Cooperative Behavior and the Frequency of Social Interaction

John Duffy* and Jack Ochs

Department of Economics
University of Pittsburgh
Pittsburgh, PA 15260 USA

This Draft: July 2008

Abstract: We report results from an experiment that examines play in an indefinitely repeated, two-player Prisoner's Dilemma game. Each experimental session involves N subjects and a sequence of indefinitely repeated games. The main treatment consists of whether agents are matched in fixed pairings or matched randomly in each indefinitely repeated game. Within the random matching treatment, we elicit player's strategies and beliefs or vary the information that players have about their opponents. Contrary to a theoretical possibility suggested by Kandori (1992), a cooperative norm does not emerge in the treatments where players are matched randomly. On the other hand, in the fixed pairings treatment, the evidence suggests that a cooperative norm does emerge as players gain more experience.

***JEL* Codes:** C72, C73, C78, C92, D83.

Keywords: Cooperation, Matching, Repeated Prisoner's Dilemma, Folk Theorem, Community Enforcement, Social Norms, Experimental Design.

∗ Corresponding author. Email: jduffy@pitt.edu, Tel: 412-648-1733 Fax: 412-648-1793.

“Sometimes cooperation emerges where it is least expected.” -Robert Axelrod, *The Evolution of Cooperation* (1984, p. 73).

1. Introduction

Cooperative behavior can expose individuals to possible exploitation by others who are willing to act opportunistically. Nevertheless, cooperation can be sustained if opportunistic behavior triggers a punishment that makes ‘cheating’ unattractive. Much cooperative behavior is sustained by decentralized, informal enforcement mechanisms. These mechanisms rely on individuals having an interest in how their current actions affect future social interactions. As Kandori (1992) notes, there are two general classes of informal mechanisms: Personal enforcement, where opportunistic behavior today destroys the possible benefits of future cooperation between the individual who has been cheated and the cheater; Community enforcement, where a cheater is sanctioned by other members of the community who have not themselves been victims of that cheater, but nevertheless refuse to engage in cooperative endeavors with any cheater. The Folk Theorem holds for personal enforcement when a particular pair of agents has an indefinite number of future interactions with one another and the discount factor is sufficiently large. Kandori shows that public observability is sufficient for the folk theorem to hold for community enforcement mechanisms as well. If the identity of the cheater is common knowledge then it does not make any difference if the cheater will have repeated future interactions with any particular member of the community or if the cheater will simply have an indefinite number of future interactions with various members of the community. That is, *regardless of matching protocol,* with public observability, there exist equilibrium strategy profiles that will support the same payoffs as are attainable under a fixed matching protocol.[1] More remarkably, Kandori shows that for groups of any fixed size there exist payoff

1 Under public observability, an individual who has been labeled a ‘cheater’ has no incentive to act cooperatively in any particular interaction, as long as s/he believes that others will not cooperate with a known ‘cheater’. Non-cheaters, in turn, will have little incentive to risk taking a cooperative action with a known ‘cheater’, if they also believe that others will not cooperate with this particular individual in the future. Kandori provides two different ways of making individual deviations from cooperation unprofitable. One way depends upon individuals having group labels, independent of their actions. Individuals of one group are always matched with individuals of another group and once any member of a given group deviates, all members of that group are treated as ‘cheaters’, by all members of the other group. Alternatively, individuals need not have group labels and only ‘cheaters’ are labeled. To prevent those who have not yet been labeled as a ‘cheater’ from finding it profitable to avoid acquiring this label,

functions such that there is a community enforcement mechanism that will sustain cooperation in an indefinitely repeated number of plays of a Prisoner's Dilemma stage game even when individual histories are purely private, individuals are anonymous, and the matching mechanism for each stage game is purely random.

In the anonymous, random matching case, it is the threat that one deviation from cooperation will trigger a contagious process of future defections by all who have experienced a defection – rather than the threat of being branded a 'cheater' – that acts as a deterrent. For this threat to be credible, an individual who has experienced a defection must find it more profitable to defect at the next opportunity, even though this will keep the contagious process going, than to continue to act cooperatively and stop the process. Of course, if only a small fraction of the population has already been infected, an individual may lose a considerable amount of future benefit from the eventual destruction of the cooperative norm. Therefore, for large groups, the threat of starting a contagious process will not be a credible unless either the one time gain from cheating is very large, or there is some means of stopping the contagious process thereby limiting the loss from the destruction of the social norm. Ellison (1994) shows how the availability of a publicly observable randomization device may be used as a correlation device to signal the end of a punishment phase and resumption of a potentially new cooperative phase.[2]

While Kandori shows that a social norm of cooperation can be sustained as a non-cooperative equilibrium even when individual pairings are both random and anonymous, whether or not such norms are likely to emerge under such conditions is clearly an empirical question. The experiment described below was specifically designed to address that question.

Kandori's (1992) theorem applies to indefinitely repeated two-person games with minimal observation of the past actions of the individuals with whom a player is currently paired to play a stage game. Our experiment is designed, therefore, to study the behavior of individuals drawn from a fixed population who play an indefinite sequence of two-person Prisoner's

---

it must be the case that a non-cheater can expect to meet other non-cheaters sufficiently frequently in the future to make it profitable to give up the one time gain from cheating. This can be assured with strategies in which a cheater's label is removed after some finite number of stage games.

2 As Ellison shows, it is not necessary to have a public randomization device to limit the period in which a contagion process is operating. But such a device does serve as a signal upon which all individuals can coordinate their departure from a punishment phase.

Dilemma games under different matching protocols and different amounts of information transmission. The objective of varying the matching protocol (fixed pairings versus random matching) is to determine empirically how much difference in the level of cooperative play is associated with these different matching protocols. The objective of varying the information transmitted to players (under the random matching protocol only) is to determine whether information on the payoff or action history of a player's opponent, prior to play of the stage game, has any effect on the level of cooperative play. The design also incorporates a randomly generated, publicly observable signal that could be used by agents as a device to coordinate the end of a punishment phase, if they were to choose strategies of the type described by Ellison (1994).

To foreshadow our experimental results, we find that the initial play of subjects is quite similar under both fixed and random matching protocols. With experience, under fixed pairings, cooperation emerges as a norm. However, under random pairings, non-cooperation is quickly established as a norm, despite the presence of some conditional cooperators. Indeed, we provide direct evidence in one random pairings treatment of the presence of such conditional cooperators using a procedure introduced by Fischbacher and Gächter (2006) that builds upon the strategy method for identifying player types introduced by Fischbacher, Gächter and Fehr (2001). Under random pairings, there is also evidence that some individuals attempt to coax members of their group to break out of a non-cooperative norm, as Ellison suggests. However, under random pairings, efforts to establish a social norm of cooperation prove futile, despite the presence of a number of conditional cooperators. We further find that a social norm of cooperation that is established under fixed pairings will be immediately broken when the matching protocol is switched to random pairings, while a social norm of non-cooperation under random pairings is also easily displaced when the same group is switched to fixed pairings. Finally, we find little evidence that smaller group sizes or the provision of information about an opponent's payoff history or past actions increases the likelihood of a cooperative social norm developing under a random matching protocol.

2. Related Work

2.1 Indefinitely Repeated Stage Games

By definition, indefinitely repeated stage games have no (predictable) last stage. Therefore, if cooperative play is reciprocated in the early stages, a belief in future reciprocity will be reinforced and cooperative behavior may be sustained indefinitely. The concept of an indefinitely repeated stage game is implemented experimentally by use of a randomization device to determine after each stage game is played whether the game has ended or another stage game is to be played. The probability of continuation determines whether or not there exists a cooperative equilibrium in the supergame. This device was introduced by Roth and Murnighan (1978).[3] There are surprisingly few experiments that have been conducted with indefinitely repeated stage games.[4] Van Huyck et al. (2002) report an experiment conducted with supergames constructed of an indefinite sequence of repetitions of dominance solvable stage games, followed by a small fixed number of repetitions of the same stage game. They observed that during the probabilistic continuation phase of a supergame whose stage game has an equilibrium in strictly dominant strategies, the cooperation level rose dramatically with experience. In the Van Huyck et al experiment, all observations were made under a fixed pairings protocol.[5] To our knowledge, the only other experiment with an indefinitely repeated game played under different matching protocols is reported in Palfrey and Rosenthal (1994). They conducted an experiment using an indefinitely repeated N-person provision point voluntary

3 In their experiments subjects played an indefinitely repeated Prisoners' Dilemma stage game against a pre-programmed strategy, either tit-for tat, or grim response. They were interested in testing how responsive cooperative play was to variations in the continuation probability. While they found the rate of cooperative play to be positively related to the continuation probability, the levels of cooperation they observed were quite far from 100%.

4 Dal Bó (2005) considers whether the responsiveness of cooperation to an increase in the continuation probability observed by Roth and Murnighan is simply a reflection of the increase in the expected number of repetitions of the stage game before a relationship is terminated, or whether behavior in games of indefinite length is fundamentally different from behavior in games of finite repetitions of a prisoners' dilemma stage game. He finds that the percentage of cooperative play in finitely repeated games of a given length is lower than indefinitely repeated games of the same expected length.

5 Holt (1985) also reports an experiment with an indefinitely repeated duopoly game conducted under fixed pairings. Like Van Huyck et al, a basic treatment variable was whether the game was an indefinitely repeated supergame or a finitely repeated game. In the supergame treatment, the median of the distribution of final period outputs was much closer to the Nash equilibrium than to a collusive, joint profit maximizing output. Aoyagi and Frechette (2005) also report an experiment conducted with supergames constructed from an indefinite sequence of a prisoner's dilemma game with noisy public signals. Their subjects also played under fixed pairings. They found that cooperation increases as the noise in the signal decreases.

contribution game. This game was played under both fixed and anonymous random matching protocols. Unlike a Prisoners' Dilemma, non-cooperation is not a dominant strategy in their stage game and the stage game has a multiplicity of cooperative equilibria. They found that under random matching, subjects "adhere to cut-point decision rules that are, on average, very close to those predicted by the Bayesian equilibrium (of the one-shot game). Repetition (i.e., fixed matching) leads to more cooperative behavior (than observed with random matching)...(but) the observed magnitudes of improvement are much smaller than predicted (assuming that random matching corresponds to the play of a one-shot game)." The fact that subjects did not come close to fully exploiting the opportunities for coordination and cooperation under either the fixed or random matching protocols in the Palfrey-Rosenthal experiment is, perhaps, not surprising. Their stage game is a game of incomplete information. The symmetric cut-point strategy that maximizes expected joint profits is not transparent. That game has a multiplicity of equilibria in non-symmetric pure strategies and there is no evidence of any effort to coordinate on a pure strategy equilibrium. By using a Prisoner's Dilemma as the stage game in our own experiment we expect the cooperative equilibrium of the supergame to be much more salient.

2.2 Information

The Kandori theorem implies that a cooperative norm can develop even in the absence of any information being transmitted about one player's past actions or experience to the other player with whom he is currently paired. Nevertheless, one might expect players to act differently if such information is transmitted than if it is not. Bolton, Katok and Ockenfels (2005) report an experiment with a finitely repeated stage game in which each player may carry with him an *image score* that reflects some information about the past experience of that player. This information, but not the identity of the player, is observable to the other person with whom s/he is matched in the current stage game. At the beginning of each stage game, individuals are randomly paired and then a random draw determines the choice of 'dictator'. The dictator can either 'Give' the other player ('receiver') a large payoff and receive no payoff himself, or 'Take' a small payoff himself and give the receiver nothing. The issue they explore is whether a concern for one's future image influences the dictator's current behavior. They find that when the opportunity cost of being nice is high, 'giving' is much lower when there is no image score than when there is an image score. Because their stage game is finitely repeated, Kandori's theorem does not directly apply. Indeed, in their game, the only sequentially perfect equilibrium is to 'Take'. Nevertheless, this experiment suggests that information transmission may make a difference in the play of a game under a random matching protocol.

Another related experiment was conducted by Schwartz, Young and Zvinakis (2000). They use a modified Prisoner's Dilemma as the stage game in their experimental design. Their stage game is played for an indefinite number of times under a random matching protocol. Subjects remain anonymous. However, under different treatments, different portions of a player's past history are revealed to the person with whom s/he is currently matched. They find that these disclosure conditions have a large effect on the initial levels of cooperation observed. However, under either information condition, they observe a decline in cooperation as subjects gain experience. Like Bolton, et al., the results of Schwartz and his associates indicate that among inexperienced subjects, in environments where cooperation can only be reciprocated indirectly, information transmission can have significant effects on behavior. What is left open is whether these effects can sustain a cooperative equilibrium as subjects gain experience. Conversely, it is still unknown whether anonymous random matching with no information transmission about one's current partner's history will reliably produce a non-cooperative equilibrium.[6]

3. The Experiment

3.1 Experimental Design

In all sessions of our experiment we use an indefinite repetition of the Prisoner's Dilemma stage game shown in Figure 1.

---

6 Andreoni and Croson (2002) assess the effect of random re-matching (strangers) versus fixed matchings (partners) in the context of finitely repeated, N>2-player voluntary public good game experiments. Examining several different studies, they find inconsistent evidence on the effect that random (fixed) matching has on the level of contributions to a public good. By contrast, in this paper, we focus on two-player, indefinitely repeated prisoner dilemma games under fixed and random matching protocols, and we vary the level of information in the case of random pairings.

