possible choices:

Stage 1.3 Player 1 can *c*onfirm or not the preceding strategy profile, i.e. c_1^1 can be *Yes* or *No*.

- if he *agrees* (c_1^1 = *Yes*) with the strategy profile (s_1^1, s_2^1), he communicates to player 2 that he *confirms* his proposal (i.e., he confirms he will follow s_1^1 knowing that player 2 will follow s_2^1). In that case, the bargaining process ends in stage 1 with a "confirmed agreement" (the contract is subscribed) and the two players receive the payoffs corresponding to the strategy profile (s_1^1, s_2^1) in the constituent game;

- if he *does not agree* ($c_1^1 = $ *No*) with the strategy profile (s_1^1, s_2^1), they move to the next stage.

Stage 2.1 The reply of player 2 in stage 1.2 becomes the proposal of this new bargaining stage, that is $(s_2^2 = s_2^1)$.

Stage 2.2 Player 1, *conditionally* to player 2's previous strategy s_2^1, replies by indicating a strategy $s_1^2 \in S_1$, which must be *different* from his proposal in stage 1.1, that is $s_1^2 \neq s_1^1$.

Then, player 2 has two possible choices:

Stage 2.3 Player 2 can *c*onfirm or not the strategy profile (s_2^1, s_1^2), i.e. c_2^2 can be *Yes* or *No*.

- if she *agrees* (c_2^2 = *Yes*) with the strategy profile (s_2^1, s_1^2), she communicates to player 1 that she *confirms* her proposal (i.e., she confirms that she will follow s_2^1 knowing that player 1 will follow s_1^2). In that case, the bargaining process ends with a "confirmed agreement" (the contract is signed) and the two players receive the payoffs corresponding to the strategy profile (s_2^1, s_1^2) in the constituent game;

- if she *does not agree* (c_2^2 = *No*) with the strategy profile (s_2^1, s_1^2), they move to the next stage.

<u>Stage 3.1</u> The reply of player 1 in stage 2.2 becomes the proposal of this new bargaining stage, that is ($s_1^3 = s_1^2$).

<u>Stage 3.2</u> Player 2, *conditionally* to player 1's previous strategy s_1^2, replies by indicating a strategy $s_2^3 \in S_2$, which must be *different* from her proposal in stage 2.1, that is $s_2^3 \neq s_2^2$.

Then, player 1 has two possible choices:

<u>Stage 3.3</u> Player 1 can *c*onfirm or not the strategy profile (s_1^2, s_2^3), i.e. c_1^3 can be *Yes* or *No*, and so on.

Notice that, in Games with Confirmed *Conditional* Proposals (GC*C*P, henceforth):

- the bargaining stages are, by construction, *overlapping*: in each bargaining stage, the proposal (reply) of the second-mover represents at the same time the proposal of the first-mover in the subsequent bargaining stage; hence, these games are by construction games with alternating offers, in which the power of confirmation has to be *symmetric*, in the sense that it is given to both players, with the possibility, for each of them, to exert it when being the first-mover in a stage, i.e. every two stages;

- each time a player proposes the same strategy in two consecutive stages, the game ends with a confirmed agreement, given that "*re-proposal*" means "*confirmation*". For instance, suppose that in stage *t* player $i = 1,2$ proposes s_i^t and that player $-i$, with respect to s_i^t, proposes s_{-i}^t. Hence, (s_i^t, s_{-i}^t) forms a possible agreement, that can be confirmed or not by player *i*. In case she disagrees with it, the next (overlapping) bargaining stage starts with a proposal of her opponent that *is conditioned to be* $s_{-i}^{t+1} = s_{-i}^t$. If she replies with $s_i^{t+1} = s_i^t$ the game ends already in bargaining stage *t*, since that re-proposal of s_i^t would mean confirmation of (s_i^t, s_{-i}^t). So, if she does not agree with (s_i^t, s_{-i}^t), her proposal at the end of stage *t* (being at the same time her reply to s_{-i}^t in stage *t* + 1) *has to* be a $s_i^{t+1} \neq s_i^t$. Suppose now that, in stage *t* + 1, player $-i$ does not agree with (s_{-i}^t, s_i^{t+1}), with $s_i^{t+1} \neq s_i^t$ Hence, at the end of this stage she *has to* propose a $s_{-i}^{t+1} \neq s_{-i}^t$, since that re-proposal of s_{-i}^t would mean confirmation of (s_{-i}^t, s_i^{t+1}).

2.1.2 Games with Confirmed Unconditional Proposals

Consider the game structure presented in the previous subsection (GC*C*P). Suppose that, for each $t = 1,2,\ldots,+\infty$, the proposal a player has to make in bargaining stage $t + 1$ *does not have to be conditional* to the opponent's reply strategy in t. In that case, we have a Game with Confirmed *Unconditional* Proposals (GC*U*P, henceforth). More precisely, GC*U*P differ from GC*C*P from step 2.1 onward of the bargaining procedure described in the previous subsection. In a GC*U*P, the first-mover in t, when deciding, at the end of this stage, not to confirm her proposal once known the strategy proposed as a reply by her opponent, is *not* obliged to play a *different* strategy in $t + 1$, since the new strategy she plays is not *conditioned* (contrarily to GC*C*P) to the opponent's reply in t. Thus, the bargaining stage $t + 1$ starts without any 'link' to the previous stage. That leads to two important differences in the game rules.

First of all, the bargaining stages are not overlapping and the first-mover (proposer) in stage $t + 1$ is not constrained to coincide with the second-mover (responder) in t. Hence, the proposer in each stage t can be randomly chosen or picked-up according to a predetermined rule. We call GC*U*P with *unilateral* (*asymmetric*) power of confirmation those in which the proposer in each bargaining stage is always the same player (chosen at the beginning of the game). Hence, only this player has the power to end the game, by confirming an agreement reached in a stage. We call GC*U*P with *alternating* (*symmetric*) power of confirmation those in which, once a player is randomly selected to be the proposer in stage 1, she will play as proposer in stage 1 and in each odd stage; the opponent will play as proposer in each even stage. Hence, players alternate in exerting the power to end the game (by confirming the agreement reached in a stage). [7]

Secondly, in each bargaining stage t, there are no restrictions on the set of players' feasible strategies. Hence, the proposer in t can choose to communicate to the responder any strategy from her initial set of feasible strategies and the latter can reply to this proposal by picking any strategy from his/her set of feasible strategies. The non-confirming player (in case the game is asymmetric) or the opponent (in case the game is symmetric) re-starts the bargaining procedure by communicating to the other player the will to follow a certain strategy (which *could* be the same proposed and not confirmed in the previous period or a different one), waiting until the other player replies with a strategy proposal before deciding whether to confirm or not the resulting strategic profile. To clarify this last point, let us consider the symmetric case. In a GC*C*P, in each bargaining stage $t = 1,2,\ \ldots,\ +\infty$, if the strategy proposed by player i in t is $s_i^t \in S_i$ and the reply of the opponent is $s_{-i}^t \in S_{-i}$, then, if this strategy profile is not confirmed in t, $-i$'s set of feasible

[7] Notice that unilateral power of confirmation is not possible, by construction, in a GC*C*P.

proposals in $t+1$ is restricted to s_{-i}^{t} and i's set of feasible replies to s_{-i}^{t} is restricted to $S_i \setminus s_i^t$. On the contrary, in a GC*U*P, for each t = 1,2, …, +∞, and for each (s_i^t, s_{-i}^t) in t, $-i$'s set of feasible proposals in $t+1$ is S_{-i} and i's set of feasible replies is S_i.

2.2 Prisoner's Dilemma with Confirmed Proposals

Let us now analyze the GC*C*P version of the most well-known social dilemma game, a Prisoner's Dilemma (PD). This means that the constituent game is a standard PD and the bargaining game built on it is an infinite dynamic game with perfect and complete information. The set of player i's feasible proposals (coinciding with her/his actions) is $S_i = \{A,B\}$, for $i = 1,2$. Figure 1 below shows, both the one shot simultaneous constituent game and, at the same time, all the possible outcomes of the bargaining (super)game with confirmed proposals one can build on it.

	A	*B*
A	**10, 10**	**25, 5**
B	**5, 25**	**15, 15**

Figure 1. Payoff matrix of the one-shot PD game

The constituent game has only one Nash equilibrium in dominant strategies, the profile (*A*,*A*). The same equilibrium outcome would be found in the standard two-stage game (without bargaining and without confirmation) in Figure 2, where player i = 1,2 moves first and player $j \neq i$ observes her "proposal" before choosing her own.

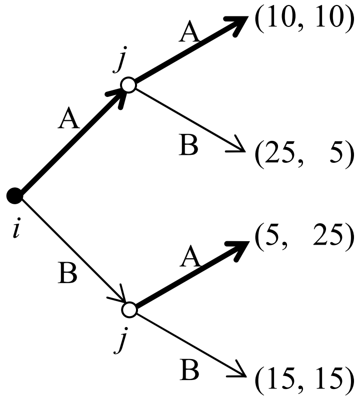

Figure 2. Two-stage Prisoner's Dilemma

In the next two sub-sections, we calculate the subgame perfect equilibrium outcome(s) of the GCP version of this game. More precisely, in the next subsection we analyze the confirmed conditional proposal version, whereas in subsection 2.2.2 we consider the two "unconditional" versions, the symmetric and the asymmetric one, respectively.

