

**The Evolution of Cooperation in Infinitely Repeated Games:
Experimental Evidence
Pedro Dal Bó and Guillaume R. Fréchette
Web Appendix**

1. Instructions

Welcome

You are about to participate in a session on decision-making, and you will be paid for your participation with cash vouchers, privately at the end of the session. What you earn depends partly on your decisions, partly on the decisions of others, and partly on chance.

Please turn off pagers and cellular phones now. Please close any program you may have open on the computer.

The entire session will take place through computer terminals, and all interaction between you will take place through the computers. Please do not talk or in any way try to communicate with other participants during the session.

We will start with a brief instruction period. During the instruction period you will be given a description of the main features of the session and will be shown how to use the computers. If you have any questions during this period, raise your hand and your question will be answered so everyone can hear.

General Instructions

1. In this experiment you will be asked to make decisions in several rounds. You will be randomly paired with another person for a sequence of rounds. Each sequence of rounds is referred to as a match.

2. The length of a match is randomly determined. After each round, there is a 50% probability that the match will continue for at least another round. This is as if we would flip a coin after each round and continue if tails and end if heads. So, for

instance, if you are in round 2, the probability there will be a third round is 50% and if you are in round 9, the probability there will be another round is also 50%.

3. Once a match ends, you will be randomly paired with another person for a new match.

4. The choices and the payoffs in each round are as follows:

	the other's choice	
your choice	1	2
1	32, 32	12, 50
2	50, 12	25, 25

The first entry in each cell represents your payoff, while the second entry represents the payoff of the person you are matched with.

- As you can see, this shows the payoff associated with each choice. Once you and the person you are paired with have made your choices, those choices will be highlighted and your payoff for the round will appear.

That is, if:

You select 1 and the other selects 1, you each make 32.

You select 1 and the other selects 2, you make 12 while the other makes 50.

You select 2 and the other selects 2, you make 50 while the other makes 12.

You select 2 and the other selects 2, you each make 25.

- At the end of the experiment (the first match to end after 1 hour of play) you will be paid \$0.006 for every point scored. There is no show-up fee for this experiment.

- Are there any questions?

Before we start, let me remind you that:

- The length of a match is randomly determined. After each round, there is a 50% probability that the match will continue for at least another round. You will play with the same person for the entire match.

- After a match is finished, you will be randomly paired with another person for a new match.

2. Additional analysis

In this appendix we include additional tables and analysis (some of which are mentioned in the main text of the paper). We organized this appendix following the sections in the main text.

III. Main Experimental Results

Some additional descriptive statistics of the sessions are presented in Table A2.1.

B. Do subjects learn to defect when it is the only equilibrium action?

We show next that the results described in the text also hold if we study linear trends by interaction and round.¹ We aggregate data by interaction and equilibrium or risk dominance conditions. We estimate first the linear trend in the evolution of cooperation in the treatment with $\delta=1/2$ & $R=32$. Since the session with the minimum number of interactions had only 119 interactions, we focus on interactions not greater than that number, so as not to give more importance to some of the sessions.²

As columns (1) and (3) in Table A2.2 show, the estimated coefficient for the linear trend is negative and significantly different from zero. This is the case for first rounds only and all rounds considered together. See also the top two graphs in Figure A2.1 which show the evolution of cooperation across interactions.

Table A2.3 shows the estimates of linear trends for rounds 2 to 4 (we do not study rounds greater than four as the number of observations is small). Columns (1) to (3) show that the trend is negative and significant at the 10% for rounds 2 and 3 but positive and non-significant for rounds 4 when cooperation cannot be supported in equilibrium.

¹ As in the text, we used interactions (decision stage counted from the beginning of the session regardless of repeated game) instead of matches so as to be able to compare sessions with different number of repeated games. Similar results hold if we focus on matches instead of interactions.

² Remember that we use the word interaction to number each decision stage regardless of the repeated game.

If we calculate a linear trend for the first rounds in each of the session in the treatment with $\delta=1/2$ & $R=32$ we find that the linear trend is negatively sloped in all three sessions. See Table A2.4 and the first graph in Figure A2.3.