		Column Player	
		X	Y
Row	X	20,20	0, 30
Player	Y	30, 0	10,10

Figure 1: The stage game

The infinite horizon supergame was constructed as follows. Following play of the stage game, a random draw was made from a uniform distribution over the range [1, 100]. The draw was made by the computer program that was used to carry out the experiment (students made their choices and observed the outcomes on networked computer workstations) and the number chosen was displayed in a pop-up box on all player's computer screens to reinforce the random nature of the draws. If the draw was less than or equal to 90, players were matched according to the given protocol and the stage game was repeated. If the random draw exceeded 90 the supergame was ended. Thus, the probability, p, that a supergame continues is .90 and the expected number of future rounds to be played from the perspective of any round reached is always 1/(1-p) or 10. This is equivalent to an infinite horizon where the discount factor attached to future payoffs is .90 per round. Once a supergame ended, depending on the time available, another supergame would begin with the same stage game, matching protocol and population of players used in all previous supergames of the experimental session.[7]

Note that the random draw is useful not only for implementing an indefinite horizon. It can also be used as a publicly observable randomization device enabling players to implement the kinds of strategies found in Ellison (1994), where the randomization device is used to coordinate a halt to a contagious defection phase of play and resumption of a potentially new cooperative phase of play of the repeated game.

The stage game payoffs and the discount factor were chosen such that for the population size of players we consider, $N \leq 14$, there exists a perfect, sequential equilibrium that supports

7 Despite our instructions, subjects may have conceived of the sequence of indefinitely repeated games actually played as one single repeated game. If this were the case, the relevant discount factor would be greater than the induced value of .9. Nevertheless, the contagious equilibrium would continue to exist in this case. As we shall see later, there is evidence of a substantial spike in cooperation frequencies at the start of each new supergame indicating that subjects did conceive of each indefinitely repeated game as a separate game.

perfect cooperation even if there is anonymous random matching after each stage game and all information about an individual's prior history is strictly private. Under the same parameters, these games also have perfect, sequential equilibria that support perfect coordination when some information about an individual's history is transmitted to the individual with whom that person is currently matched. Appendix A provides further details. Given that the expected length of a supergame is 10 repetitions of the stage game, subjects have experience with several supergames over the course of a two-hour session.

There are three treatment variables in our main design. The first treatment variable is the matching protocol (fixed pairings; random pairings). The second treatment variable is the size of the population (N=14 or N=6). The third treatment variable is the information conveyed to each member of a pair playing a stage game regarding the history of the other member of the pair. Since, in fixed pairings each member of a pair shares a complete history with the other member and is aware of this fact, the information treatment is varied only in the random matching protocol sessions.

3.2 Hypotheses

The basic hypothesis to be tested is that there is no significant difference in the level of cooperative play observed under anonymous-fixed and anonymous-random matching protocols. In a given session, the matching protocol is made public through the instructions that are read out loud. In most sessions, the protocol does not change during the course of a session. Subjects are either assigned to a fixed pairing at the beginning of a supergame or are randomly paired after each stage of a supergame. In sessions in which a fixed pairing protocol is used, the fixed pairings changed from one supergame to the next in a round robin format; prior to the first round of each new supergame, each player was anonymously matched with one of the N-1 players with whom s/he had not previously played a supergame.[8] In some sessions, one matching protocol is

8 Hence, for the fixed-pairings treatment, up to N-1 supergames could be played where in each supergame, the players in a anonymous fixed pairing have not previously met one another. This consideration motivated our choice of N=14, as we typically played no more than 13 supergames in a session (see Table 1). If more than N-1 supergames were played in a session, (as in sessions where N=6), then players were matched with players with whom they had played before. However, as these matchings were anonymous, players could not condition on their past history of play with any other player.

used at the beginning of a session and then the second protocol is used for the remainder of the session. In those sessions in which two protocols are used, subjects are not informed of the change in protocol until the point in the session at which the switch is made. This treatment allows us to observe how a given group of subjects responds to a change in matching protocol.

A second hypothesis to be tested is that in an anonymous random matching environment, the relative frequency of cooperative play is unaffected by the amount of information about an individual's own history that is available to the person with whom that individual is matched. A competing hypothesis is that the more information an individual has available to label a player a non-cooperator, the greater likelihood that individuals will refrain from non-cooperative play in the random matching environment. The amount of information about one player's history that is transmitted to the other player is a second treatment variable in our design. This variable, I, can take on one of three values: 0 (no information is transmitted); 1 (the average payoff of the two players in the individual's last stage game is transmitted); 2 (the action chosen by the individual last period is transmitted). Under all conditions, the matching and information transmission are done so as to preserve the anonymity of each person. When $I = 0$, each player can only condition his/her own strategy on his/her own history. Only an individual who has actually experienced non-cooperative play has any reason to update his/her own priors about the relative frequency of playing future games with another individual who has had the same experience. This is the case considered by Kandori (1992). When $I > 0$, each player can condition his action on not only her/his own history, but also on the information provided about his opponent's history. Furthermore, when $I>0$, each player knows that the player with whom they will be matched next period will possess information that *may* (when $I= 1$) or *will* ($I = 2$) be sufficient to label her/him a 'non-cooperator' if s/he chooses to defect this period. Intuitively, giving players information on other players' histories prior to the play of a stage game should serve to speed up the onset of the contagious equilibrium thereby strengthening the threat by which a norm of cooperative behavior is sustained. On the other hand, Kandori's theorem does not require that players posses such information on their opponent's immediate past history of play. Hence our null hypothesis is that the frequency of cooperative play is unaffected by either $I>0$ treatment.

3.3 Results

We report results from 28 experimental sessions involving a total of 344 subjects. A description of the characteristics of these experimental sessions is given in Table 1. In most sessions we used a population size of N=14. In six sessions we considered a smaller population of size N=6 as a robustness check on our results with the larger population size.

[Insert Table 1 here]

Our aim was to get approximately 100 rounds of data per session. As the length of each indefinitely repeated game (supergame) should average 10 rounds, our goal of 100 rounds per session was satisfied by playing an average of 10 indefinitely repeated games per session. Of course, due to the random end of each indefinitely repeated game, there is some variation in the number of games and rounds as indicated in Table 1. Subjects were not told of our objective of 100 rounds, nor were they told in advance which indefinitely repeated game would be the last one played. Subjects were recruited for a two-hour session but our goal of 100 rounds was always achieved well before this two-hour limit, typically after around 90 minutes.

The subjects were recruited from the undergraduate population at the University of Pittsburgh. Each group of subjects had no prior experience participating in any treatment of our experiment. Subjects were read instructions pertaining to the single treatment they were participating in and then began playing the Prisoner's Dilemma game shown in Figure 1, entering their choices, X or Y, on a computer screen when prompted. All of our treatments involved the same parameterization of the stage game shown in Figure 1. Copies of the instructions used in the fixed and the random pairings (I=0) treatments are included in Appendix B. Following their choice of action, X or Y, subjects were informed of the other player's action and their payoff. The payoff numbers for the game, as shown in Figure 1 were interpreted as monetary payoffs in terms of cents (US$). Thus, if two players chose Y,Y in a round, each player earned 10 cents, etc. Subjects were paid their payoffs from all rounds of all games played and in addition were given a show-up payment of $5. Average total earnings depended on the treatment. In the fixed pairings treatment, subject's total earnings (including the $5 showup fee) averaged $18.64. In the random pairings treatment (I=0, N=14) subjects' total earnings averaged $14.86.

3.3.1 Fixed Versus Random Pairings with No Information, 14 Subjects

[Insert Figure 2 here]

The left column of Figure 2 presents data on the aggregate frequency of cooperation in each round of each game played in four sessions that were conducted under a fixed pairings matching protocol with 14 subjects. The horizontal axis reports round numbers. A round number of 1, represented in Figure 2 (and subsequent figures) by a vertical bar, indicates the start of a new supergame. The right column of Figure 2 presents the aggregate frequency of cooperation in each round of each game played in the four sessions that were conducted under the random pairings matching protocol when subjects received no information (I=0) regarding the past experience of anyone with whom they were currently matched. While subjects who played under the random pairings protocol were randomly paired after each round of play, the procedure was to terminate a sequence of rounds with the same stopping rule as was used in the fixed pairings matching protocol sessions. When a sequence ended, the end of the 'game' was announced. If our criteria of obtaining approximately 100 rounds of play had not yet been reached, we announced that a new game would begin. Therefore, a round number of 1 on these graphs also indicates when a new sequence of rounds was begun. In the graphs shown in Figure 2, we report both the aggregate frequency of cooperation—% choice of action X— together with a fitted line from a regression of %X on a constant and time, t,=1,2…T (where T is the total number of rounds played in all supergames of the session). A tabular display of the aggregate frequencies of cooperation for the fixed and random, I=0 matching sessions with 14 subjects is presented in Table 2.

[Insert Table 2 here]

The column in Table 2 labeled "Game 1, Round 1" reports the aggregate frequency of cooperative play (i.e., choice of X) in the first round of the first game played in each session involving 14 subjects. According to nonparametric, robust rank-order tests[9], there is no

9 See Siegel and Castellan (1988) or Feltovich (2003) for a discussion of the robust rank order test. This test is used

significant difference ($p>.10$) in the distribution of these Game 1, Round 1 cooperation frequencies between the fixed and random, I=0 treatments. Thus, the difference between the fixed and random matching protocols is not immediately taken into account by subjects.

While there is no difference in the way inexperienced subjects first play these games, experience under the fixed pairings protocol drives each group of subjects to a much higher level of cooperative play than is observed under random pairings. Indeed, under random pairings, as subjects gain experience, the frequency of cooperation plummets towards zero. By contrast, under fixed pairings, as a session progresses, the frequency of cooperative play increases. As subjects gain experience, the difference in cooperative frequencies between fixed and the random pairing treatments increases.

More precisely, robust rank-order tests of the null hypothesis of no difference in cooperation rates between treatments confirm that the aggregate cooperation frequencies over the first half, over the second half, and over all rounds of a session (as reported in Table 2) are significantly higher in the fixed pairings treatment than in the random pairings (I=0) treatment ($p = .014$, smallest critical value for paired samples with 4 observations each). Furthermore, in the fixed pairings treatment, the cooperation frequencies in the second half of the sessions are significantly higher than those in the first half ($p=.014$). By contrast, in the Random I=0 treatment, the cooperation frequencies in the second half of the sessions are marginally lower than those in the first half ($p=.10$).

An interesting property of the data in all sessions is the increase in cooperation observed in the first round of many of the supergames relative to the level of cooperation in the final rounds of the preceding supergame. This ‘restart’ phenomenon shows up in all of our treatments as revealed in Figure 2 (but see also Figures 6-9). It is clearly illustrated in Figure 3a, which shows the aggregate frequency of cooperation in the first round of all supergames as well as the aggregate frequency of cooperation in all other supergame rounds, excluding the first round across all fixed or random I=0 sessions with 14 subjects. The figure reveals that on average, cooperation is greater in the first round than over all subsequent rounds of each supergame.

[Insert Figures 3a-3b here]

throughout this paper, wherever p-values are reported in the text.

In the random pairings treatment, the restart effect reflects repeated efforts by just a few subjects to encourage a social norm of cooperation. Some evidence for the existence of heterogeneous subject types is given in Table 3, which reports the cumulative number (cumulative percent) of players in each of the fixed or random I=0 pairings sessions involving 14 subjects whose individual frequencies of cooperation fell below various threshold levels, using data from all rounds of all supergames of a session, as well as for the first and second halves of a session. For instance, in Session 2 of the Random, I=0 pairings treatment, there were 2 subjects who cooperated (chose action X) in 10 to 25 percent of all rounds played, while in session 4 of this same treatment, there were 3 subjects who cooperated in 25 to 50 percent of all rounds played. However most of these cooperating-types are cooperating with low frequency, typically in the first round of each new supergame. As a random-pairings session continues, some of these first-round cooperators get discouraged, and shift to defecting in the first round. This has the effect of dampening out the restart effect and reducing first-round cooperation frequencies in the random matching (I=0) treatment as shown in Figure 3b.[10]

To see that this is the case, let us (arbitrarily) label a player who cooperates more than 10 percent of the time in the first half of a Random I=0 session a "hopeful" player. For instance, in session 1 of the Random, I=0 treatment, Table 3 reveals that there is exactly 1 hopeful player; in session 2 there are 3 hopeful players, in session 3, there are 6 hopeful players and in session 4 there are 12 hopeful players. In all 8 of these sessions, Table 3 reveals that the number of hopeful players always declines from the first to the second half of the session.

This reduction in the number of hopeful players tends to dampen out the 'restart' phenomenon as additional supergames are played, as the hopeful players are mainly cooperating in the first round of a supergame. This dampening out of the restart phenomenon in the random matchings treatments is illustrated in Figure 3b, which shows the aggregate frequency of cooperation in the first rounds of supergame numbers 1-10 using pooled data from all sessions of a treatment.[11]

---

10 A similar dampening out of the restart effect occurs in the random pairings treatments with I=1 or I=2.
11 As noted above in the discussion of Table 1, some sessions had more than 10 supergames, and some had less. In Figure 3b, we have reported the average frequency of cooperation in supergame number 1,2,…10 for all sessions of a treatment for which that supergame was actually played.