2.2.1 Prisoner's Dilemma with Confirmed Conditional Proposals

The Prisoner's Dilemma with confirmed *conditional* proposals is represented in Figure 3. The payoff structure of the bargaining game is the same as the PD in Figure 1. The first of the two payoffs refer to player 1 (randomly selected to be the proposer in stage 1 and in each odd stage) and the second refers to player 2 (proposer in each even stage).

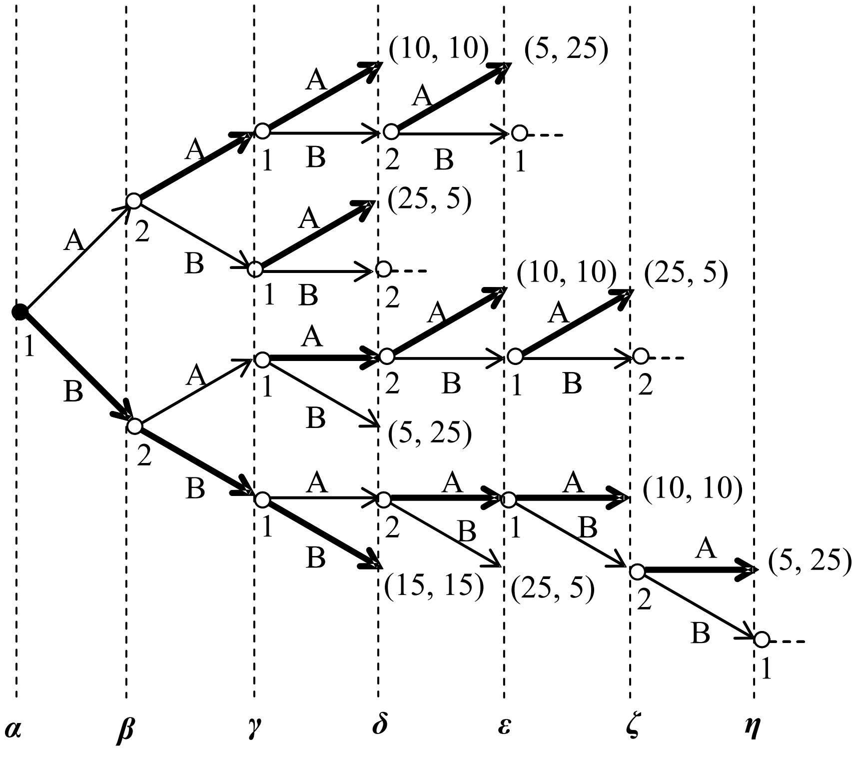

Figure 3. PD with confirmed conditional proposals

Notice that with confirmed conditional proposals, the bargaining stages are overlapping. For example, in the game depicted in Figure 3, bargaining stage 1 goes from time α to time δ; bargaining stage 2 goes from β to ε; bargaining stage 3 goes from γ to ζ, and so on. According to the notation in the figure, re-proposing the same action inside the same bargaining stage means confirming this action, hence ending the game.

Remember that we call '*confirmed agreement*' the subgame perfect equilibrium contract agreed upon by the two players, i.e. the proposal $a_i^t \in \{A,B\}$ and the counter-proposal $a_j^t \in \{A,B\}$ of the confirmed equilibrium outcome $(a_i^t,\ a_j^t \mid a_i^t,\ Yes_i^t \mid (a_i^t, a_j^t))$, given that *i* and *j* play, respectively, as proposer and responder in the bargaining stage *t*, in which the contract (a_i, a_j) is confirmed.

Proposition 1. The prisoner's dilemma with confirmed *conditional* proposals has a *unique* subgame perfect equilibrium, inducing the cooperative confirmed agreement in the *first* bargaining stage.

Proof. Let us consider the infinite game depicted in Figure 3. Each tree not marked with lines in bold indicates a partial history outside the equilibrium path, because in the previous stage the proposer chooses the action (confirmation) that does not lead to it. Let us explain why each non-marked tree is outside the equilibrium path. We use the following *weakly dominance criterion*: if the proposer in stage *t* can obtain her highest possible payoff by confirming a strategy profile in that stage, then she weakly prefers confirming instead than continuing the game, given that: (*i*) in every stage $t + k$, with k = 1,2, …,+∞, the highest payoff she can obtain is equal to the one obtained by ending the game (through a confirmation) in stage *t*; (*ii*) this highest payoff can be obtained only by confirming the same strategy profile confirmed in stage t. Given that the constituent game is symmetric, the highest possible payoff is the same for both players and is equal to 25. Let us look at Figure 3, there are four decision nodes where the proposer can confirm the strategy profile giving her 25. At time δ, after history $h = \{(A),(A),(B)\}$, player 2 can obtain 25 by choosing A, hence confirming strategy profile (A,B) as outcome of the bargaining game. If player 2, instead of confirming, would choose to continue the game, in any subgame in the continuation game he cannot obtain more than 25. Moreover, she can obtain 25 only by confirming the same strategy profile he can already confirm at time δ. Therefore, confirming (B,A) at time δ weakly dominates continuing the game. The same occurs to player 1 at time γ after history $h = \{(A),(B)\}$ and at time ε after history $h = \{(B),(A),(A),(B)\}$, and to player 2 at time ζ after history $\{(B),(B),(A),(A),(B)\}$. Disregarding the actions which are dominated according to the preceding criterion in each stage *t* and applying backward induction, we find that there is only one subgame perfect equilibrium, which leads to the terminal history $h = \{(B),(B),(B)\}$.■

Thus, in the unique subgame perfect equilibrium of the game, the proposer in stage 1 proposes strategy *B*, the responder (knowing the opponent's proposal) answers by indicating strategy *B* and the proposer (knowing the opponent's best-reply) re-proposes and confirms strategy *B*, such that the constituent game strategy profile (*B*,*B*) is the unique equilibrium outcome of the conditional confirmed proposal game. This agreement is confirmed in the first bargaining stage.

This result holds independently of the payoff structure of the constituent social dilemma game represented in Figure 1, as far as the payoff ranking does not change.

2.2.2 Prisoner's Dilemma with Confirmed Unconditional Proposals

The Prisoner's Dilemma with confirmed *unconditional* proposals is represented in Figure 4. The payoff structure of the PD is the same as in Figure 1, with the first of the two payoffs referring to player 1 and the second referring to player 2. The game in Figure 4 with $i = 1$ and $j = 2$ is a PD with *asymmetric* confirmed proposals: the proposer in each bargaining stage is always player 1 (randomly chosen at the beginning of the game). Hence, only this player has the power to end the game, by confirming an agreement reached in a certain stage. The game in Figure 4 with $i = 2$ and $j = 1$ is a PD with *symmetric* confirmed proposals: player 1 is randomly selected to be the proposer in stage 1 and so also in each odd stage; player 2 will play as proposer in each even stage.

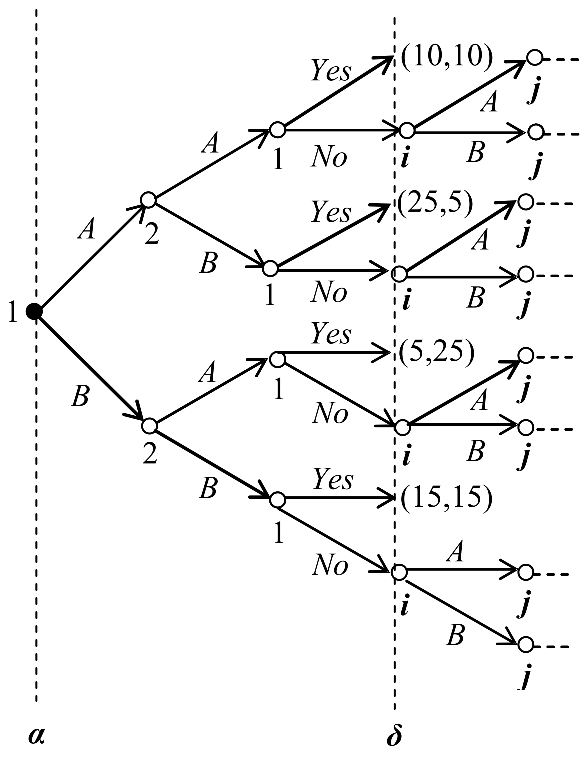

Figure 4. PD with confirmed unconditional proposals

Proposition 2. The Prisoner's Dilemma with confirmed *unconditional* proposals and *symmetric* power of confirmation has a *unique* subgame perfect equilibrium *outcome*, the cooperative agreement, that can be confirmed in *any* bargaining stage $t = 1,2, \ldots,+\infty$.