C. Do subjects learn to cooperate when it is an equilibrium action?

When cooperation is an equilibrium action the estimated coefficient for the linear trend is positive and significantly different from zero for both first rounds and all rounds – columns (2) and (4) in Table A2.2. See also the top two graphs in Figure A2.1. The slope of the linear trend for treatments when cooperation is an equilibrium action is significantly larger than the slope for treatments under which cooperation is not an equilibrium action (p-values <0.001 for both first rounds and all rounds)

If we calculate the linear trends for rounds 2 to 4 separately we also find positive and significant linear trends when cooperation can be supported in equilibrium – see columns (4) to (6) in Table A2.3.

For treatments under which cooperation is an equilibrium action but is not risk dominant the result varies depending on whether we focus on first or all rounds. For all rounds there is a slight but significant positive trend, while it is not significant for first rounds – see columns (5) and (7) in Table A2.2. The difference with the other statistical analysis of the evolution of cooperation in these treatments can be explained by the sharp decrease in cooperation at the beginning of these sessions followed by a small increase thereafter.

Consistently with the analysis in the text, we find that if we consider each session separately it is clear that being a possible equilibrium action does not necessarily lead to increasing levels of cooperation as subjects gain experience. Of the 15 session in the treatment in which cooperation can be supported in equilibrium, 10 have a significantly increasing linear trend, 2 have a significantly decreasing linear trend, and 3 have no significant trend at the 10% level – see Table A2.4 and Figure A2.3. If we focus on sessions under which cooperation can be supported in equilibrium but is not risk dominant, 2 sessions have significantly increasing linear trends, one significantly decreasing and 3 have no significant linear trend at the 10% level.

D. Do subjects learn to cooperate when it is risk-dominant to do so?

When we focus on treatments for which cooperation is the risk dominant action, we find that the linear trend is significantly positive for both first rounds and all rounds—see columns (6) and (8) in Table A2.2.³ The same is true for rounds 2 to 4 considered separately – see Table A2.3 and Figure A2.2. The slope of the linear trend for these treatments is significantly larger than for treatments for which cooperation is an equilibrium action but is not risk dominant (p-values < 0.001 for first rounds and all rounds).

However, and consistent with the analysis in the text, we find that if we consider each session separately it is clear that being a possible equilibrium action does not necessarily lead to increasing levels of cooperation as subjects gain experience. Of the 5 sessions in the treatment in which cooperation can be supported in equilibrium, 4 have a significantly increasing linear trend but one has a significantly decreasing trend – see Table A2.4 and Figure A2.3.

E. The last repeated game

The evolution of cooperation can also be seen when we study cooperation rates in the last repeated game (and comparing them with those in the first repeated game). Table A2.5 provides the cooperation rates in the last repeated game of each treatment.⁴ Increases in the probability of continuation and the payoff of cooperation result in increases in cooperation rates. Also, cooperation is higher when it can be supported as part of an equilibrium (p-values < 0.001 for both first and all rounds). However, when cooperation can be supported in equilibrium, it is not the case that cooperation is always greater in the last repeated game than in the first repeated game (see in Tables 5 and A2.5, all treatments with $\delta=1/2$ and the treatment with $\delta=3/4$ & $R=32$).

Similarly, when considering together all sessions in treatments where cooperation can be supported in equilibrium, rates of cooperation are, on average, higher in the last repeated game if cooperation is risk-dominant (p-values < 0.01 for both first and all

³ In this part we consider data up to interaction 128 since this is the lowest number of interactions reached by one of the sessions in these treatments.

⁴ The top panel in Table 3 in the text has the same information but for the first repeated game. Hence the sample size for the left panel is exactly the same as in the top left panel of Table 3, and the sample size for the right panel is about the same as for the top right panel of Table 3.

rounds). But for sessions in these latter treatments considered separately, it is also true that cooperation is not always greater in the last repeated game when compared to the first (see the treatment with $\delta=1/2$ & $R=48$). However, note that cooperation can reach very high levels in the last repeated game, as it is the case in the treatment with $\delta=3/4$ & $R=48$.