[Insert Table 3 here]

In the fixed pairings treatment, the *aggregate* level of cooperation within a given sequence of rounds (supergame) tends to diminish as the number of rounds played in that sequence increases, as can be seen in Figure 2 or in the aggregate frequencies shown in Figure 3a. The decline in the aggregate frequency of cooperation over time is due to the presence of just a few players, who very frequently chose to defect, despite being in the fixed pairings treatment. The presence of these defecting players can again be seen in Table 3. For instance, in Sessions 1, 2 and 3 of the fixed pairings sessions, we see that there are always 1 or 2 individuals who were choosing action X (cooperating) in less than 10 percent of all rounds played, (defecting more than 90 percent of the time). As in the random pairings treatment, there is a "restart" phenomenon where the aggregate level of cooperation increases at the beginning of a new sequence with new pairings, from the level observed at the end of the previous sequence. Unlike the random pairings treatment, there is an upward trend in the aggregate level of cooperation observed the first time new pairings interact, in the first round of each supergame – see Figure 3b.

[Insert Figure 4 here]

As Figure 4 makes clear, *on average*, the aggregate frequency of cooperation is a little more than 10 percent *lower* at the end of each supergame relative to the start of that supergame. The reason for this finding is that in each fixed pairing session there is typically a small core of players – 'defectors' – who defect with a high frequency as can be seen in Table 3. In the first rounds of play of a new supergame, these defectors' impact on the aggregate frequency of cooperation is at its weakest. However, if the defectors are in *fixed pairings* with subjects playing conditionally cooperative strategies, these conditional cooperators will quickly switch from cooperating to defecting, thereby *further lowering* the aggregate frequency of cooperation as the supergame proceeds. Nevertheless, the upward trend in the frequency of cooperation in the first round of each new supergame is sufficiently strong that the aggregate frequency of cooperation increases over time.

This upward trend in first-round cooperation under fixed pairings – as shown in Figure 3b- is due to a reduction in the number of 'near-unconditional' defectors as a session proceeds. To see that this is the case, let us (again, arbitrarily) label a player who cooperates less than 10 percent of the time in the first half of a fixed pairing session a "pessimistic" player. As Table 3 reveals, in two of the four fixed pairing sessions, (numbers 2 and 4), the number of these pessimists drops from 1 or 2 in the first half of the session to 0 in the second half of the session. If pessimists were alternatively defined as those who cooperated less than 50% of the time, a starker drop-off in the number of pessimists would be found from the first to the second half of *all* four fixed pairings sessions.[12]

We conclude that, under fixed pairings there appears to develop a social norm of cooperation as a given group of subjects gains experience, while under the random, I=0 pairing treatment, experience tends to drive groups toward a far more competitive norm.

3.3.2 Further evidence of heterogeneity and player 'types'

While we have provided some evidence for heterogeneity in player types, the evidence has been obtained using data on *observed actions*. At the suggestion of referees, we have conducted a modified version of our random, I=0 pairings treatment that allows us to more carefully identify player types and the extent of heterogeneity among player types in this treatment. In this modified treatment, in addition to observing subjects' actions we also elicit their *strategies and beliefs* regarding the play of other 13 subjects in the room, with the aim of characterizing each player as either a conditional cooperator or an unconditional defector. To do so, we adopt a design proposed by Fischbacher et al. (2001) and Fischbacher and Gächter (2006) aimed at measuring the extent of "social" (or "other- regarding") preferences in finitely repeated public good games.[13]

---

12 We have verified that this drop-off in the number of pessimists is due to changes in the actions chosen by the players labeled as pessimists in the first half of the session. Players who were *not* labeled as pessimists in the first half of a session are almost never labeled as pessimists in the second half of a session.

13 In the indefinitely repeated games that we study, conditionally cooperative play need not be indicative of "social" preferences; rather such behavior is fully consistent with standard, self-interested preferences.

Specifically, in this new treatment, referred to in Table 1 as the "one shot then Random I=0" treatment, we divide our random, I=0 pairing treatment into two parts. In the first part, the 14 subjects are presented with the Prisoner's Dilemma game shown in Figure 1. They are then asked to provide their *strategy* for playing that game *once* against a randomly chosen opponent. Specially, they are asked: "what is the smallest number of the other 13 people with whom you might be matched who must choose X before you would choose to play X?" They are further instructed that this number can be any integer between 0 and 13 inclusive and that "Never choose X" is also a choice; under the latter choice, they would *never* choose X (cooperate) regardless of the number of others who chose X – a strategy of unconditional defection. *After* specifying a strategy, subjects were asked to provide a *forecast* of the number of other 13 subjects who will play X. Finally, they were instructed to *state* their action choice (X or Y) for the one-shot game. The 14 subjects were then randomly matched. In each pair, one randomly chosen member had his strategy played for him using the actual number of the other 13 subjects who stated that they would play X. The action played by the other member of the pair was his stated action for the one-shot game.[14] Subjects were then informed of the outcome of the one-shot game, learning whether the action they chose was as they had stated or was determined using their strategy, the action chosen by their opponent and their payoff in points (0, 10, 20 or 30 points). They were further informed of the number of the other 13 subjects who chose X and were awarded an additional 10 points if their forecast of the behavior of the other 13 subjects was correct (0 points otherwise).[15]

In the second part of this new treatment (which was not revealed in advance), subjects played a number of indefinitely repeated supergames under the same random matching I=0 conditions that we examined earlier, but with one change. Prior to making their decision in every round of every game, they had to predict the number of the other 13 subjects who they thought would choose X in the round they are about to play (an integer between 0 and 13

---

14 This design, which was made clear to subjects in advance in the written instructions, insures that they have incentives to specify strategies that best characterize the actions they would choose given their beliefs (forecasts). Instructions for this treatment are available in the supplementary materials that accompany this paper or from the authors.

15 Thus, at the start of the second half of the session, subjects in this treatment were informed of the frequency of cooperation by other subjects in the one-shot game.

inclusive). Subjects were further instructed that "since the person with whom you are matched is selected randomly, your prediction reflects your assessment of the likelihood that you will be matched with someone who plays X." To incentivize subjects to provide accurate forecasts, we told them we would choose one round of one supergame at random at the end of the session and pay them 100 points if their pre-play forecast of the behavior of the other 13 subjects was correct in that round (and 0 points otherwise). As in our prior sessions, subjects also received their payoff in points from all rounds of all supergames played in the second part along with their first part earnings. The conversion rate for points awarded in both parts of this treatment was the same, 1 point equaled 1 cent.[16]

[Insert Figure 5 here]

[Insert Table 4 here]

Figure 5 shows the cumulative frequency of player "types" in each session of this treatment, where a player's type is defined by his cut-off strategy – the number of the other 13 players who would have to play X in order for the player to play X (cooperate). The figure reveals that with a single exception (in session 3), no subject specified a threshold of 0, indicating unconditional cooperation. However, in each session there is a substantial fraction of *unconditional defectors* in the one-shot game – those specifying a threshold of 14 (which corresponds to the strategy "Never choose X" (cooperate)) in Figure 5 range from 21.4% (3 out of 14 subjects in Session 3) to 50% (7 out of 14 subjects in Session 1). The remaining subjects, between 50% in Session 1 and 78.6% in Session 3 may be labeled as *conditional cooperators*, who cooperate provided that some number $0 \leq k \leq 13$ of the other 13 subjects cooperate. The mean and median threshold among all conditional cooperators is 8. If we also include the unconditional defectors (threshold of 14) the mean threshold rises to 10 and the median threshold rises to 11. A two-sided Kolmogorov-Smirnov test confirms the impression given by Figure 5 that there are no significant differences between any pair of cumulative frequency distributions (p>.10 in all pairwise comparisons), suggesting that our subject samples come from the same population distribution. Our findings regarding the distribution of player types is similar to that of Fischbacher and Gächter (2006) who report that 55 percent of subjects in their finitely

16 We did not attempt a strategy elicitation treatment under the fixed matching protocol as the space of possible strategies under that matching protocol is too large to characterize in a simple fashion.

repeated public good game may be classified as conditional cooperators while 22.9 percent are "free-riders" (unconditional defectors).

Table 4 shows aggregate cooperation frequencies in the first and second parts of the three sessions of this 'one shot then random, I=0 pairings' treatment. Cooperation frequencies in the second part of this treatment are further disaggregated according to player types, as determined again by the cut-off strategy elicited from subjects in the first part of the session for play of the one-shot game. Subjects were characterized as "conditional cooperators" if their strategy specified a cut-off level of 13 or less. Otherwise they were characterized as unconditional defectors. The numbers of each type of subject in each session are reported in left-most column of Table 4.

Notice first in Table 4 that session-level cooperation frequencies by *all* subjects in the initial one-shot game (average .262) are somewhat lower than is observed in Game 1, Round 1 of the indefinitely repeated random matching I=0 treatment with 14 subjects (average .429) as reported in Table 2, but this difference is not significant (p>.10 using a rank-order test). Similarly, cooperation frequencies by all subjects in the second part of this new treatment (all games all rounds, average .043) is also a little lower than found in all games, all rounds of the indefinitely repeated random matching I=0 treatment with 14 subjects (average .075) as reported in Table 2, but again this difference is not significant (p>.10). The more striking finding of Table 4 is that cooperation rates by conditional cooperators are typically 3 to 5 times greater than the corresponding cooperation rates of unconditional defectors within the same session. This evidence serves to confirm our earlier intuition that heterogeneity among subjects accounts for both the restart effect and the non-convergence of cooperation levels to zero in the random-matching I=0 treatment, but it also makes clear that a large fraction of subjects would stand ready to cooperate if only the aggregate cooperation frequencies were sufficiently high. Finally, the last column of Table 4 reports the average accuracy of subjects' first part strategy and second-part beliefs in predicting their actual second part behavior.[17] While there is no reason to

17 This measure is constructed as follows. For each subject we took the cut-off strategy we elicited from them in the first part of the session and combined that with their forecast (belief) prior to the play of each round in the second part of the session regarding the number of the other 13 players who would choose X in that round. We counted the number of times subjects' strategy and forecast accurately predicted their actual play and used these numbers to calculate the average accuracy measures reported in the final column of Table 4.

believe that subjects' strategies for playing a one-shot game should predict their behavior in the second part of these sessions which involve a sequence of indefinitely repeated games, the accuracy of subjects' one-shot strategies in predicting their play in the indefinitely repeated game is quite high averaging in excess of 75 percent in all sessions.[18] Notice further that those specifying a strategy of unconditional defection in the initial one-shot game nearly always defect in the second part of the session (on average 98.6 % of the time) whereas those who specify conditionally cooperative strategies in the initial one-shot game are somewhat less likely to follow those same cut-off strategies in the second part of the session, though consistency remains high (on average 84.3%). Taken together, this evidence suggests that subjects may approach play of an indefinite sequence of prisoner dilemma games under random matching in the same way that they would play a one-shot version of that game. In other words, subjects behave as though they understand that random matching will work to frustrate opportunities for collusive outcomes even in an indefinitely repeated game with a finite population size.

3.3.3 The Effect of Group Size

A group size of 14 is, theoretically, sufficiently small for the existence of a cooperative equilibrium under random matching with no information transmission (I=0). Indeed, as detailed in Appendix A, our parameterization of the indefinitely repeated Prisoner's Dilemma game admits a cooperative equilibrium under random pairings and no information for any group of size 2-30. However, the threat of setting off a contagion process does not appear to be sufficient to sustain cooperation in random matching environments with a group of size 14. Figure 6 below displays the results observed in sessions in which a smaller group of 6 subjects were matched either in fixed pairings for the duration of a supergame or randomly in each round of a supergame with no information about their opponent's prior history of play – 3 sessions of each treatment. In the experimental sessions with groups of 6 subjects, we followed the same experimental procedures as in the sessions with 14 subjects. With a smaller group size, a contagion process will get back to its originator much more quickly and the threat of setting off

18 Of course, we cannot rule out the possibility that the high degree of consistency observed between the actions subjects took and the actions that their elicited strategies indicated they would take, conditional on their forecasts, may be an artifact of the design in which they are first asked to articulate their strategies.

such a process should provide a correspondingly larger incentive to cooperate. As the data in Figure 6 reveal, when there is no information feedback, under random matching the smaller groups behave as competitively as the larger groups.

[Insert Figure 6 here]

Table 5 gives the aggregate frequencies of cooperation in the eight sessions with 6 subjects.

[Insert Table 5 here]

As in the sessions with 14 subjects, robust rank-order tests reveal that the null hypothesis of no difference in the distribution of game 1, round 1 cooperation frequencies between the fixed and random pairings I=0 treatments cannot be rejected ($p>.10$). Under the fixed matching protocol, the aggregate frequency of cooperation increases with experience in all three sessions with 6 subjects, while under the random matching protocol the aggregate frequency of cooperation diminishes with experience in two of the three sessions.[19] Rank-order tests further confirm that the aggregate cooperation frequencies over the first half, over the second half, and over all rounds of a session (as reported in Table 5) are significantly higher in the fixed pairings treatment than in the random, I=0 treatment. ($p\leq .029$). A comparison of the aggregate cooperation frequencies (over the first half, second half, or all rounds of a session) achieved by groups of 14 subjects in the random, I=0 treatment with those achieved by groups of 6 subjects in the same treatment (cf. Tables 2 and 5) yields no significant differences ($p>.10$). Similarly, a comparison of the cooperation frequencies achieved by groups of size 14 or 6 under the fixed pairings protocol also yields no significant differences. We conclude that group size has no statistically significant effect on aggregate cooperation rates.