Proof. First of all, notice that players cannot agree in equilibrium on the contract (B, A), giving the proposer a payoff of 5. In each bargaining stage t, the proposer in that stage will never confirm that contract, given that she can always commit to play the strategy (*A*, *Yes*), allowing her a payoff of at least 10 in stage *t*. For the same reason, players cannot agree in equilibrium on the contract (*A*, *B*), giving the responder a payoff of 5. In each bargaining stage *t*, the responder in that stage will never reply to a *A* proposal with a *B* proposal: committing on replying with *A* in stage *t* and then, becoming the proposer in $t + 1$, playing the strategy (*A*, *Yes*), allows her a payoff of at least 10 in stage $t + 1$. Moreover, the contract (*A*, *A*) cannot be an equilibrium outcome and this can be verified by using a stationarity argument. Given that the game horizon is infinite, all subgames beginning on even dates are identical and the same holds for all subgames beginning on odd dates. Since the players are rational, strategy profiles confirmed at stage *t* will be the same as the ones that would have been confirmed at $t + 2$, with $t = 1,2, \ldots,+\infty$. Hence we can characterize a subgame perfect equilibrium based solely on stationary strategies. Suppose that (*A*, *A*, *Yes*) is an equilibrium outcome, leading to the payoff profile (10, 10) in stage 1. In a stationary equilibrium, the payoff profile at the end of stage $t = 1$ has to coincide with the payoff profile at the end of stage $t > 1$, for each $t = 1,2, \ldots,+\infty$. Therefore, given that the game starting in α and the one starting in δ are isomorphic (the set of strategies in the two games are the same and the constituent game is symmetric), we can assign to each non-terminal node at the end of bargaining stage 1 (time δ) the payoff profile (10, 10). That would lead the proposer, player 1, to choose *Yes* at the end of stage 1 in all but one nodes in which he is active, hence confirming (not declining) every strategy profile apart from (*B*, *A*). In particular, it is $(A, A, Yes) \sim_1 (A, A, No)$, since he obtains the same payoff in the non-terminal node (*A*, *A*, *No*) and in the terminal node (*A*, *A*, *Yes*). Going backward, the responder (player 2) would best-reply to *A* with *A* and to *B* with *B*. Hence, at the beginning of the first stage, player 1 would propose *B*, player 2 responds with *B* and player 1 confirms, so leading to the payoff profile (15, 15) at the end of stage 1, which contradicts that the confirmed agreement (*A*, *A*) is a stationary equilibrium outcome. Therefore, only (*B*, *B*) can be an equilibrium outcome. Let us verify that the game ends in some stage *t* with the following plan of actions: (*B*, *B*, *Yes*). Given that we assign to each non-terminal node at the end of bargaining stage 1 (time δ) the payoff profile (15, 15), the proposer does not confirm (*A*, *A*) and (*B*, *A*), confirms (*A*, *B*) and does not decline (*B*, *B*): notice that $(B, B, Yes) \sim_1 (B, B, No)$, which implies the confirmation of (*B*, *B*) in stage *t* or in any subsequent stage. Going backward, the responder would best-reply to *A* with *A* and to *B* with *B*, thus leading the proposer to start bargaining with a *B* proposal. ■

Proposition 3. The set of subgame perfect equilibrium outcomes (confirmed agreements) in the Prisoner's Dilemma with confirmed *unconditional* proposals and *asymmetric* power of confirmation is $\{(A,A),(A,B),(B,B)\}$. Each of these three outcomes can be confirmed in *any* bargaining stage $t = 1,2, \ldots,+\infty$.

Proof. Notice again that players cannot agree in equilibrium on the contract (*B*, *A*), giving the proposer a payoff of 5 (same argument used in the proof of Proposition 2). Players can instead agree in any stage *t* in equilibrium on the contract (*A*, *B*), giving player 2 a payoff of 5. This is because player 2, being the responder in each bargaining stage, cannot commit on any counter-proposal allowing her a higher payoff. In fact, assigning to each non-terminal node at the end of bargaining stage 1 (time δ) the payoff profile (25, 5) leads player 1 to choose *No* at the end of stage 1 in all but one node in which he is active: given that (*A*, *B*, *Yes*) $\sim_1$ (*A*, *B*, *No*), according to the same weakly dominance criterion used in Proposition 1, he confirms (*A*, *B*). Going backward, the responder (player 2) is indifferent between *A* and *B* in both nodes in which she is active. Given that, the proposer has a weak preference for starting with proposal *A* in α, in case he rationally attaches at least a small probability to player 2's counter-proposing *B* to *A*. Therefore, the agreement (*A*, *B*) can be confirmed in equilibrium in any bargaining stage *t*. This result can be used to prove that also the contract (*A*, *A*) can be an equilibrium outcome. Suppose that when player 1 chooses *No* at the end of bargaining stage 1, the payoff profiles in the continuation game are (25, 5) in correspondence to (*B*, *B*, *No*) and (10, 10) for all the other non-terminal histories. With this payoff structure of the continuation game, at the end of stage 1, player 1 chooses *No* after (*B*, *A*) and (*B*, *B*), confirms after (*A*, *B*) and is indifferent between confirming or not after (*A*, *A*). Going backward, the responder (player 2) replies with *A* in both nodes in which she is active (ensuring herself a payoff of 10), hence leading player 1 to be indifferent between proposing *A* or *B* at the beginning of the stage. Therefore, the agreement (*A*, *A*) can be confirmed in equilibrium in each bargaining stage *t*. The fact that also the agreement (*B*, *B*) can be confirmed in equilibrium in any bargaining stage *t* can be easy proved through the same stationarity argument used for Proposition 2. ■

Therefore, when proposals are not conditional but the power of confirmation is symmetric, we find the same equilibrium outcome as in the conditional confirmation case (Proposition 2 and Proposition 1): the two players agree on the cooperative outcome.

However, the equilibrium is unique when proposals are conditionals and the game ends with cooperation already in the first bargaining stage. When instead proposals are not conditional, in equilibrium the two players could cooperate immediately or they could need two, more or infinite periods to reach the cooperative outcome.

The same can happen also when the power of confirmation is asymmetric (Proposition 3). However, in this last case, the cooperative agreement is not the unique equilibrium outcome, given that all outcomes, except for the worst one for the proposer, can be sustained in equilibrium at each stage. Notice that only in this last case the Nash equilibrium of the constituent PD is an equilibrium outcome of the bargaining game. When the power of confirmation is symmetric, the Nash-outcome of the social dilemma game cannot be confirmed as an equilibrium of the infinite bargaining game. In order for this result to hold, conditionality of the proposals is not a necessary condition.

3. An Experimental Prisoner's Dilemma with Confirmed Unconditional Proposals

Let us consider a world with two persons with two strategies available to each one of them: *Defect* (*A*) and *Cooperate* (*B*). Assume also that the combinations of the two players' strategies yield a one-shot Prisoners' Dilemma (PD henceforth) type of situation, as the one represented in Figure 1. From traditional game theory, we know that one of the most appropriate environments for cooperation to emerge in a PD is one in which the game is infinitely repeated. In fact, the zero-discounting, infinitely repeated version of the PD in Figure 1 has a subgame perfect equilibrium generating the history ((*B*,*B*), (*B*,*B*), (*B*,*B*), …), i.e. leading to cooperation in each repetition.

From the side of games with confirmed (unconditional) proposals, the power of confirmation does play a role: when only one of the two players plays as proposer in each bargaining stage, she can reach the asymmetric confirmed proposal which is advantageous for her. When, instead, the power of confirmation is symmetric, it is not possible to reach an asymmetric agreement in equilibrium.

Moreover, when the power of confirmation is asymmetric, the Nash Equilibrium outcome of the one-shot PD is still a subgame perfect equilibrium outcome of the game, while when players alternate in exerting the power to end the game, the non-cooperative outcome is not a subgame perfect equilibrium: all the subgame perfect equilibria lead to the cooperative (symmetric) confirmed agreement. Therefore, despite the fact that the constituent PD game has a unique Nash non-cooperative equilibrium, the confirmed proposal version always includes the cooperative Pareto-superior outcome among its subgame perfect equilibria, independently of the allocation of the power of confirmation. The non-cooperative outcome is not an equilibrium when the power of confirmation is symmetric. In that case, the Pareto-optimal symmetric outcome is the only subgame perfect equilibrium outcome of the game. The only confirmed proposal structure allowing the proposer to gain from his power of confirmation is the asymmetric one. Only in this game the asymmetric outcome involving the highest payoff for the proposer is a subgame perfect equilibrium.

Therefore, our first research question is: Does the GCP version of social dilemma games, lead unambiguously to the cooperative outcome? To empirically evaluate the success of our framework as a cooperation inducing device, we compare it to the well known collusion facilitating environment of an Indefinitely (almost infinitely) Repeated PD Game.

Our second research question is: Can a GCP with asymmetric power of confirmation as easily as predicted lead to symmetric confirmed agreements? Alternatively, we want to test whether this bargaining structure leads to a less cooperative behavior with respect to the case in which both players can alternatively make a proposal and confirm it. This comparison would shed some light on which features among those of the GCP structures (*conditional* vs *unconditional*, *symmetric* vs *asymmetric* power of confirmation) are crucial for the hypothesized higher level of cooperation.