IV. Discussion

Table A2.6 shows probit estimates of the effect on cooperation rate in the last repeated game of the size of the basin of attraction, the number of rounds over the expected number and dummy variables for whether cooperation is a possible equilibrium outcome and whether it is risk dominant. The table shows that whether cooperation is a possible equilibrium outcome and whether it is risk dominant are not significant determinants of cooperation. Moreover, the inclusion of these two variables does not affect the significance of the other variables.

Table A2.7 show the probit estimates of the effect of past observations on Round 1 Cooperation that are used to generate the marginal effects reported in Table 6 of the main text.

We describe next the methodology for the estimation of the importance of strategies presented in the text in Table 7. As described in the text we focus on a set of six strategies and the importance of each strategy is estimated by maximum likelihood. The likelihood that the data corresponds to a given strategy is obtained by allowing a deviation in any round between the strategy and the observed choices:

$y_{imr}(s^k) = 1\{s_{imr}(s^k) + \gamma\varepsilon_{imr} \geq 0\}$ where y is the choice (1 for cooperate or 0 for defect), $1\{\cdot\}$ is an indicator function, imr stands for subject i , match m and round r , s^k is a specific strategy, and $s_{imr}(s^k)$ specifies the choice implied by that strategy given the history of the repeated game up to that round in that match (it is coded with 1 if the strategy would cooperate and -1 otherwise), ε is the error term and γ is the variance in the error. The error term is such that the resulting likelihood has the usual logistic form and this results in the likelihood

$$p_i(s^k) = \prod_M \prod_R \left(\frac{1}{1 + \exp(-s_{imr}(s^k)/\gamma)} \right)^{y_{imr}} \left(\frac{1}{1 + \exp(s_{imr}(s^k)/\gamma)} \right)^{1-y_{imr}}$$
 for a given subject and strategy (where M and R represent the sets of all matches and rounds). From this we obtain a loglikelihood $\sum_I \ln \left(\sum_K p(s^k) p_i(s^k) \right)$ where K represents the set of strategies we consider, labeled s^1 to s^K , and $p(s^k)$ is the proportion of the data which is attributed to strategy s^k .

Repeated games that lasted only one round provide little information to identify strategies (while they allow distinguishing between cooperative and non-cooperative strategies they do not help us to distinguish among the cooperative strategies we consider). We show in Table A2.8 that the results presented in Table 7 are robust to eliminating repeated games that lasted only one round.

The learning model is as follows. Subjects in the first repeated game have beliefs about the probability their partner uses either AD or a cooperative strategy like TFT. These beliefs are tracked by two variables: β_{it}^D and β_{it}^C such that the belief by subject j at time t that his partner will play AD is $\beta_{it}^D / (\beta_{it}^D + \beta_{it}^C)$. In the first repeated game each subject has a given β_{i1}^D and β_{i1}^C . After the first repeated game is played they update their beliefs as follows $\beta_{it+1}^k = \theta_i \beta_{it}^k + 1(a_j^k)$ where θ_i discounts past beliefs ($\theta_i = 0$ gives Cournot dynamics and $\theta_i = 1$ is fictitious play), k is the action and $1(a_j^k)$ is an indicator function that takes value 1 if subject j (with whom i is paired) took the action k (TFT or AD). In fact, we will abstract from the complexities of the repeated game by reducing it to the choice in round 1: defect corresponds to AD and cooperate corresponds to TFT. Given those beliefs, subject i is modeled as a random utility maximizer where each choice yields $U_{it}^a = \frac{\beta_{it}^D}{\beta_{it}^D + \beta_{it}^C} u^a(a_j^D) + \frac{\beta_{it}^C}{\beta_{it}^D + \beta_{it}^C} u^a(a_j^C) + \lambda_{it} \varepsilon_{it}^a$ where $u^a(a_j^k)$ is the average payoff from taking action a when subject i is paired with j and takes action k . The expected return from each choice is given by the theoretical values. The parameter λ_{it} is a scaling parameter that measures how well the subject best-responds to his beliefs where $\lambda_{it} = \lambda^F_i + \phi_i^t \lambda^V_i$ with $\phi_i \in [0,1]$, and with λ^F_i and λ^V_i being positive