3.3.4 The Effect of Prior Conditioning

19 In session #3 of the random, I=0 treatment with 6 subjects, there was a slight increase the aggregate frequency of cooperation over time-see the bottom right panel of Figure 6; the slope of the fitted line for this session is positive, though not significantly different from zero. This slight upward trend is owing to the increase, over time, in cooperation frequencies in the first few rounds of each new supergame. However, there continues to be a dramatic fall-off in the cooperation over the course of each supergame, contrary to the findings in the fixed pairings treatment.

A group of subjects who gain experience with the fixed pairings protocol tends to exhibit a high degree of cooperation. It is natural to ask whether the social norm of cooperation such a group had exhibited under fixed pairings will be sustained when the group is switched to a random matching protocol. Conversely, if a group has exhibited a social norm of non-cooperation under a random matching protocol will that experience inhibit the formation of a cooperative norm if they are switched to a fixed pairings protocol? To study the effect of prior conditioning on the nature of the social norm developed under a given matching protocol we conducted four sessions in which subjects were first matched under one protocol and then, sometime during the middle of each experimental session, they were switched to another matching protocol. This type of design is referred to as a "within-subjects" design and stands in contrast to the "between-subjects" design we have used up to now.[20] The switch in matching protocols was *not* announced in advance. When the switch was made, we handed out and read aloud a brief change in the instructions, which explained to subjects the new matching protocol that would be in effect in all subsequent rounds. We then played several supergames under this new protocol. All other procedures were as before.

Figure 7 shows data on cooperation frequencies from the four within-subjects sessions we conducted with 14 inexperienced subjects per session. The left column of Figure 7 displays data from two sessions in which subjects were first matched according to the fixed pairings protocol and then, without prior announcement switched to a random pairings protocol in the manner describe above. The right column of Figure 7 displays data from two sessions with the opposite order of use of protocol.

[Insert Figure 7 here]

When subjects are first matched under fixed pairings, they quickly achieve a high level of cooperation. However, the switch to the random matching protocol produces an immediate, dramatic decline in the rate of cooperation and, as the session continues, the rate of cooperation

20 Within-subject designs yield findings that are less susceptible to individual differences than between-subject designs, e.g., a subject who had too little sleep before an experimental session is nevertheless present in all treatments of a within-subjects design, but is only in one treatment of a between-subjects design. On the other hand, between-subject designs allow subjects to acquire more experience with a particular treatment than within-subject design sessions of the same duration.

quickly tends to zero. In short, there is no evidence that a group of people who have learned to cooperate under fixed pairings will develop a social norm of cooperation that persists when matched randomly. Conversely, as shown in the right column of Figure 7, experience with random matching that has led members of a group to behave competitively does not prevent the group from immediately making a marked increase in the cooperation rates in response to a switch to the fixed pairing protocol and, with experience, achieving very high sustained levels of cooperation. Indeed, the data suggest that a group that has experienced the competitive outcomes under random matching may learn to cooperate under fixed pairings even more rapidly than groups who have not had such experience.

3.3.5 The Effect of Information Transmission

In the case of the random pairings treatment with 14 subjects, we have also considered variation in the amount of information that players have regarding their opponents relative to our baseline I=0 treatment, where players have no such information. In the random pairings, I=1 treatment, prior to play of the stage game, both players are informed of the average payoff earned in the last two-player stage game played by their opponent. There are just three possibilities for this average payoff: 10, 15 or 20. If the report is 10 (20), then it is known that the opponent played Y (X) in the last stage game. If the report is 15, then it is known that either the opponent, or his matched pair, but not both, played Y in the last period. In this information treatment, a player who was seeking to signal to a future opponent her determination (say) to play the cooperative action X by choosing X this period would be unable to do so *unambiguously*. Alternatively, this information treatment can be viewed as a particular form of imperfect monitoring.

In the random pairings, I=2 treatment, prior to play of the stage game, players are perfectly informed of the action (X or Y) that their opponent chose in the previous round of play when matched with another player. This is a different kind of information than is given in the I=1 treatment; in the I=2 treatment, there is *no ambiguity* about the action chosen by a player’s opponent in the previous period. Providing unambiguous information on an opponent’s action choice prior to play of the stage game as in the I=2 treatment makes it straightforward to label an opponent as a cooperator/defector *before* playing the game. On the other hand, the payoff information in the I=1 treatment reveals whether the opponent *or his partner* defected in the last round and therefore provides more information about whether a contagious process has started than does the information provided in the I=2 treatment.[21]

[Insert Figures 8-9 here.]

Figures 8 and 9 show the time paths of the frequency of cooperative play in all random pairings sessions with 14 subjects under the I=1 (Figure 8) and I=2 (Figure 9) treatments. Table 6 provides aggregate cooperation frequencies for these two treatments analogous to Table 2. For comparison purposes, the aggregate cooperation frequencies in the random pairings I=0 treatment with 14 subjects, reported earlier in Table 2, are also reproduced in Table 6.

[Insert Table 6 here.]

Figures 8-9 reveal that, as in the random matching treatment with no information feedback (I=0), in both the I=1 and I=2 random matching treatments, there is no indication of any significant trend increase in cooperation rates with experience. Indeed, in most sessions there is a slight decrease in cooperative behavior over time. Further, the level of cooperation achieved in these treatments is quite low relative to that observed under the *fixed* pairings matching protocol with 14 subjects (compare Figure 8 or 9 with the left panel of Figure 2).

Analyzing the session level averages for the I=1 and I=2 treatments reported in Table 6, we find that there is no significant difference in the distribution of Game 1, Round 1 (initial) cooperation frequencies between the Random I=0 and I=1 treatments, between the Random I=0 and I=2 treatments or between the Random I=1 and I=2 treatments ($p>.10$ in all pairwise comparisons using nonparametric rank-order tests).[22] Thus, the additional information provided

---

21 Unambiguously labeling a player as a cooperator or defector may enhance reputational concerns as stressed, e.g., by Bolton et al. (2004). However, such reputational concerns play no role in Kandori's theory.

22 There is also no significant difference in initial cooperation frequencies between the Random matchings I=1 or I=2 treatments and the Fixed matchings treatment ($p>.10$).

in the random matching I=1 and I=2 protocols is not immediately taken into account by subjects, nor is there any initial perceived difference between these two types of information.

Though initial frequencies of cooperation are similar across all random pairings treatments, average cooperation frequencies are significantly higher in the I=1 treatment as compared with the I=0 treatment using session-level averages over all rounds (p=.05) or from the first half of each session (p= .10) or from the second half (p<.05). Further, in the I=1 treatment, there is no significant decrease in cooperation frequencies from the first half to the second half of sessions (p >.10), though cooperation frequencies in this treatment are low, averaging less than 20% in both halves of a session. By contrast, we are unable to reject the null of no difference in average cooperation frequencies between the I=2 and I=0 random pairings treatments using session averages from all rounds, or from the first or second half of a session (p>.10 in all cases). As in the I=1 treatment, there is no significant decrease in cooperation frequencies from the first to the second half of I=2 sessions (p>.10), but cooperation rates in the I=2 sessions average less than 20% in both halves of those sessions. Finally, we are unable to reject the null hypothesis of no difference in average cooperation frequencies between the I=1 and I=2 random pairings treatments using session averages from all rounds, or from the first or second half of sessions (p>.10 in all cases).

While there is no significant difference in cooperation frequencies between the I=1 and I=2 random pairings treatments, the cooperation frequencies in these two treatments (as in the I=0 treatment) remain *well below* those found in the fixed pairings treatment, where cooperation frequencies over all rounds averaged more than 50% . Cooperation frequencies in the fixed pairings treatment are significantly higher than in either the I=1 or I=2 treatments using session level averages from all rounds, or from the first or second half of sessions (p=.014 in all cases). The conclusion that emerges from this analysis is clear: regardless of the additional information we provide in the random pairings treatment, this information does not enable subjects to sustain cooperation rates that are anywhere close to those observed in the fixed pairings treatments. This finding suggests that it is the *matching protocol rather than information about an opponent's history* that plays the more important role in the achievement of high frequencies of cooperation.

4. Concluding Observations

Ellison (1994) observed that Kandori's theorem cast doubt on claims that the development of private local institutions, providing information on the reputations of individuals who participate in trade with various partners in the absence of enforceable contracts, were essential for the continued success of medieval trade fairs (Milgrom, et al. (1990)) and the international trading ventures of Maghribi traders (Grief (1989)). The experiment we report on in this paper shows that under anonymous random matching there is no evidence of the development of a cooperative norm even under the conditions of small group interaction or limited information about an opponent's past actions or histories. This finding gives weight to the argument that without the development of the kinds of institutions that can make an individual's reputation public, the systems of trading at medieval fairs and international trade conducted amongst members of the same tribe could not have been sustained.[23]

Ellison also observed that Kandori's theorem cast doubt on claims by experimenters that random, anonymous matching was sufficient to prevent subjects from treating all repetitions of a game played during an experimental session as a single supergame.[24] Our findings indicate that, as a matter of fact, the behavior of subjects who are in fixed pairings for the duration of a supergame is markedly different from the behavior of subjects who play a sequence of one shot games with random re-matching of anonymous players after each game is played. This finding suggests that random matching amongst anonymous players does, in fact, tend to suppress the inclination of subjects to treat all trials in a given session as a single supergame.

23 Similarly, our experimental findings cast doubt on the main proposition of Aliprantis et al. (2007) who (applying the logic of Kandori's result to a search model of money) argue that anonymous random pairings may not suffice to generate an essential role for money.

24 Experimenters know that subject behavior changes with experience. They wish to give subjects experience with a game without creating a supergame. A concern that subjects may treat all of the repetitions of a game played during an experimental session as a supergame, even when the intent is simply to give subjects experience with the game, has a long history. In commenting on the early Prisoner's Dilemma experiment conducted by Flood and Dresher, John Nash claimed that "The flaw in this experiment as a test of equilibrium point theory is that the experiment really amounts to having the players play one large multimove game…..Since 100 trials are so long that (backward induction is likely to fail) …it is fairly clear that one should expect (behavior) which is most appropriate for indeterminate end games..." Flood (1958, p. 16). If subjects further considered the existence of some conditionally cooperative player types, they would have an incentive to adjust their behavior with experience, even in experiments that utilized a random matching protocol, so as to learn about the percentage of the group with whom they are interacting that were conditional cooperators. Consequently, unless there is some way of making the number of conditional cooperators common knowledge at the outset of an experiment, supergame effects cannot be completely eliminated when a random matching protocol is used.

Finally, the results of our experiment establish empirically that in indefinitely repeated PD games played with fixed pairings, a community norm of cooperation becomes possible as subjects gain experience. This norm is achieved in both large and small groups, despite the anonymity of pairings and the presence of some `pessimistic' players, and even in cases where players have prior experience, under a random matching protocol, with a competitive norm.

Acknowledgements

The authors gratefully acknowledge financial support for this project from the National Science Foundation under grant SES-0213658. We thank the associate editor, the two referees and seminar participants at various institutions for their helpful comments and suggestions on earlier drafts. We also thank Scott Kinross and Jonathan Lafky for expert research assistance.

Appendix A

In this appendix we explain how we verified the existence of a "contagious" equilibrium as described in Kandori (1992, section 4) for the parameterization of the Prisoner's Dilemma game shown in Figure 1 and used in all experimental sessions reported in this paper. We also establish that under this same parameterization, the cooperative outcome can be supported as equilibrium of the indefinitely repeated game if both players in a fixed pairing adhere to a grim trigger strategy.

Let the stage game be described by the following symmetric payoff table showing the payoffs to the row player only

	C	D
C	w	x
D	y	z

Here C is the cooperative action and D is the defect action (labeled X and Y in the experiment).

In our experimental environment (unlike Kandori (1992)), we restricted w, x, y and z to be strictly nonnegative. Specifically, as noted in the text, we chose $w=20$, $x=0$, $y=30$ and $z=10$ so that the game is a Prisoner's Dilemma. To translate into Kandori's notation, the gain from defection, $g = y - w$, and the loss when cheated, $\ell = z - x$. Given our parameterization, $g = \ell = 10$.

A.1 Cooperative Equilibrium With Random Pairings

As in Kandori, let δ be the period discount factor and let M denote the population size. The M players are randomly paired in each round of an indefinitely repeated game. Suppose there are just two types of players in the population. Type c players are those whose history of play includes no defections; otherwise, a player is a type d player forever. The "contagious strategy" is for players to play the action corresponding to their type, i.e., type c's play C and type d's play D. Kandori (1992 Theorem 1) shows that the contagious strategy is a sequential equilibrium strategy for any given *g* and *M* provided that δ and ℓ are sufficiently large.

Following Kandori's (1992) notation, let X_t be total number of type d players in period t and let *A* be an $M \times M$ transition matrix with elements $a_{ij} = \Pr\{X_{t+1} = j \mid X_t = i\}$. Similarly, let *B* be an $M \times M$ transition matrix with elements $b_{ij} = \Pr\{X_{t+1} = j \mid X_t = i \ \ \textit{and} \text{ one type d player deviates to playing C at time } t\}$. The matrix *H=B-A* characterizes how the diffusion of d types is delayed if one d type unilaterally deviates from the contagious strategy. The conditional probability that a type d player randomly meets a type c player when there are *i* d types is given by the *i*th element of the column vector

$$\rho = \frac{1}{M-1}[M-1, M-2, \ldots, 1, 0]^T .$$

Finally, let e_i be a $1 \times M$ row vector with the *i*th element equal to 1 and all other elements equal to 0. Using the notation given above, we restate Kandori's Lemma.