From our theory, we know that modeling a PD as a game with conditional proposals leads with certainty to the cooperative outcome. The same happens without proposal conditionality when the power of confirmation is symmetric, even though in the latter case we cannot be sure that the cooperative outcome is obtained in the first interaction among players. Therefore, from a theoretical point of view, conditionality is not a necessary condition for cooperation in a social dilemma GCP. Rather, it is a necessary condition only for "immediate" cooperation. Suppose that, from an experimental point of view, we find the same (high) level of cooperative agreements confirmed already in the first bargaining stage both in the asymmetric unconditional environment and in the symmetric one. This would lead us to state that neither conditionality is a necessary condition for immediate cooperation nor symmetric confirmation power is a necessary condition for bargainers' cooperation at all. At that point, we could conclude that the mechanism of proposal – counterproposal – confirmation itself is the key feature for cooperation.

Finally, we are interested in GCP as a form of communication between players. Through the bargaining structure developed, we believe that it could be understood how players communicate, without using more 'explicit' communication devices. Moreover, we could shed some light on what a subject *wants* to communicate to another, on what she *is able* to communicate as well as to understand from the proposals received.

3.1 Experimental design

Participants were voluntary students recruited at the *Universitat Jaume I* in Castellón (Spain), at *Bocconi University* and the *Catholic University* in *Milan* (Italy). Sessions were conducted in appropriate rooms where subjects were seated in isolated cubicles in front of computer terminals which were connected through a computer network. A total of 324 experimental subjects participated in our experiments, with each subject participating only once. Average earnings were approximately €15 per subject. The experiment was programmed using the z-Tree software (Fischbacher, 2007).

Three treatments were run, all of which are built starting from the one-shot PD depicted in Figure 1. The first two treatments were Unconditional GCP versions of the game, one with Asymmetric and the other with Symmetric power of confirmation, denoted, respectively, by GCP*A* and GCP*S*. The third treatment, an *Indefinitely Repeated* (*IR* henceforth) PD, with a very low end-game probability of 2%, was run for comparison purposes and it constitutes our collusive benchmark.

- **Treatment GCP*A***

The game (with $i = 1$ and $j = 2$) we propose to subjects participating in this treatment is shown in Figure 4. At the beginning of the experimental session, pairs are randomly formed. Within each pair, each player is randomly chosen to play either the role of 'Proposer' or 'Responder'. Pairs and roles are fixed during the whole session. In each bargaining stage, the Proposer plays first and has the power to confirm the strategy profile in that stage, hence ending the game. Moreover, proposals in a bargaining stage are not conditional to the opponent's strategy in the previous bargaining stage.

Therefore, every time the Proposer decides not to confirm her proposal (once known her opponent's choice), she starts the next period by making a new proposal, which can be the same or different to the one made - and not confirmed - in the previous period. Subjects do not have to wait for the other pairs to end the game. Once a pair of them reach an agreement, they have to leave their cubicles, and proceed to a separate room in which they are individually paid.

- **Treatment GCP*S***

Figure 4, with $i = 2$ and $j = 1$, depicts the infinite game proposed to subjects participating in this treatment. At the beginning of the session, pairs are randomly formed. Within each pair, each player is randomly selected to play the role of a Proposer or a Responder in the first bargaining stage. Pairs are fixed during the whole session. In this treatment, however, roles change every period, that is, within each pair, each time the Proposer in period t does not confirm the strategy profile proposed, he plays as a Responder in period t +1, and vice versa. Hence, the two players alternate in exerting their power to end the game. As in treatment *GCPA*, proposals in a bargaining stage are not conditional to the opponent's strategy in the previous bargaining stage. Also, subjects do not have to wait for the other pairs to end the game. Once a pair reaches an agreement, they leave the room in order to proceed with individual rewards.

- **Treatment *IR***

In this treatment, we implement the almost infinite repetition of the simultaneous PD represented in Figure 1. It is a multistage game with perfect monitoring and complete information. At the end of each period, after each player knows the action played by her rival, the *z-Tree* program makes a random draw among 100 possible numbers: 1, 2, 3,…, 97, 98, 99, 100. If the outcome of the random draw is a number in the interval 1 to 98 (both included), then the game goes on with the partner matching. In case the outcome of the random draw is 99 or 100, then the game ends and players' payoffs inside each pair are determined according to the action profile played by the subject pair in the last period. The players know in advance that their behavior in the last period will be the one determining their payment. The random draw is the same for each pair of subjects participating in the same session, so that the game randomly ends at the same period for all pairs. In some way, this protocol can be seen as a version of the confirmed strategy setup, with nature being the "player" who randomly determines whether or not to confirm the strategy profile simultaneously proposed by the pair in that period.

3.2 Experimental Results

The same number of subject pairs (*N* = 54) have participated in each one of the three treatments. In this section we discuss the facts obtained from an exhaustive analysis of our data.

3.2.1 Cooperation and Length of the Bargaining Process

The simplest way of looking at our data is by observing the frequency of cooperation and the speed of reaching the corresponding agreement. Table 1 informs us on the first of these two issues. Almost all pairs reach the cooperative (Pareto-superior) confirmed agreement in the two *GCP* treatments. Specifically, the cooperation ratio is around 93% in treatment *GCPA* and around 91% in treatment *GCPS*.

Outcome	*No. of pairs in GCPA*	*No. of pairs in GCPS*
Cooperation (*B*,*B*)	50	49
Nash (*A*,*A*)	2	2
Proposer 'grabs' (*A*,*B*)	2	2
Responder 'grabs' (*B*,*A*)	0	1
TOTAL	54	54

Table 1. Outcomes and confirmed agreements in treatments *GCPA* and *GCPS*

In both treatments, only two pairs over 54 (3.5%) reach a confirmed agreement to behave à la Nash (*A*, *A*). Specifically, in one of the two cases in which this occurred in the *GCPA* treatment, this happened in the first period while in the second case it happened after 33 bargaining periods. In both cases, the proposer starts playing the non-cooperative action and the responder does not play the cooperative action at any bargaining period. In fact, in the first case, the proposer plays the non cooperative action, and is imitated by the responder. Then, the proposer confirms. Hence, it seems as if the Nash equilibrium was imposed by the proposer, given that he does not try any cooperation at all. In the second case, the proposer plays the non cooperative action 33% of the times and the cooperative one 66% of the times. However, independently of the proposal, the responder replies with the non cooperative action. Hence, it seems as if the Nash equilibrium here was imposed by the responder. It is striking that only in the cases in which one of the two players was absolutely committed to non cooperative behaviour the non cooperative outcome was observed.

Similarly, the asymmetric outcome involving one of the players defecting while the other cooperates appeared in another two occasions in treatment *GCPA* and in another three in treatment *GCPS*. In both cases in which this occurred in *GCPA*, the agreement was favourable to the proposer, who grabs the occasion since the responder has accepted such an unfavourable asymmetric outcome. Also, in both cases the agreement was reached in the first bargaining period and during their payment by the experimenters, the two responders involved recognized that they had accepted the asymmetric payoff by mistake!

We have spent these first lines of the discussion of our results on the few observations contradicting our theory in order to make as clear as possible that our theoretical prediction of full cooperation receives very strong support. As predicted by our theoretical analysis, the cooperation obtained in our framework is pervasive.

In order to quantify the success of the *GCP* framework as a cooperation-inducing device, we compare the results of the two *GCP* treatments with those obtained from the *IR* treatment. To facilitate the comparison, in each bargaining period t (for $t = 1,2,\ldots,65$) of the two *GCP* treatments, we report the percentage of cooperative pairs who have agreed on (B,B) in this or in previous periods. This percentage is reported in Figure 5 against the percentage of pairs cooperating in the same period of the *IR* treatment. It should be noted that there is a large literature on cooperation in repeated experimental PDs, but to our knowledge, the end-game probability of 2% used here is the highest of all those used in this type of design. Given that each period could have been the last one (hence the one determining the agreement and the relative payment), the percentage of pairs whose action profile has been (B,B) in period t (for $t = 1,2,\ldots,97$) represents the level of cooperation in that period.[8]

[8] We run four sessions over the *IR* treatment and each session had a different end period, randomly drawn by the z-Tree program. In particular, *IR* session 1 ended in period 97, *IR* session 2 in period 4, *IR* session 3 in period 91 and *IR* session 4 in period 88.