and representing the fixed and variable parts of the scaling parameter. That is, we allow for the noise in decision making to decrease with experience. Finally, ε_{it}^a is an idiosyncratic error term assumed to have a type I extreme value function. Given the distributional assumption on the error terms, this gives rise to the usual logistic form for

the probabilities:
$$p_{it}^a = \frac{\exp\left(\frac{1}{\lambda_{it}} U_{it}^a\right)}{\exp\left(\frac{1}{\lambda_{it}} U_{it}^D\right) + \exp\left(\frac{1}{\lambda_{it}} U_{it}^C\right)}$$
. Thus the parameters to be estimated

are β_{it}^D , β_{it}^C , θ_i , λ^F_i , λ^V_i and ϕ_i .

Table A2.9 shows the averages and medians of the estimated parameters from the learning model. P(C) is the belief of cooperation in the first round of the first repeated game implied by the beta parameters. All P(C) is the belief of cooperation in the first round of the first repeated game implied by the beta parameters if we also add to the sample the subjects excluded from the estimation due to the fact that their round 1 decisions never change. λ^F is the estimated of the fixed noise parameter for all subjects. The next column shows the percentage of subjects for whom noise is constant across repeated games. Almost half of the subjects fall into this category. The columns for λ^V and ϕ show the estimates only for subjects who displayed decreasing noise as they gained experience.⁵ The last two columns show the estimates of noise in the first and fiftieth repeated games. There is a large decrease in the median total noise parameter as the subjects gain experience.

To better understand the evolution of cooperation in the simulations using the estimates from the learning model it is useful to study the simulated distribution of cooperation by treatment. Figure A2.4 shows for each treatment and repeated games 1 and 1000 the proportion of all the simulated sessions that have a given number of subjects (out of 14) choosing to cooperate in the first round. Figure A2.4 also shows the limit of the basins of attraction of AD versus TFT or G for each treatment: for example under $\delta=3/4$ & $R=48$ if the subject expects that 3 or more of the subjects in the session

⁵ Note that for subjects that displayed decreasing noise as they gained experience there is a multiplicity of parametrizations that would result in the same level of constant noise.

play the cooperative strategy then cooperation is the best response and defection is the best response otherwise.

The distributions of the number of cooperative actions are unimodal for every treatment in repeated game 1 and this mode is always located in the interior (from 4 cooperative actions in $\delta=1/2$ & $R=32$ and $\delta=3/4$ & $R=32$ to 8 in $\delta=3/4$ & $R=48$). The distribution of cooperative actions is quite different in repeated game 1000. In the treatment in which cooperation cannot be supported in equilibrium ($\delta=1/2$ & $R=32$) the mode and median is full defection. In the treatments under which cooperation can be supported in equilibrium but is not risk dominant, the mode decreases with experience and converges to 1 cooperative action over 14 in a session. In the treatments in which cooperation is both an equilibrium action and risk dominant, the results are diverse. In the case of $\delta=1/2$ & $R=48$ the distribution in repeated game 1000 is bimodal, with modes in 1 and 11 cooperative actions over 14. This bifurcation in the evolution of cooperation resembles the continental divide results from the coordination games literature (see Van Huyck et al. 1997). For the other two treatments in this group ($\delta=3/4$ & $R=40$ and $\delta=3/4$ & $R=48$) the distribution moves to higher levels of cooperation with experience with an extreme result for $\delta=3/4$ & $R=48$ where the mode and median converge to full cooperation.