The contagious equilibrium constitutes a sequential equilibrium if, first, a one-shot deviation from the equilibrium is unprofitable, i.e., if

$$\frac{w}{1-\delta} \geq \sum_{t=0}^{\infty} \delta^t \left[e_1 A^t \rho \ y + (1 - e_1 A^t \rho) z \right].$$

The left hand side is the expected payoff from cooperating forever and the right hand side is the expected payoff from defecting forever. The term $e_1 A^t \rho$ is the probability of meeting a type c player at time t given that the player was the first to defect at *t=0*. The above expression can be simplified to yield

$$\frac{w-z}{y-z} \geq (1-\delta) e_1 (I - \delta A)^{-1} \rho, \qquad (1)$$

which is comparable to equation (1) in Kandori (1992) under his normalization of *w=1, z=0* and using the definition *y=w+g*.

A second, sufficient condition for the contagious equilibrium strategy to be an equilibrium is that a one-shot deviation off the equilibrium path (a type d plays C) is unprofitable under any consistent belief. Specifically, the condition is that a type d player finds a one-shot deviation from playing D forever to be unprofitable given $X_t = k$, for all $k = 2, 3, \ldots, M$:

$$\sum_{t=0}^{\infty} \delta^t \left[e_k A^t \rho \ y + (1 - e_k A^t \rho) z \right] \geq \left(\frac{M-k}{M-1} \right) w + \left(\frac{k-1}{M-1} \right) x + \delta \sum_{t=0}^{\infty} \delta^t \left[e_k B A^t \rho \ y + (1 - e_k B A^t \rho) z \right].$$

The left hand side is the expected payoff from defecting forever when there are *k* d-type players including the player himself. The right hand side is what the player earns by deviating in the current period --playing C -- and then playing D forever; $(M-k)/(M-1)$ is the probability of meeting a type c player and $(k-1)/(M-1)$ is the probability of meeting a type d player. Finally, $e_k B$ is the distribution of the number of type d players in the next period given that in the current period there are k type d players and one of them (the player under consideration) deviates to playing C in the current period. The above expression can be simplified to yield

$$\left(\frac{M-k}{M-1}\right)(y-w)+\left(\frac{k-1}{M-1}\right)(z-x) \geq \delta\, e_k H (I-\delta\, A)^{-1} \rho\, (y-z) \text{ for } k = 2,3,\ldots M, \qquad (2)$$

which is again comparable to equation (2) in Kandori (1992) under his normalization of *w=1, z=0* and using the definition $g = y - w$ and $\ell = z - x$.

To check whether conditions (1-2) are satisfied under our parameterization of the stage game and for our choices of M and δ, we require the transition matrices *A* and *H*. Formulas for constructing these matrices are provided in Kandori (1989) and for completeness we reproduce these formulas here.

First, define the number of different ways of forming *M/2* pairs out of *M* individuals,

$$S(M) = \prod_{m=1}^{M/2} (2m-1).$$

Using this definition, a closed form solution for the $M \times M$ transition matrix *A* is given by the following formula. For *j=i, i+2, i+4,..., min[2i, M]*, if *i* is even and for *j=i+1, i+3, i+5,..., min[2i, M]* if *i* is odd,

$$a_{ij} = \frac{\binom{i}{j-i}\binom{M-i}{j-i}(j-i)!\, S(2i-j)\, S(M-j)}{S(M)},$$

otherwise

$$a_{ij} = 0.$$

A closed form solution for the $M \times M$ transition matrix $H = (h_{ij}) = B - A$ is given by the following formula. For *j=i+2, i+4,..., min[2i, M]*, if *i* is even, and *for j=i+1, i+3,..., min[2i, M]* if *i* is odd,

$$h_{ij} = \left(\frac{j-i}{i}\right) a_{ij} \text{ and } h_{i,j-1} = \left|h_{ij}\right|,$$

otherwise

$$h_{ij} = 0.$$

Using these definitions for the matrices *A* and *H*, we have verified that conditions (1-2) are satisfied for our parameter choices $\delta = .90$, *w=20, x=0, y=30 z=10* for even integer values of *M* over the range $2 \leq M \leq 30$.[25] (The maximum number of computers we have available in our computer laboratory is 30).

A.2 Cooperative Equilibrium with Fixed Pairings

When players remain paired with the same player for the duration of an indefinitely repeated game, a strategy where each player plays C in all rounds of the game is an equilibrium under our parameterization if players adhere to the "grim trigger" strategy, i.e., begin by

25 A Mathematica program that checks these conditions is available in the supplementary materials that accompany this paper or from the authors.

cooperating and if the history of play ever includes a defection, defect forever, otherwise continue cooperating.

Specifically, consider a player who decides to deviate from playing C in the current round. His one time gain from doing so, *g*=*y*-*w*. Since the other player is playing a grim trigger strategy, the deviant player faces a loss of *w*-*z* in the following period and forever after. Hence, the cooperative strategy is equilibrium provided that:

$$y - w < \delta \sum_{t=0}^{\infty} \delta^t (w - z), \text{ or}$$

$$y - w < \frac{\delta}{1-\delta}(w - z).$$

This is simply the condition that a deviation from the grim trigger strategy is unprofitable. Since *y-w=w-z=10* in our parameterization, this condition reduces to $.50 < \delta,$ which is readily satisfied by our choice of $\delta = .90.$ Hence, the grim trigger strategy supporting cooperative play is an equilibrium in the fixed pairings environment that we consider.

Appendix B

This appendix provides the written instructions used in the two main treatments of the experiment, the fixed pairings treatment and the random matching I=0 treatment. Instructions for these and all other treatments are available in the supplementary materials that accompany this paper or from the authors.

B.1 Instructions used in the fixed pairings treatment

<u>Overview</u>

This is an experiment in decision-making. The National Science Foundation has provided funds for this research. During the course of the experiment, you will be called upon to make a series of decisions. If you follow the instructions carefully and make good decisions, you can earn a considerable amount of money which will be paid to you in cash at the end of the experiment. We ask that you not talk with one another for the duration of the experiment.

<u>Specifics</u>

The experiment is divided into a series of games. A game will consist of an indefinite number of rounds. At the beginning of each game you will be paired with someone else in this room. You will be paired with this player for one game. In each round both of you will play the game described in the upper center portion of your screen. In this game each of you can make either of two choices, X or Y. The points you earn in a round depends upon both the choice you make and the choice made by the other person with whom you are matched. As the payoff table on your screen indicates:

If both of you choose X this round then: you both earn ***20*** points.

If you choose X this round and the other person chooses Y then: you earn ***0*** points and the other person earns ***30*** points.

If you choose Y this round and the other person chooses X then: you earn ***30*** points and the other person earns ***0*** points.

If you both choose Y this round then: you both earn ***10*** points.

To make your choice in each round, click the radio button next to either X or Y. You may change your mind any time prior to clicking the submit button by simply clicking on the button next to X or Y. You are free to choose X or Y in every round. When you are satisfied with your choice, click on the submit button. The computer program will record your choice and the choice made by the player with whom you are matched. After all players have made their choices, the results of the round will appear on the lower portion of your screen. You will be reminded of your own choice and will be shown the choice of the player with whom you are matched as well as the number of points you have earned for the round. Record the results of the round on your RECORD SHEET under the appropriate headings.

Immediately after you have received information on your choice and the choice of the person with whom you are matched for a given round, the computer program will randomly select a number from 1 to 100. The selected number will appear on a popup box in the middle of your screen. If this random number is less than 91, the game will continue into the next round. If the number selected is greater than 90 the game is over. Therefore, after each round there is a 90% chance that you will play another round with the same individual and a 10% chance that the game will end.

Suppose that a number less than 91 has been drawn. Then you click on the OK button, eliminating the popup box, and the next round is played. You will play the same game with the same individual as in the previous rounds Before making you choice, you may review all the outcomes of all of the prior games in the sequence by scrolling down the history record. You then choose either X or Y. Your choice and the choice of the person with whom you are matched are recorded and added to the history record at the lower portion of your screen. You record the outcome and your point earnings for the round. The computer then randomly selects a number between 1 and 100 to determine whether the game continues for another round.

If the number drawn is greater than 90 then the game ends. The experimenter will announce whether or not a new game will be played. If a new game is to be played then you will be matched with someone different from those you have been matched with in prior games. You will be matched with that person for all rounds in the new game.

Earnings

Each point that you earn is worth 1 cent ($.01). Therefore, the more points you earn the more money you earn. You will be paid your earnings from all rounds played today in cash, and in private, at the end of today's session.

Final Comments

First, do not discuss your choices or your results with anyone at any time during the experiment.

Second, your ID# is private. Do not reveal it to anyone.

Third, remember that you are paired with the same individual for the entire sequence of rounds in a given game. Since there is a 90% chance that at the end of a round the sequence will continue, you can expect, on average, to play 10 rounds with the same individual. However, since the stopping decision is made randomly, some sequences may be much longer than 10 rounds and others may be much shorter.

Questions?

Now is the time for questions. Does anyone have any questions before we begin?

B.2 Instructions used in the random pairings, no information (I=0) treatment

Overview

This is an experiment in decision-making. The National Science Foundation has provided funds for this research. During the course of the experiment, you will be called upon to make a series of decisions. If you follow the instructions carefully and make good decisions, you can earn a considerable amount of money which will be paid to you in cash at the end of the experiment. We ask that you not talk with one another for the duration of the experiment.

Specifics

The experiment is divided into a series of games. A game will consist of an indefinite number of rounds. At the beginning of each round you will be paired with someone else in this room. You will be paired with this player for one round. In each round you will play the game described in the upper center portion of your screen. In this game each of you can make either of two choices, X or Y. The points you earn in a round depends upon both the choice you make and the choice made by the other person with whom you are matched. As the payoff table on your screen indicates:

If both of you choose X this round then: you both earn ***20*** points.

If you choose X this round and the other person chooses Y then: you earn ***0*** points and the other person earns ***30*** points.

If you choose Y this round and the other person chooses X then: you earn ***30*** points and the other person earns ***0*** points.

If you both choose Y then: you both earn ***10*** points.

To make your choice in each round, click the radio button next to either X or Y. You may change your mind any time prior to clicking the submit button by simply clicking on the button next to X or Y. You are free to choose X or Y in every round. When you are satisfied with your choice, click the submit button. The computer program will record your choice and the choice made by the player with whom you are matched. After all players have made their choices, the results of the round will appear on the lower portion of your screen. You will be reminded of your own choice and will be shown the choice of the player with whom you are matched as well as the number of points you have earned for the round. Record the results of the round on your RECORD SHEET under the appropriate headings.

Immediately after you have received information on your choice and the choice of the person with whom you are matched for the round, the computer program will randomly select a number from 1 to 100. The selected number will appear on a popup box in the middle of your screen. If this random number is less than 91, the game will continue into the next round. If the number selected is greater than 90 the sequence is over. Therefore, after each round there is a 90% chance that you will play another round and a 10% chance that the game will end.

Suppose that a number less than 91 has been drawn. Then you press the OK button eliminating the popup box and the next round is played. You will play the same game as in the previous round, but **with an individual selected at random from all the individuals in the room**. Before making your choice, you may review all the outcomes of all of the prior games in the sequence by scrolling down the history record. You then choose either X or Y. Your choice and the choice of the person with whom you are matched this round are recorded and added to the history record at the lower portion of your screen. You record the outcome and your point earnings for the round. The computer then randomly selects a number between 1 and 100 to determine whether the game continues for another round.

If the number drawn is greater than 90 then the game ends. The experimenter will announce whether or not a new game will be played. If a new game is to be played then you will be matched with someone drawn at random from the other people in the room. The new game will then be played as described above.

Earnings

Each point that you earn is worth 1 cent ($.01). Therefore, the more points you earn the more money you earn. You will be paid your earnings from all rounds played today in cash and in private at the end of today's session.

Final Comments

First, do not discuss your choices or your results with anyone at any time during the experiment.

Second, your ID# is private. Do not reveal it to anyone.

Third, since there is a 90% chance that at the end of a round the sequence will continue, you can expect, on average, to play 10 rounds in a given game sequence. However, since the stopping decision is made randomly, some sequences may be much longer than 10 rounds and others may be much shorter.

Fourth, remember that after each round of a game you will be matched randomly with someone in this room. Therefore, if there are N people in the room the probability of you being matched with the same individual in two consecutive rounds of a game is 1/(N-1).

Questions?

Now is the time for questions. Does anyone have any questions before we begin?

References

Aliprantis, C.D., Camera, G., Puzzello, D., 2007. Contagion equilibria in a monetary model. Econometrica 75, 277-282.

Aoyagi, M., Frechette, G., 2005. Collusion as public monitoring becomes noisy: Experimental evidence. Working paper, New York University and Osaka University.

Andreoni, J., Croson, R., 2002. Partners versus strangers: The effect of random rematching in public goods experiments. To appear in Plott, C., Smith, V., (Eds.), Handbook of Experimental Economics Results. Elsevier, Amsterdam/New York.