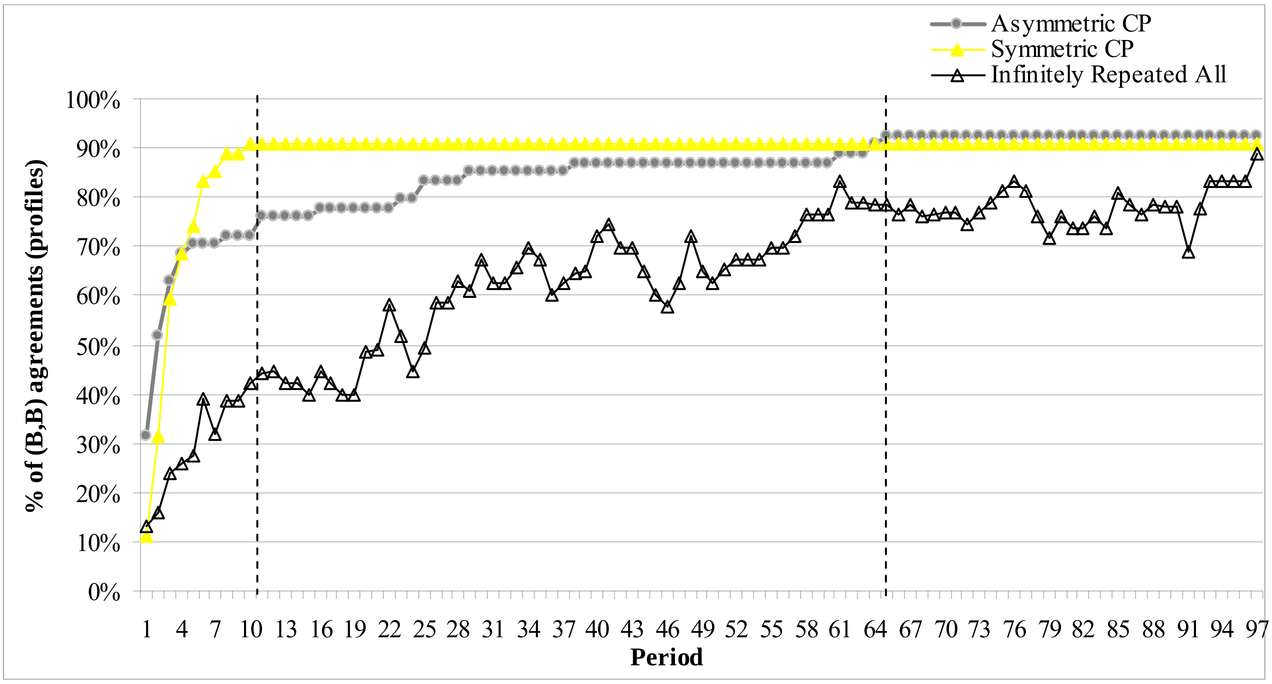

Figure 5. Cooperation rates for the three treatments, by period

Note that, overall, the cooperation rate in the *GCP* treatments is higher than in the *IR* treatment. Specifically, in the *IR* treatment the cooperation rate never exceeds the 90% threshold, which is reached in *GCPS* and even exceeded in *GCPA*.

Apart from the prediction of cooperation, there are other interesting patterns in our data which are more complex and are not predicted as such by our theory. We analyze first the length of the bargaining process in the case of the *GCP* Treatments. In Figure 6, we represent the relative frequencies of the agreement according to the period in which they were signed.

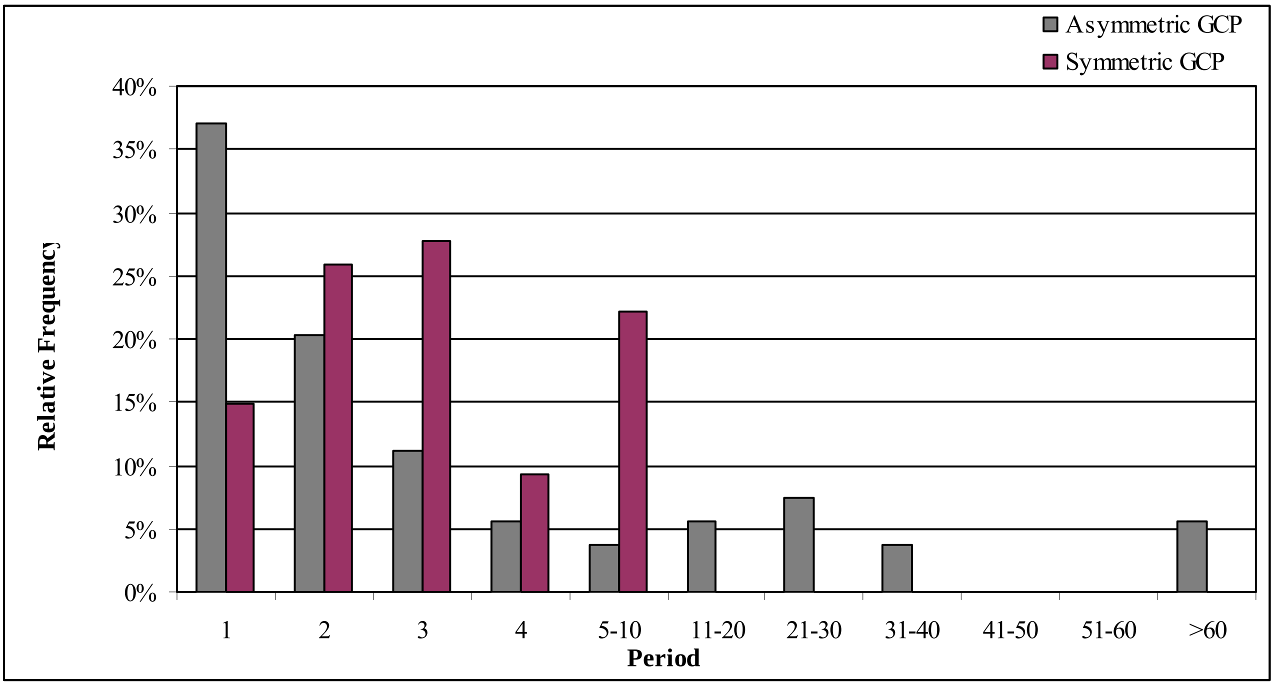

Figure 6. Agreement period distributions for *GCPA* and *GCPS*

First of all, it is worth observing that, despite the almost infinite nature of *GCP* structure, in the vast majority of cases, cooperation is reached in few periods. However, there seems to be some difference in the timing of agreements obtained under the two protocols, the symmetric and the asymmetric one. Specifically, more than a third (37%) of asymmetric proposers (20 pairs) starts with a cooperative strategy which is also adopted immediately by most (17) responders and the cooperative outcome is confirmed. In any case, the remaining three cooperative proposals were finally confirmed by period t=3 at most. Thus, the first is the most frequent agreement period in the *GCPA* treatment, whereas the corresponding modal agreement period in *GCPS* is t=3, despite the fact that in this treatment 19 first period proposers also started with a cooperative proposal, but they were followed by only 6 first period responders. Therefore, the existence of an exogenous leader more than doubles the likelihood of an agreement in the first period. However, it is also worth mentioning that the symmetric alternating proposals protocol has never lead to a bargaining process lasting more than 10 periods, whereas the asymmetric treatment has produced some very long games, including a small percentage of cases in which bargaining lasted for over 65 (!) periods. In summary, the existence of asymmetric confirmation power has a dramatic positive effect on the frequency of immediate cooperation agreements, but it also entails some low risk of extremely long bargaining processes. We feel that this finding needs further investigation in future research, because of its implications for the organization of bargaining processes aiming either at maximizing the likelihood of immediate agreements or at minimizing the average or the maximal time of the negotiations.

3.3 Beyond GCP testing: Proposal/confirmation strings as bargaining semantics

In order to give a complete explanation of all the patterns and dialogues emerging from the experimental data of the *GCP* Treatments, let us consider first different types of agents, concerning factors like their preferences on outcomes, their level of rationality, patience, understanding of the rules of the game, etc., as well as their beliefs on their opponents' personality. In that sense, if both agents are self-interested, rational expected utility maximizers, patient, have complete information and they believe that the opponent is self-interested, rational, patient and has complete information, they should propose (*B*,*B*), followed by a confirmation by the proposer. This is independently from the distribution of the power of confirmation. If (*B*, *B*, *Confirm*) is not obtained in the first period**,** it means that at least one of the previous hypothesis is not satisfied. Nonetheless, obtaining (*B*, *B*, *Confirm*) in the first period does not necessarily mean that all five hypotheses mentioned are satisfied. For example, an irrational agent, choosing randomly, could propose the 'right' action and/or confirm the 'right' action profile. Our experimental data show that (*B*, *B*, *Confirm*) is not obtained in the first period for over 2/3 of all pairs. This means that one or more of the five aforementioned hypotheses are not satisfied. The hypothesis that could be most easily weakened is the last one: at least one of the two players thinks that his/her opponent is not rational and/or not patient, etc.

Let us assume that players can be:

self-interested, i.e. (25, 5) $\succ_1$ (15, 15) $\succ_1$ (10, 10) $\succ_1$ (5, 25) and (5, 25) $\succ_2$ (15, 15) $\succ_2$ (10, 10) $\succ_2$ (25, 5);

- motivated (also) by 'moderate' *altruism*, i.e. (15, 15) $\succ_1$ (25, 5) $\succ_1$ (10, 10) $\succ_1$ (5, 25) and (15, 15) $\succ_2$ (5, 25) $\succ_2$ (10, 10) $\succ_2$ (25, 5);
- motivated (also) by *inequality aversion*, i.e. (15, 15) $\succ_1$ (10,10) $\succ_1$ (25, 5) $\succ_1$ (5, 25) and (15, 15) $\succ_2$ (10, 10) $\succ_2$ (5, 25) $\succ_2$ (25, 5).