Axelrod, R., 1984. The Evolution of Cooperation. Basic Books, New York.

Bolton, G.E. Katok, E., Ockenfels, A., 2004. How effective are electronic reputation mechanisms? An experimental investigation. Management Science 50, 1587-1602

Bolton, G.E., Katok, E., Ockenfels, A., 2005. Cooperation among strangers with limited information about reputation. Journal of Public Economics 89, 1457-1468.

Dal Bó, P., 2005. Cooperation under the shadow of the future: Experimental evidence from infinitely repeated games. American Economic Review 95, 1591-1604.

Ellison, G., 1994. Cooperation in the Prisoner's Dilemma with anonymous random matching. Review of Economic Studies 61, 567-588.

Feltovich, N., 2003. Nonparametric tests of differences in medians: comparison of the Wilcoxan-Mann-Whitney and Robust Rank-Order tests. Experimental Economics 6, 273-297.

Fischbacher, U., Gächter, S., 2006. Heterogeneous social preferences and the dynamics of free-riding in public goods. Institute for Empirical Research in Economics, University of Zurich, Working Papers Series ISSN 1424-0459.

Fischbacher, U., Gächter, S., Fehr, E., 2001. Are people conditionally cooperative? Evidence from a public goods experiment. Economics Letters 71, 397-404.

Flood, M.M., 1958. Some experimental games. Management Science 5, 5-26.

Greif, A., 1989. Reputation and coalitions in medieval trade: Evidence on the Maghribi traders coalition. Journal of Economic History 49, 857-882.

Holt, C.A., 1985. An experimental test of the consistent-conjectures hypothesis. American Economic Review 75, 314-325.

Kandori, M., 1989. Social norms and community enforcement. CARESS Working Paper No. 89-14, University of Pennsylvania.

Kandori, M., 1992. Social norms and community enforcement. Review of Economic Studies 59, 63-80.

Milgrom, P.R., North, D.C., Weingast, B.R., 1990. The role of institutions in the revival of trade: The law merchant, private judges, and the Champagne fairs. Economics and Politics 2, 1-23.

Palfrey, T.R., Rosenthal, H., 1994. Repeated play, cooperation and coordination: An experimental study. Review of Economic Studies 61, 545-565.

Roth, A.E., Murnighan, K., 1978. Equilibrium behavior and repeated play of the Prisoner’s Dilemma. Journal of Mathematical Psychology 17, 189-198.

Schwartz, S., Young, R., Zvinakis, K., 2000. Reputation without repeated interaction: A role for public disclosures. Review of Accounting Studies 5, 351-375.

Siegel, S., Castellan, Jr., N.J., 1988. Nonparametric Statistics for the Behavioral Sciences, 2nd Ed. New York, McGraw Hill.

Van Huyck, J.B., Wildenthal, J.M., Battalio, R.C., 2002. Tacit cooperation, strategic uncertainty, and coordination failure: Evidence from repeated dominance solvable games. Games and Economic Behavior 38, 156-175.

Table 1: Characteristics of Experimental Sessions

Treatment	Treatment Session No.	Number of Subjects	Number of Supergames	Number of Rounds
Fixed	1	14	11	59
Fixed	2	14	10	96
Fixed	3	14	13	131
Fixed	4	14	10	115
Random I=0	1	14	10	112
Random I=0	2	14	12	104
Random I=0	3	14	8	97
Random I=0	4	14	8	89
One Shot then Random I=0	1	14	9	84
One Shot then Random I=0	2	14	12	78
One Shot then Random I=0	3	14	11	79
Random I=1	1	14	9	75
Random I=1	2	14	9	106
Random I=1	3	14	16	99
Random I=1	4	14	9	105
Random I=2	1	14	14	110
Random I=2	2	14	8	101
Random I=2	3	14	7	100
Fixed Then Random I=0	1	14	15	134
Fixed Then Random I=0	2	14	15	127
Random I=0 Then Fixed	1	14	13	113
Random I=0 Then Fixed	2	14	11	133
Fixed	1	6	10	109
Fixed	2	6	9	108
Fixed	3	6	13	108
Random I=0	1	6	12	104
Random I=0	2	6	12	100
Random I=0	3	6	7	88
	28 Sessions	Total: 344 Subjects	Avg.=10.8	Avg.=102.3

Table 2: Aggregate Frequencies of Cooperation
Fixed or Random (I=0) Matching Sessions with 14 Subjects

Fixed	Game 1, Round 1	All Rounds of All Games	First Half of the Session	Second Half of the Session
Session 1	0.786	0.548	0.478	0.617
Session 2	0.286	0.576	0.457	0.695
Session 3	0.571	0.477	0.408	0.545
Session 4	0.286	0.608	0.520	0.695
All Sessions	0.482	0.549	0.462	0.634

Random I=0	Game 1, Round 1	All Rounds of All Games	First Half of the Session	Second Half of the Session
Session 1	0.214	0.022	0.034	0.010
Session 2	0.429	0.042	0.080	0.004
Session 3	0.429	0.063	0.097	0.029
Session 4	0.643	0.173	0.211	0.137
All Sessions	0.429	0.075	0.105	0.045

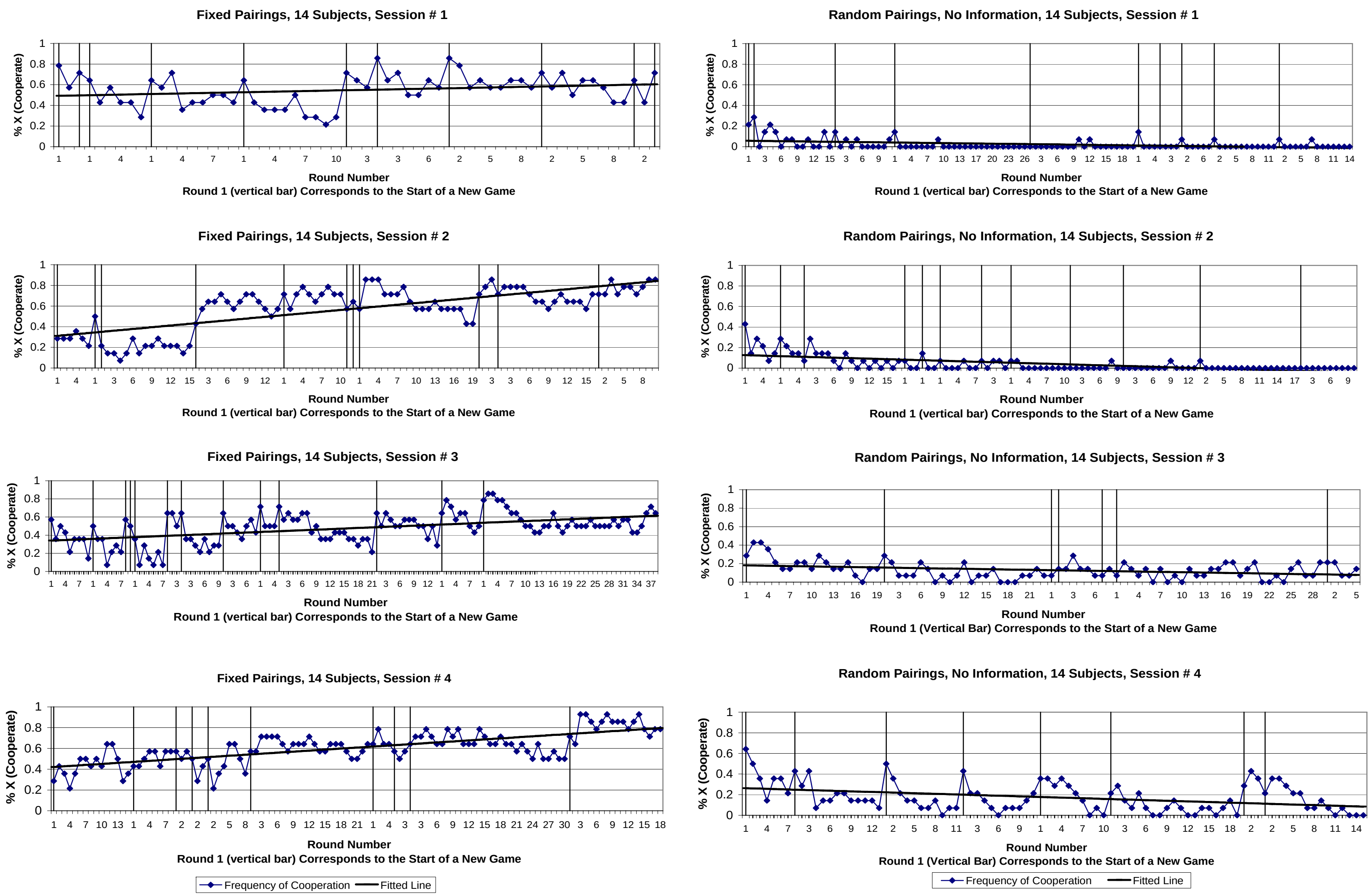


Figure 2: Aggregate Frequency of Cooperation in Fixed (left column) and Random, I=0 (right column) Matching Sessions with 14 Subjects

Figure 3a: Frequency of Cooperation in First Rounds Versus All Remaining Rounds of All Supergames, Pooled Data, All Sessions of a Treatment

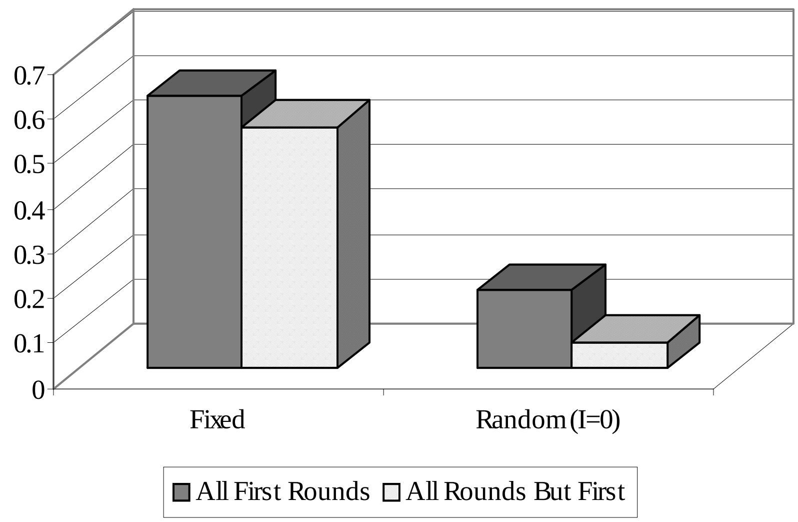


Figure 3b: Frequency of Cooperation in First Rounds of Supergames 1-10, Pooled Data from All Sessions of a Treatment

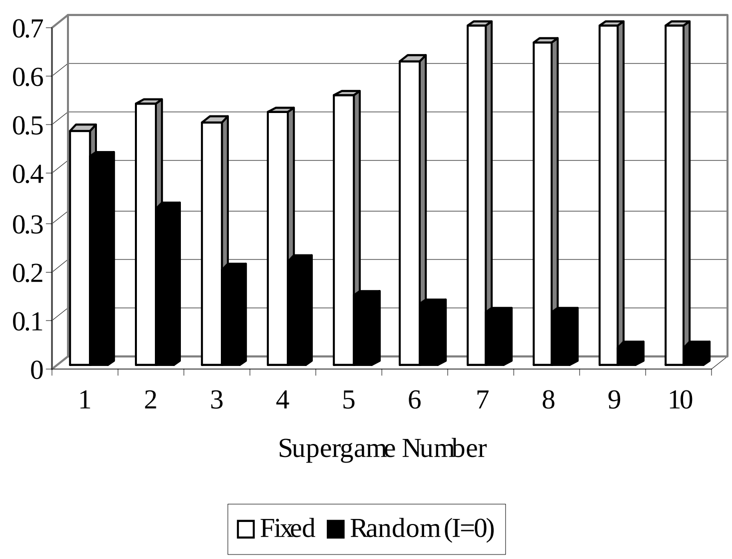
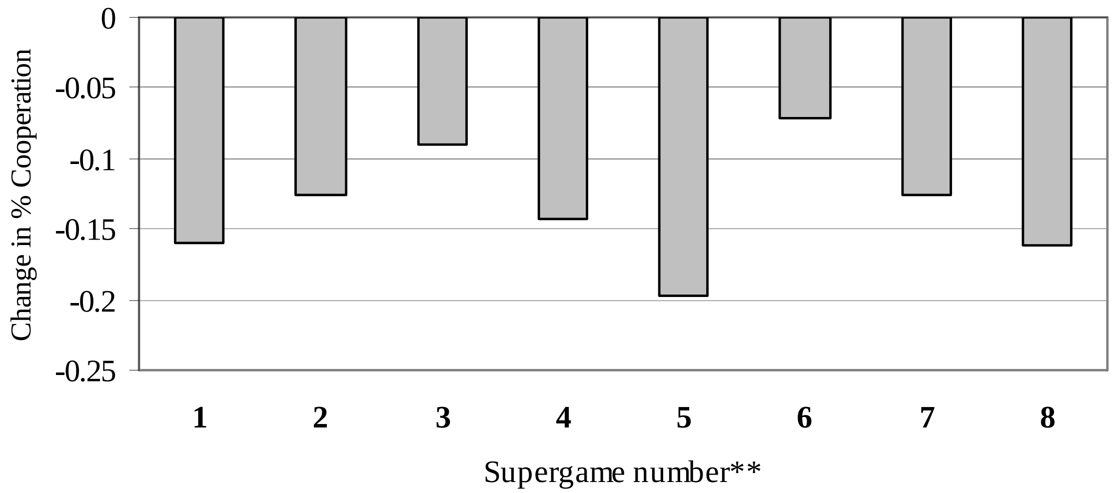