If both players are rational, patient, and have perfect information and correct beliefs on their opponent, whatever combination of the three types of preferences above (i. e., *i* and *j* both self-interested, *i* and *j* both moderately altruist, *i* and *j* both inequality adverse, *i* inequality adverse and *j* altruist, and so on, for $i, j = 1,2$ and $i \neq j$) leads again to the Pareto efficient equilibrium outcome (*B*, *B*, *Confirm*) both in the asymmetric confirmed proposal treatment and in the asymmetric one.

Therefore, the reason of the continuation of the game after the first period for approximately 2/3 of all pairs has to be explained by weakening some other assumptions.

A careful look at the strings of strategies obtained from our experiments reveals that all the dynamic patterns observed can be interpreted as dialogues between the two negotiating parties. The question we address in this section is how different types of signals can be sent by each player to his opponent in a cheap-talk bargaining context before one of them confirms a given strategy profile. In fact, we argue that in the specific case of a PD game, both players aim at eliminating the asymmetric outcomes belonging to the set {(*Cooperate*, *Defect*), (*Defect*, *Cooperate*)}. Thus, it is of little if any relevance whether both players, or just one of them, have the right to confirm an announced strategy profile. For reasons which become clear by the end of the section, we concentrate on the case in which the power of confirmation is asymmetric. The bargaining semantics we elaborate for the asymmetric case can be easily extended to the symmetric one.

Before we proceed with the analysis of observed bargaining patterns, we establish a glossary of 'proposal-response' *strings* with their corresponding verbal interpretations. The rationale behind the dialogues constructed below is founded on the following basic heuristics which we claim are used by all pairs of humans who interact with each other in a strategic context like ours:

1. *Rational agents wish to know whether the others with whom they interact are rational*.
2. *If not, they take advantage of this*.
3. *Otherwise, they realize the benefits from cooperation, but they still fear that they may be fooled*.
4. *Once this fear is removed by the freedom of re-negotiation following unfair or inefficient outcomes, cooperation is the unique rational equilibrium*.

In terms of the PD context considered here, the following *strings of dialogues* provide an exhaustive list of basic conversations which can be used to understand the bargaining dynamics we observe.

*S*1. *"A-B-Confirm":*

- "Are you rational?"
- "No, I am not".
- "Then, I'll take advantage of this".

It is straightforward that, in order to take advantage of the responder's irrationality, the "*A-B*" sub-string should be followed by an immediate confirmation by the proposer.

*S*2. *"A-A-Withdraw":*

- "Are you rational?"
- "Yes, I am".
- "Then, we can both do better than being competitive to each other".

It is also straightforward to see why the "*A-A*" sub-string will not be confirmed by the proposer, who realizes that he can do better than obtaining the non-cooperative payoff of the game. This is a very strong and clear cut prediction of our framework, because the Nash equilibrium of the non-cooperative game is ruled out as one of the least expected outcomes. Nevertheless, in order to refer to a tiny percentage of outliers contradicting this prediction, we use the string *"A-A-Confirm"* denoted by *S*2-Nash.

Following the observation of the proposer's rationality, *S*2 could be followed either by a proposal to cooperate or by another *S*2 or a whole series of them. In the latter case, we will refer to a series of *S*2 by the term *patience challenge*, aimed at eliciting the responder's patience or time-discount factor. It is straightforward to see why a patience challenge followed by a *S*1 string should be interpreted as the responder's lack of patience or at least as a large difference in the two players' time discount factors. According to the same reasoning, the proposer's commitment to a very long series of *S*2 repetitions should be a result of her belief that the responder's impatience is significantly higher than her own.

As stated already, if the "*A*-*A*" sub-string belongs to a patience challenge, it will not be confirmed. Rather, it will be followed by an invitation of the proposer to cooperate. However, the responder may now want to elicit the proposer's rationality, taking at the same time a 'revenge' on her rival's initial doubt concerning her own rationality. This will give rise to a third type of string:

*S*3-*A*. *"B-A-Withdraw"*:
- "Ok, then. Let's cooperate".
- "(Wait! It is my turn to know:) Are you rational?"
- "Yes I am!"

It is very unlikely that this string of dialogue will uncover the proposer's irrationality if a string like *S*2 has preceded *S*3, revealing the proposer's perfect understanding of the situation. However, the *revenge motive* may still hold strong. In a similar manner as in the case of *S*2, a series of *S*3 may be observed corresponding to a "patience challenge" by the responder. Also, the incentive of checking responder's irrationality may even emerge after both players proposing *B*. Then, in some occasions the proposer may check the responder's willingness to cooperate, then ensure that the latter is a rational player. This will give rise to an alternative *S*3:

*S*3-*B*. *"B-B-Withdraw"*:
- "Ok, then. Let's cooperate".
- "Ok, then. Let's cooperate".
- "(Wait! It is my turn to know:) Are you rational?"

The predicted end point of all types of such dialogues between pairs of rational players will be:

*S*4. *"B-B-Confirm"*:
- "Ok, then. Let's cooperate".
- "Ok, then. Let's cooperate".
- "Confirmed".

This will be the end point of the bargaining process leading to the confirmation of the "*B*-*B*" sub-string of strategy proposals yielding the predicted strategy profile (*B*,*B*). It should be observed that the abstract setting of the PD used in our experiments contains several of the aspects which are central in more generic bargaining situations involving payoff asymmetries. Those aspects should be relevant in the presence of fairness considerations and Pareto dominance, which should guide the agents' endeavours towards economic efficiency.

In the appendix we classify all the data obtained in our experiments according to short or longer dialogues consisting of these four strings. We identify 8 dialogues which provide an exhaustive list of the bargaining histories observed in the two GCP treatments. A significant part (1/3) of them totally coincides with *S4*, whereas over 90% of them end with *S4*.

While we do not agree with the strategy of calling subjects' mistakes those observations which contradict a theory, we feel that the ability of our setup to organize such a large percentage of our observations, legitimates some final remarks on decision making errors. The possibility of making mistakes, which can derive from the weakening of one of the five hypotheses introduced in the beginning of this section, is the reason underlying the responder's behavior in dialogues 3 and 3-bis of the *GCPA* treatment (see Appendix 1) and in dialogues 3 of the *GCPS* (see Appendix 2). That possibility concerns both agents, although for the proposer only confirmation mistakes are relevant.[9] For the responder, making a mistake in the proposal is crucial, given that he cannot confirm or withdraw it. Notice that when at least one agent is sure that he/she will never make mistakes but he/she believes that his/her opponent will make mistakes with positive probability, then the game could never end, even if both agents were extremely patient. However, as the play unfolds both players can signal, through their proposals, counterproposals, and withdrawals, that they are rational and/or they have complete information of the rules of the game, thus influencing the opponent's belief or hope that they could make a mistake in a subsequent step of the bargaining process. In other words, they have the means to convince the opponent that he/she holds a wrong belief about their probability of an erroneous decision. Once both players are sure that their opponent will not make a mistake and is not impatient, they both agree on a (*B*, *B*, *Confirm*) outcome, at some subsequent stage of the bargaining game.

A skeptical reader may hurry to argue that the four strings of dialogue and their variations mentioned above are simply all possible combinations of the two players' strategies. However, it should be noted that our analysis involves predictions on the timing of the strings in a given dialogue, on the outcome of the confirmation/withdrawal choice and even on the possible repetitions of each one of them before the end point, *S*4.

Contrary to explicit verbal communication, we claim that the '*bargaining semantics*' proposed here contain the necessary and sufficient syntaxes for a dialogue between bargaining agents in a non cooperative context. Many authors have explored the role of communication on the ability of agents to reach cooperative outcomes. Open or controlled verbal protocols have been

[9] Making a mistake in the proposal or in the withdrawal has the only consequence of sending an involuntary wrong signal to the opponent. However, in Dialogue 5-bis we observe a proposer making a mistake in the confirmation, as recognized by himself during his payment by the experimenters.

often used by experimentalists to establish the cooperation-enhancing potential of communication. Cheap-talk signalling of agents' cooperative intentions or publicizing subjects' preference or belief elicitation results have also been studied and shown to enhance cooperation.

However, our approach achieves *full* cooperation in *all* cases in which *rational and patient* players are involved. Our experimental results confirm that the confirmed proposal framework offers the natural and minimal semantic bridge between cooperative and non-cooperative behavior, implementing cooperation with at least 90% of success. From now on we know that the minimal semantic charge required for an individual's *vocabulary* to support non cooperative bargaining leading to full cooperation is as little as two sentences:

- *"Are you rational?"*
- *"Let's cooperate"*.

The remaining job is done by a context in which the proposer can withdraw or confirm a potentially cheap-talk string of signals into an actual strategy profile.

4. Conclusions and further applications

Throughout the paper, we have defined Games with Confirmed Proposals and shown their positive effect on agents' ability of coordinating on Pareto efficient outcomes not belonging to the set of equilibrium outcomes of the constituent non cooperative game.

From an experimental point of view, this particular bargaining structure applied to social dilemma games seems to lead to *cooperation* more than do other cooperation-enhancing mechanisms[10], like the corresponding indefinitely repeated game with a tiny (2%) end-game probability. Moreover, in the Prisoner's Dilemma modeled as a Game with Confirmed Proposals, the existence of asymmetric *power of confirmation* does not significantly affect the frequency of cooperation. On the contrary, the existence of an exogenous leader increases the likelihood of immediate cooperation, although it entails some risk of very long negotiation games.

Finally, our experimental results show how Games with Confirmed Proposals can be used to create a glossary of bargaining semantics for tacit communication among agents concerning their *rationality/irrationality*, *patience/impatience*, *social and psychological preferences*, etc. through the signals contained in their proposals and confirmation strategies.

[10] For a recent example and overview of related references, see Nikiforakis et al. (2009).

References

K. Binmore, 1987. Perfect Equilibria in Bargaining Models. In K. Binmore and P. Dasgupta, editors, Economics of Bargaining. Cambridge University Press, Cambridge.

Binmore, K., Swierzbinski, J. and C. Tomlinson, 2007. An Experimental Test of Rubinstein's Bargaining Model, ELSE Working Paper #260.

Cubitt, R. and R. Sugden, 1994. Rationally Justifiable Play and the Theory of Non-cooperative Games, *Economic Journal*, 104, 798-803.

Fischbacher, U., 2007. Z-Tree: Zurich Toolbox for Readymade Economic Experiments, *Experimental Economics*, 10, 171-178.

Friedman, J. W., 1971. A Non-cooperative Equilibrium for Supergames, *Review of Economic Studies*, 38, 1-12.

Harsanyi, J., 1956. Approaches to the Bargaining Problem before and after the Theory of Games: A Critical Discussion of Zeuthen's, Hicks', and Nash's Theories, *Econometrica*, 24, 144–157.

Harsanyi, J., 1961. On the Rationality Postulates Underlying the Theory of Cooperative Games, *Journal of Conflict Resolution*, 5, 179–96.

Harsanyi, J., 1962. Bargaining in Ignorance of the Opponent's Utility Function, *Journal of Conflict Resolution*, 6, 29–38.

Innocenti, A., 2008. Linking Strategic Interaction and Bargaining Theory: The Harsanyi-Schelling Debate on the Axiom of Symmetry, *History of Political Economy*, 40, 111-132.

Muthoo, A., 1991. A Note on bargaining over a finite number of feasible agreements, *Economic Theory*, 1, 290-292.

Nash, J., 1950. The Bargaining Problem, *Econometrica*, 18, 155–162.

Nash, J., 1953. Two-person Cooperative Games, *Econometrica*, 21, 128–140.

Neelin, J., Sonnenschein, H. and M. Spiegel, 1988. A further Test of Noncooperative Bargaining Theory: Comment, *American Economic Review*, 78, 824–836.

Nikiforakis, N., Normann, H.-T. and B. Wallace, 2009. Asymmetric Enforcement of Cooperation in a Social Dilemma, *Southern Economic Journal*, forthcoming.

Ochs, J. and A. Roth, 1989. An Experimental Study of Sequential Bargaining, *American Economic Review*, 79, 355–384.

Rubinstein, A., 1982. Perfect Equilibrium in a Bargaining Model, *Econometrica*, 50, 97–109.

Smale, S., 1980. The Prisoner's Dilemma and Dynamical Systems Associated to Non-cooperative Games, *Econometrica*, 48, 1617-1634.

Sutton, J., 1986. Non-cooperative Bargaining Theory: An Introduction, *Review of Economic Studies*, 53, 709-724.

Sutton, J., 1987. Bargaining experiments, *European Economic Review*, 31, 272-284.

Weg, E. and R. Zwick, 1999. Infinite Horizon Bargaining Games: Theory and Experiments. In D. Budescu, I. Erev, and R. Zwick, editors, Games and Human Behavior: Essays in Honor of Amon Rapoport. Laurence Erlbaum Associates, Mahwah, NJ.

Appendix 1: Experimental Results – Treatment *CP-Asymmetric*

Dialogue 1 (*S*4):
Common belief of rationality

Proposer: Would you cooperate?
Responder: Yes!
Proposer: Then, let's cooperate.

No. of pairs
17

Period	Proposer	Responder	Confirmation
1	**B**	**B**	**Yes**

Dialogue 2 (*S*2-*Nash*):
P tests R's rationality but P is (extremely) impatient or irrational.

Proposer: Are you rational?
Responder: Yes, I am!
Proposer: I am not patient (or irrational).

No. of pairs
1

Period	Proposer	Responder	Confirmation
1	**A**	**A**	**Yes**

Dialogue 3 (*S*1):
P tests R's rationality and R is irrational.

Proposer: Are you rational?
Responder: No, I am not!
Proposer: Then, I'll take advantage of this!

No. of pairs
2

Period	Proposer	Responder	Confirmation
1	**A**	**B**	**Yes**

Dialogue 3-bis (*S*1 & *S*4):
P tests R's rationality and R is irrational;
but P does not take advantage of this.

Proposer: Are you rational?
Responder: No, I am not!
Proposer: I don't want to take advantage of this! = Would you cooperate?
Responder: Yes!
Proposer: Then, let's cooperate.

No. of pairs
1

Period	Proposer	Responder	Confirmation
1	A	B	No
2	**B**	**B**	**Yes**

Dialogue 4 (*S*2 & *S*4):
P tests R's rationality and R is rational.

Proposer: Are you rational?
Responder: Yes, I am!
Proposer: Would you cooperate?
Responder: Yes!
Proposer: Then, let's cooperate.

No. of pairs
9

Period	Proposer	Responder	Confirmation
1	A	A	No
2	**B**	**B**	**Yes**

No. of pairs
2

Period	Proposer	Responder	Confirmation
1-2	A	A	No
3	**B**	**B**	**Yes**

No. of pairs
1

Period	Proposer	Responder	Confirmation
1-22	A	A	No
23	**B**	**B**	**Yes**

No. of pairs
1

Period	Proposer	Responder	Confirmation
1-60	A	A	No
61	**B**	**B**	**Yes**

Dialogue 5 (*S*3-*A* & *S*4):
R tests P's rationality and P is rational.

Proposer: Would you cooperate?
Responder: Wait a minute! Are you rational?
Proposer: I am rational!
Proposer: Would you cooperate?
Responder: Yes!
Proposer: Then, let's cooperate.

No. of pairs
1

Period	Proposer	Responder	Confirmation
1	B	A	No
2	**B**	**B**	**Yes**

Dialogue 6 (*S*2, *S*3-A & *S*4) :
P tests R's rationality & R tests P's rationality;
both players are rational and patient.

No. of pairs
2

Period	Proposer	Responder	Confirmation
1	A	A	No
2	B	A	No
3	**B**	**B**	**Yes**

No. of pairs
1

Period	Proposer	Responder	Confirmation
1	A	A	No
2-3	B	A	No
4	**B**	**B**	**Yes**

No. of pairs
1

Period	Proposer	Responder	Confirmation
1	A	A	No
2-4	B	A	No
5	**B**	**B**	**Yes**

No. of pairs
1

Period	Proposer	Responder	Confirmation
1	B	A	No
2	A	A	No
3	**B**	**B**	**Yes**

No. of pairs
1

Period	Proposer	Responder	Confirmation
1	A	A	No
2	B	A	No
3	A	A	No
4	**B**	**B**	**Yes**

No. of pairs
1

Period	Proposer	Responder	Confirmation
1	A	A	No
2-6	B	A	No
7	A	A	No
8	**B**	**B**	**Yes**

No. of pairs
1

Period	Proposer	Responder	Confirmation
1-11	A	A	No
12-13	B	A	No
14	A	A	No
15	B	A	No
16	A	A	No
17	B	A	No
18-20	A	A	No
21	B	A	No
22-47	A	A	No
48	B	A	No
49	A	A	No
50	B	A	No
51-53	A	A	No
54-60	B	A	No
61-64	A	A	No
65	**B**	**B**	**Yes**

Dialogue 6-bis (*S*2, *S*3-*A* & *S*2-*Nash*) :
P tests R's rationality and R tests P's rationality;
P is impatient.

No. of pairs
1

Period	Proposer	Responder	Confirmation
1	A	A	No
2-5	B	A	No
6	A	A	No
7-10	B	A	No
11-12	A	A	No
13	B	A	No
14-15	A	A	No
16-17	B	A	No
18-19	A	A	No
20-23	B	A	No
24	A	A	No
25-29	B	A	No
30	A	A	No
31-32	B	A	No
33	**A**	**A**	**Yes**

Dialogue 7 (*S*2, *S*3-*B* & *S*4):
P tests R's rationality and R's willingness to cooperate.