Table 3: Individual Frequencies of Cooperation

Fixed Pairings	Cumulative Number (Cum %) of the 14 Subjects Whose Frequency of Cooperation Falls Below Various Thresholds Over All Rounds and Over the 1st Half and the 2nd Half of Each Session												All 4 Sessions Combined Cumulative Frequency All Rounds
Frequency of Cooperation Is:	Session 1			Session 2			Session 3			Session 4			
	All Rounds	1st Half	2nd Half	All Rounds	1st Half	2nd Half	All Rounds	1st Half	2nd Half	All Rounds	1st Half	2nd Half	
<.05	1 (.071)	1 (.071)	1 (.071)	0 (.000)	2 (.143)	0 (.000)	0 (.000)	0 (.000)	1 (.071)	0 (.000)	0 (.000)	0 (.000)	0.018
<.10	1 (.071)	1 (.071)	1 (.071)	1 (.071)	2 (.143)	0 (.000)	2 (.143)	2 (.143)	2 (.143)	0 (.000)	1 (.071)	0 (.000)	0.071
<.25	2 (.143)	3 (.339)	2 (.143)	1 (.071)	2 (.143)	1 (.071)	3 (.214)	4 (.286)	4 (.286)	0 (.000)	1 (.071)	0 (.000)	0.107
<.50	4 (.286)	7 (.500)	5 (.357)	5 (.357)	9 (.643)	2 (.143)	7 (.500)	9 (.643)	7 (.500)	3 (.214)	7 (.500)	3 (.214)	0.339
<.75	11 (.786)	11 (.786)	8 (.571)	11 (.786)	13 (.929)	7 (.500)	10 (.714)	14 (1.00)	8 (.571)	10 (.714)	12 (.857)	6 (.429)	0.750
<=1.0	14 (1.00)	14 (1.00)	14 (1.00)	14 (1.00)	14 (1.00)	14 (1.00)	14 (1.00)	14 (1.00)	14 (1.00)	14 (1.00)	14 (1.00)	14 (1.00)	1.000

Random Pairings I=0	Cumulative Number (Cum %) of the 14 Subjects Whose Frequency of Cooperation Falls Below Various Thresholds Over All Rounds and Over the 1st Half and the 2nd Half of Each Session												All 4 Sessions Combined Cumulative Frequency All Rounds
Frequency of Cooperation Is:	Session 1			Session 2			Session 3			Session 4			
	All Rounds	1st Half	2nd Half	All Rounds	1st Half	2nd Half	All Rounds	1st Half	2nd Half	All Rounds	1st Half	2nd Half	
<.05	13 (.929)	10 (.714)	13 (.929)	9 (.643)	8 (.571)	13 (.929)	4 (.286)	1 (.071)	10 (.714)	1 (.071)	1 (.071)	4 (.286)	0.482
<.10	13 (.929)	13 (.929)	14 (1.00)	12 (.857)	11 (.786)	14 (1.00)	14 (1.00)	8 (.571)	14 (1.00)	3 (.214)	2 (.143)	5 (.357)	0.750
<.25	14 (1.00)	14 (1.00)	14 (1.00)	14 (1.00)	13 (.929)	14 (1.00)	14 (1.00)	14 (1.00)	14 (1.00)	11 (.786)	9 (.643)	12 (.857)	0.946
<.50	14 (1.00)	14 (1.00)	14 (1.00)	14 (1.00)	14 (1.00)	14 (1.00)	14 (1.00)	14 (1.00)	14 (1.00)	14 (1.00)	14 (1.00)	14 (1.00)	1.000
<.75	14 (1.00)	14 (1.00)	14 (1.00)	14 (1.00)	14 (1.00)	14 (1.00)	14 (1.00)	14 (1.00)	14 (1.00)	14 (1.00)	14 (1.00)	14 (1.00)	1.000
<=1.0	14 (1.00)	14 (1.00)	14 (1.00)	14 (1.00)	14 (1.00)	14 (1.00)	14 (1.00)	14 (1.00)	14 (1.00)	14 (1.00)	14 (1.00)	14 (1.00)	1.000

Figure 4: Change in the Aggregate Frequency of Cooperation from the First to the Last Round of Each Supergame*
(Pooled Data from 4 Fixed Pairing Sessions with 14 Subjects Each)



* Change is: Avg. % Cooperation in Final Round of Supergame - Avg. % Cooperation in First Round of Supergame. ** Supergames are in the order played, excluding any 1-round supergames.

Figure 5: Cumulative Frequency Distribution of Elicited Strategy Thresholds in the First Part of the "One Shot Then Random I=0" Treatment. Data from 3 Sessions, 14 Subjects Per Session.

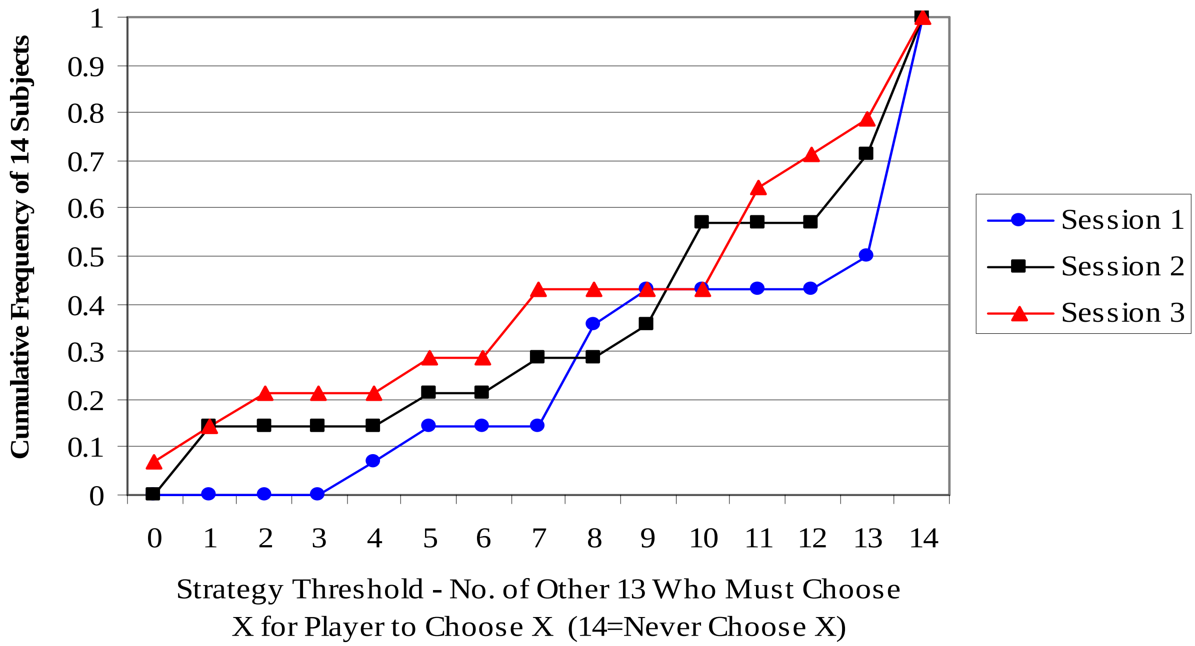


Table 4: Aggregate Cooperation Frequencies and Strategy Prediction Accuracy
One-Shot Then Random (I=0) Matching Sessions with 14 Subjects

Subject Group	First Part One Shot %C	Second Part Game 1, Round 1 %C	Second Part All Rounds of All Games %C	Second Part First Half of the Session %C	Second Part Second Half of the Session %C	Second Part Strategy Prediction Accuracy All Rounds of All Games
All Subjects						
Session 1	0.143	0.143	0.096	0.110	0.083	0.899
Session 2	0.286	0.071	0.015	0.026	0.004	0.955
Session 3	0.357	0.214	0.019	0.027	0.012	0.806
All Sessions	0.262	0.143	0.043	0.054	0.033	0.887
Conditional Cooperators Only						
Session 1: 7 Players	0.286	0.143	0.153	0.173	0.133	0.837
Session 2: 10 Players	0.400	0.100	0.019	0.033	0.005	0.938
Session 3: 11 Players	0.455	0.273	0.024	0.035	0.014	0.753
All Sessions	0.380	0.172	0.065	0.081	0.050	0.843
Unconditional Defectors Only						
Session 1: 7 Players	0.000	0.143	0.039	0.044	0.034	0.961
Session 2: 4 Players	0.000	0.000	0.003	0.006	0.000	0.997
Session 3: 3 Players	0.000	0.000	0.000	0.000	0.000	1.000
All Sessions	0.000	0.048	0.014	0.017	0.011	0.986

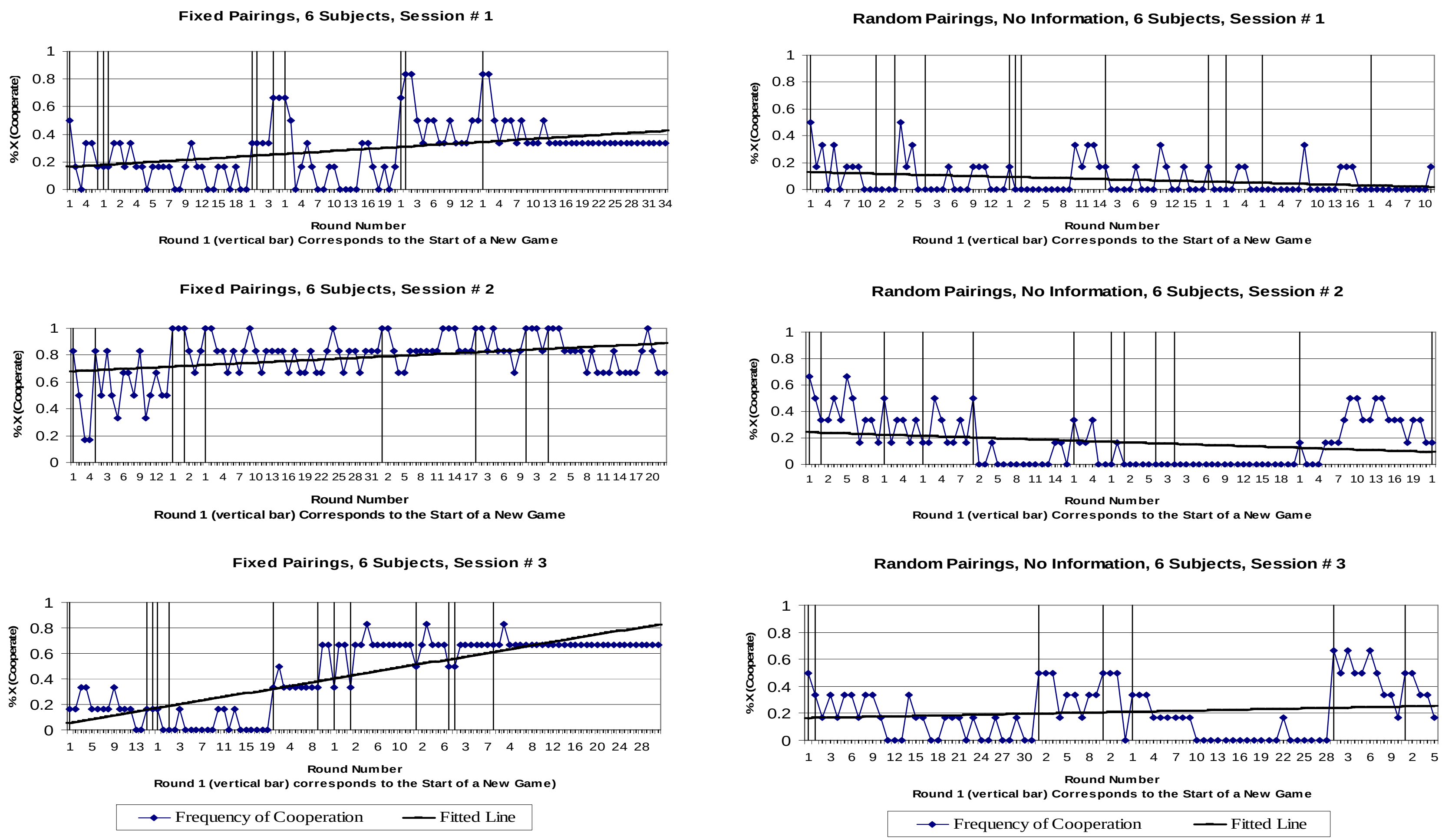


Figure 6: Aggregate Frequency of Cooperation in Fixed (left column) and Random, I=0, (right column) Matching Sessions with 6 Subjects

Table 5: Aggregate Frequencies of Cooperation
All Sessions with 6 Subjects and a Single Matching Protocol

Fixed	Game 1, Round 1	All Rounds of All Games	First Half of the Session	Second Half of the Session
Session 1	0.500	0.292	0.189	0.418
Session 2	0.833	0.782	0.729	0.840
Session 3	0.167	0.440	0.280	0.663
All Sessions	0.500	0.505	0.399	0.634

Random I=0	Game 1, Round 1	All Rounds of All Games	First Half of the Session	Second Half of the Session
Session 1	0.500	0.072	0.102	0.045
Session 2	0.667	0.167	0.217	0.117
Session 3	0.500	0.208	0.216	0.201
All Sessions	0.556	0.149	0.178	0.121

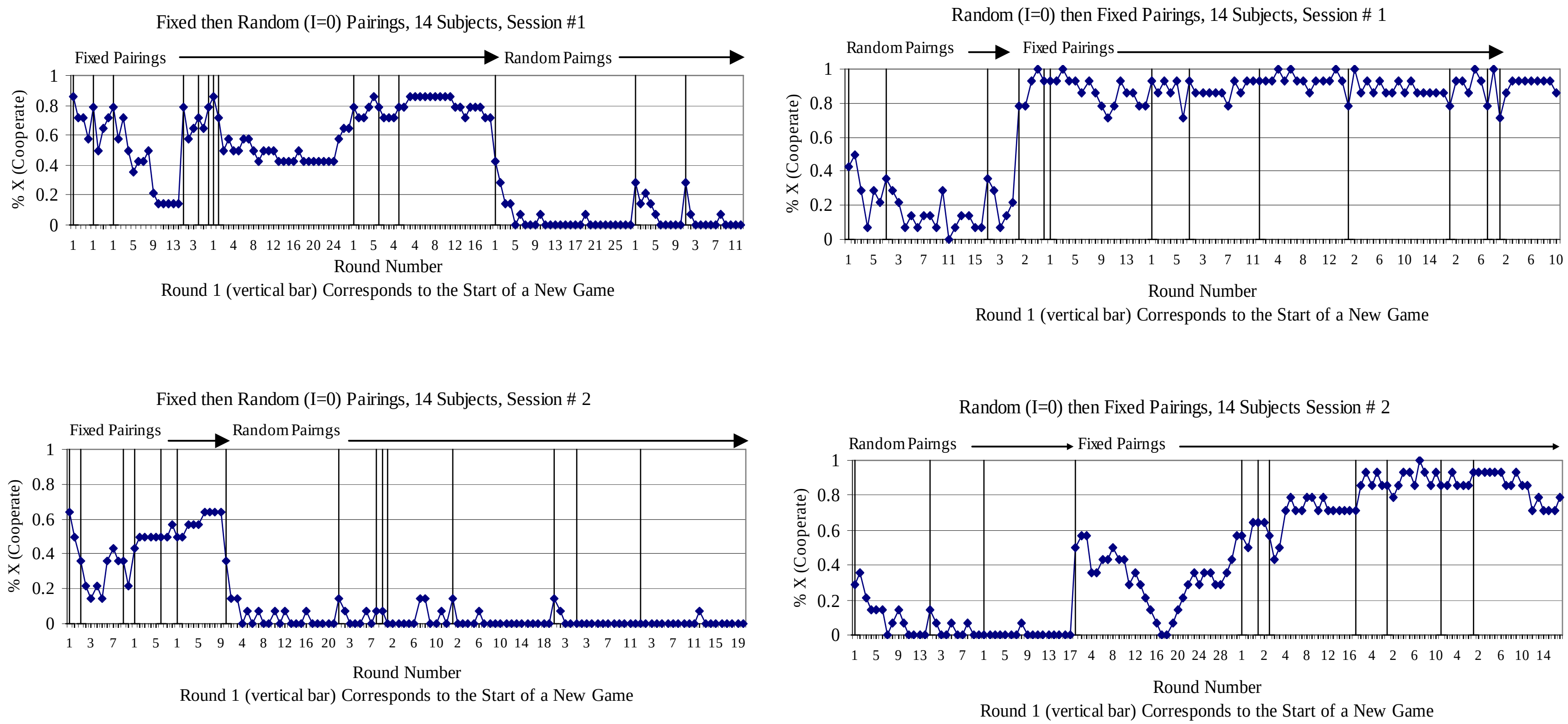


Figure 7: Within Subjects Treatments. 2 Sessions of Fixed followed by Random (I=0) Matching, 14 subjects each (left column) and 2 Sessions of Random (I=0) followed by Fixed Matching, 14 subjects each. The Timing of Each Matching Protocol is Indicated on Each Figure

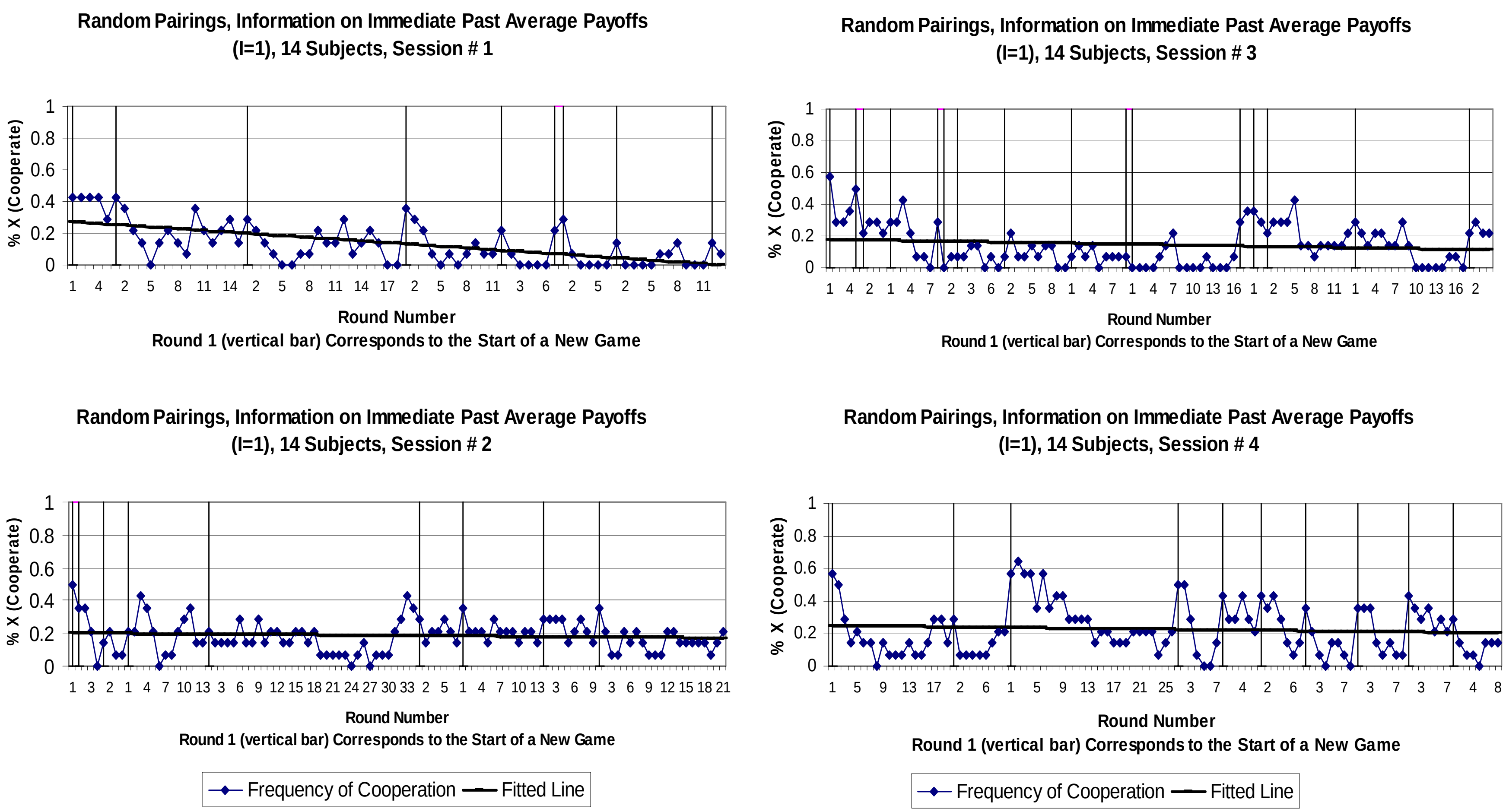


Figure 8: Aggregate Frequency of Cooperation in Four Random (I=1) Matching Sessions with 14 Subjects

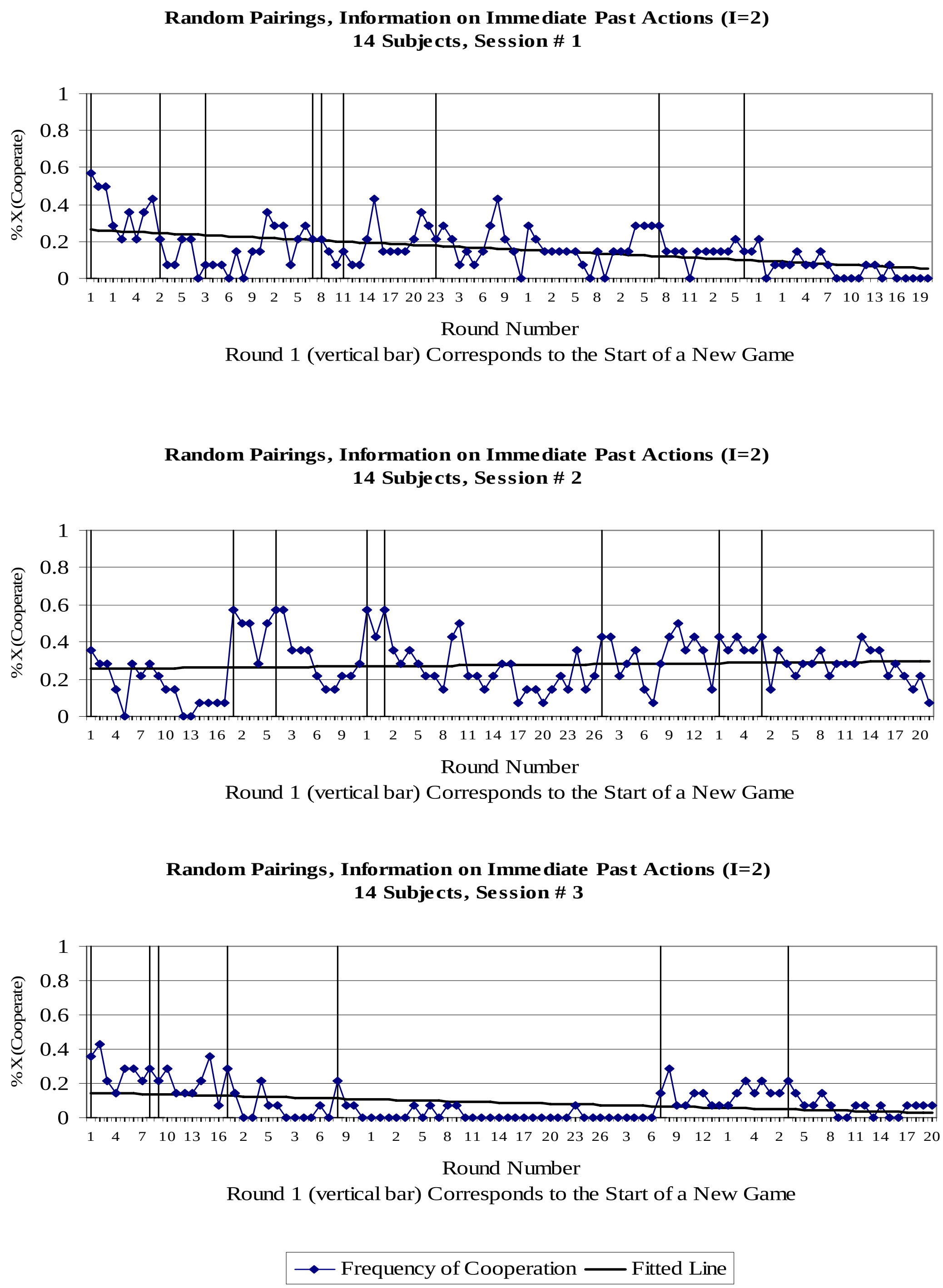


Figure 9: Aggregate Frequency of Cooperation in Three Random (I=2) Matching Sessions with 14 Subjects

Table 6: Aggregate Frequencies of Cooperation
All Random Matching Sessions with 14 Subjects, I=0, 1, 2

Random I=0	Game 1, Round 1	All Rounds of All Games	First Half of the Session	Second Half of the Session
Session 1	0.214	0.022	0.034	0.010
Session 2	0.429	0.042	0.080	0.004
Session 3	0.429	0.063	0.097	0.029
Session 4	0.643	0.173	0.211	0.137
All Sessions	0.429	0.075	0.105	0.045

Random I=1	Game 1, Round 1	All Rounds of All Games	First Half of the Session	Second Half of the Session
Session 1	0.429	0.135	0.197	0.075
Session 2	0.500	0.185	0.168	0.202
Session 3	0.571	0.144	0.143	0.144
Session 4	0.571	0.225	0.236	0.214
All Sessions	0.518	0.176	0.186	0.166

Random I=2	Game 1, Round 1	All Rounds of All Games	First Half of the Session	Second Half of the Session
Session 1	0.571	0.158	0.207	0.110
Session 2	0.357	0.276	0.277	0.275
Session 3	0.357	0.086	0.108	0.066
All Sessions	0.429	0.173	0.198	0.143