Proposer: Would you cooperate?
Responder: Yes!
Proposer: Wait a minute! Are you rational?
Responder: I am rational!
Proposer: Would you cooperate?
Responder: Yes!
Proposer: Then, let's cooperate.

No. of pairs
1

Period	Proposer	Responder	Confirmation
1	B	B	No
2	A	A	No
3	**B**	**B**	**Yes**

No. of pairs
1

Period	Proposer	Responder	Confirmation
1	A	A	No
2	B	B	No
3	A	A	No
4	**B**	**B**	**Yes**

No. of pairs
1

Period	Proposer	Responder	Confirmation
1	A	A	No
2	B	B	No
3-5	A	A	No
6	B	B	No
7-8	A	A	No
9	B	B	No
10-11	A	A	No
12	B	B	No
13	A	A	No
14	B	B	No
15-16	A	A	No
17	B	B	No
18-24	A	A	No
25	**B**	**B**	**Yes**

No. of pairs
1

Period	Proposer	Responder	Confirmation
1-2	A	A	No
3	B	B	No
4-11	A	A	No
12	B	B	No
13-29	A	A	No
30	**B**	**B**	**Yes**

Dialogue 8:
'Trembling-hand Temptation' (cheap talk):
both players test the other's rationality and willingness to cooperate;
both players are rational and patient.

No. of pairs
1

Period	Proposer	Responder	Confirmation
1	A	A	No
2	B	B	No
3-4	A	A	No
5	B	B	No
6	A	A	No
7	B	B	No
8	A	A	No
9	B	A	No
10	A	A	No
11	**B**	**B**	**Yes**

No. of pairs
1

Period	Proposer	Responder	Confirmation
1	A	A	No
2	B	A	No
3-4	B	B	No
5	A	A	No
6	B	B	No
7-8	A	A	No
9	B	A	No
10	A	A	No
11	**B**	**B**	**Yes**

No. of pairs
1

Period	Proposer	Responder	Confirmation
1-3	A	A	No
4	B	B	No
5-7	A	A	No
8-12	B	A	No
13-15	A	A	No
16	**B**	**B**	**Yes**

No. of pairs
1

Period	Proposer	Responder	Confirmation
1-6	A	A	No
7	B	B	No
8	B	A	No
9	B	B	No
10-23	A	A	No
24	B	A	No
25	**B**	**B**	**Yes**

No. of pairs
1

Period	Proposer	Responder	Confirmation
1-7	A	A	No
8	B	A	No
9-29	A	A	No
30	B	B	No
31-37	A	A	No
38	**B**	**B**	**Yes**

No. of pairs
1

Period	Proposer	Responder	Confirmation
1-6	A	A	No
7	B	A	No
8-16	A	A	No
17-18	B	B	No
19-37	A	A	No
38	B	B	No
39-40	A	A	No
41-44	B	B	No
45-47	A	A	No
48	B	B	No
49-55	A	A	No
56	B	A	No
57-58	A	A	No
59	B	B	No
60-62	A	A	No
63	B	A	No
64	**B**	**B**	**Yes**

Appendix 2: Experimental Results – Treatment *CP-Symmetric*

Dialogue 1 (*S*4):
Common belief of rationality

Proposer: Would you cooperate?
Responder: Yes!
Proposer: Then, let's cooperate.

No. of pairs: 6

Period	Proposer	Responder	Confirmation
1	**B**	**B**	**Yes**

Dialogue 2 (*S*2-*Nash*):
P tests R's rationality but P is (extremely) impatient or irrational.

Proposer: Are you rational?
Responder: Yes, I am!
Proposer: I am not patient (or irrational).

No. of pairs: 1

Period	Proposer	Responder	Confirmation
1	**A**	**A**	**Yes**

Dialogue 3 (*S*1):
P tests R's rationality and R is irrational.

Proposer: Are you rational?
Responder: No, I am not!
Proposer: Then, I'll take advantage of this!

No. of pairs: 1

Period	Proposer	Responder	Confirmation
1	**A**	**B**	**Yes**

No. of pairs: 1

Period	Proposer	Responder	Confirmation
1	A	A	No
2	**A**	**B**	**Yes**

Dialogue 4 (*S*2 & *S*4):
P tests R's rationality and R is rational.

Proposer: Are you rational?
Responder: Yes, I am!
Proposer: Would you cooperate?
Responder: Yes!
Proposer: Then, let's cooperate.

No. of pairs
8

Period	Proposer	Responder	Confirmation
1	A	A	No
2	**B**	**B**	**Yes**

No. of pairs
5

Period	Proposer	Responder	Confirmation
1-2	A	A	No
3	**B**	**B**	**Yes**

No. of pairs
1

Period	Proposer	Responder	Confirmation
1-3	A	A	No
4	**B**	**B**	**Yes**

Dialogue 5 (*S*3-*A* & *S*4):
R tests P's rationality and P is rational.

Proposer: Would you cooperate?
Responder: Wait a minute! Are you rational?
Proposer: I am rational!
Proposer: Would you cooperate?
Responder: Yes!
Proposer: Then, let's cooperate.

No. of pairs
3

Period	Proposer	Responder	Confirmation
1	B	A	No
2	**B**	**B**	**Yes**

Dialogue 5-bis (*S*3-*A*):
R tests P's rationality and P is irrational.

Proposer: Would you cooperate?
Responder: Wait a minute! Are you rational?
Proposer: No, I am not.
Proposer: Then, I'll take advantage of this!

No. of pairs	Period	Proposer	Responder	Confirmation
1	1	B	A	No
	2	**B**	**A**	**Yes**

Dialogue 6 (*S*2, *S*3-A & *S*4) :
P tests R's rationality & R tests P's rationality;
both players are rational and patient.

No. of pairs	Period	Proposer	Responder	Confirmation
4	1	A	A	No
	2	B	A	No
	3	**B**	**B**	**Yes**

No. of pairs	Period	Proposer	Responder	Confirmation
1	1-2	A	A	No
	3	B	A	No
	4	**B**	**B**	**Yes**

No. of pairs	Period	Proposer	Responder	Confirmation
6	1	B	A	No
	2	A	A	No
	3	**B**	**B**	**Yes**

No. of pairs	Period	Proposer	Responder	Confirmation
3	1	A	A	No
	2	B	A	No
	3	A	A	No
	4	**B**	**B**	**Yes**

No of Pairs
1

Period	Proposer	Responder	Confirmation
1-2	A	A	No
3	B	A	No
4-5	A	A	No
6	**B**	**B**	**Yes**

No. of pairs
1

Period	Proposer	Responder	Confirmation
1-2	A	A	No
3	B	A	No
4	A	A	No
5	B	A	No
6	**B**	**B**	**Yes**

No. of pairs
1

Period	Proposer	Responder	Confirmation
1-3	A	A	No
4	B	A	No
5	A	A	No
6	B	A	No
7	**B**	**B**	**Yes**

No. of pairs
1

Period	Proposer	Responder	Confirmation
1	B	A	No
2-4	A	A	No
5	B	A	No
6	A	A	No
7	B	A	No
8	**B**	**B**	**Yes**

No. of pairs
1

Period	Proposer	Responder	Confirmation
1	A	A	No
2	B	A	No
3	A	A	No
4	B	A	No
5	A	A	No
6	B	A	No
7	A	A	No
8	**B**	**B**	**Yes**

No. of pairs: 1

Period	Proposer	Responder	Confirmation
1	A	A	No
2	B	A	No
3	A	A	No
4	B	A	No
5	A	A	No
6	B	A	No
7-9	A	A	No
10	**B**	**B**	**Yes**

Dialogue 6-bis (*S*2, *S*3-*A* & *S*2-*Nash*) :
P tests R's rationality and R tests P's rationality;
P is impatient.

No. of pairs: 1

Period	Proposer	Responder	Confirmation
1	A	A	No
2	**A**	**A**	**Yes**

Dialogue 8:
'Trembling-hand Temptation' (cheap talk):
both players test the other's rationality and willingness to cooperate;
both players are rational and patient.

No. of pairs: 2

Period	Proposer	Responder	Confirmation
1	A	A	No
2	B	B	No
3	B	A	No
4	A	A	No
5	**B**	**B**	**Yes**

No. of pairs: 1

Period	Proposer	Responder	Confirmation
1	B	A	No
2	B	B	No
3	B	A	No
4	A	A	No
5	**B**	**B**	**Yes**

No. of pairs
1

Period	Proposer	Responder	Confirmation
1-2	A	A	No
3	B	A	No
4	B	B	No
5	B	A	No
6	**B**	**B**	**Yes**

No. of pairs
1

Period	Proposer	Responder	Confirmation
1	A	A	No
2	B	B	No
3	B	A	No
4	A	A	No
5	B	A	No
6	**B**	**B**	**Yes**

No. of pairs
1

Period	Proposer	Responder	Confirmation
1	B	A	No
2	A	A	No
3	B	A	No
4	B	B	No
5	B	A	No
6	**B**	**B**	**Yes**