Note that in the previous simulations we used the parameter estimates of a given treatment to simulate behavior in that treatment. In what follows we use estimates from other treatments, including beliefs, to simulate behavior in a given treatment.⁶ The results are presented in Figure A2.5. The simulations based on the estimated learning model track well the evolution of cooperation observed in the data in most treatments. The performance is clearly worse than when we consider estimates from the same treatment. First, note that for every treatment in which cooperation is greater in the last repeated game than in the first repeated game the same is true for the simulations. Second, of the four treatments for which cooperation is lower in the last repeated game than in the first repeated game, in three the same is the case for the simulated data. The exception is the treatment with $\delta=1/2$ & $R=48$, where the simulated data has a clear positive trend. Third,

⁶ These simulations are interesting and suggestive but do not provide a tight test of the learning model since there is no reason to expect that initial beliefs should be the same across treatments. Also note that only subjects that did not always play the same action are considered in this exercise.

the simulated evolution of cooperation the experimental data tend to be close or within the 90% interval generated by the simulations. Finally, for the range of repeated games for which we have experimental data from all three sessions, the average level of cooperation in the simulations has differences of 16% with the average observed cooperation rate in the experiments.

Table A2.10 indicates the fraction of behavior consistent with the strategies always defect (AD), always cooperate (AC), grim (G) and tit-for-tat (TFT) in all matches and matches that lasted at least 2 rounds.⁷ Note that AD and any of the other strategies we consider are mutually exclusive but AC, G and TFT are all consistent with a subject that cooperates in every round if the other player always cooperates as well. The Random Baseline column indicates what fraction of matches would have been classified as either AD or TFT if decisions were random; with 50-50 chances of either playing cooperate or defect.⁸

Focusing on AD and TFT allows us to explain an important part of the data (78% of matches with at least 2 rounds and 88% of the entire data set). This percentage increases with experience (see the last panel of Table A2.10) and is always higher than the random baseline. This is not to say that these strategies capture exactly what all subjects are doing, but rather that the evidence supports the idea that the data can be parsed in two broad categories: subjects who always defect and others who use strategies that start by cooperating, and condition current decisions on past histories, in a way that allows to support cooperation.

Next, we study how subjects changed strategies from repeated game to repeated game. Table A2.11 shows the transition probabilities by treatment when we organize the observed behavior in those consistent with AD, TFT and the rest. In all treatments, subjects that choose AD are more likely to choose AD again in the next repeated game. That is also the case for those choosing TFT when cooperation is an equilibrium action. When cooperation is not an equilibrium action, subjects are likely to abandon TFT in favor of AD. Moreover, the likelihood that TFT subjects would choose TFT again is

⁷ The reason to show matches with two rounds or more is that in matches with only one round all behavior is trivially consistent with one of the strategies we consider: defection is consistent with AD and cooperation with AC, G and TFT.

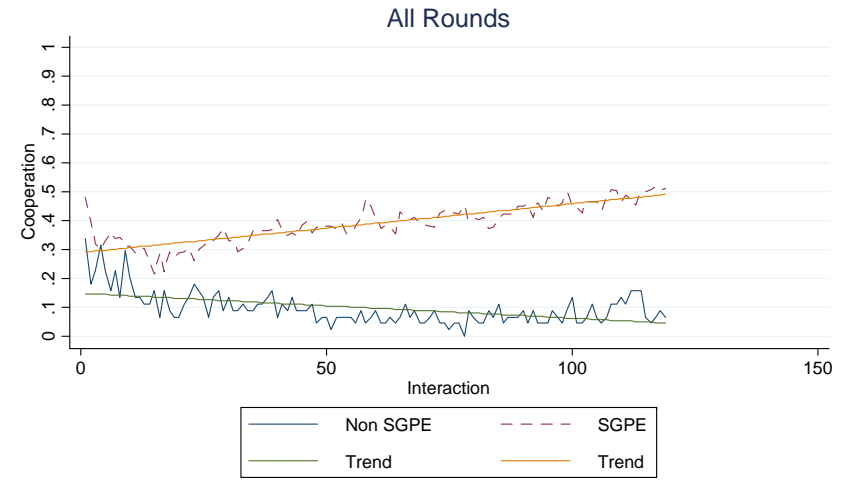
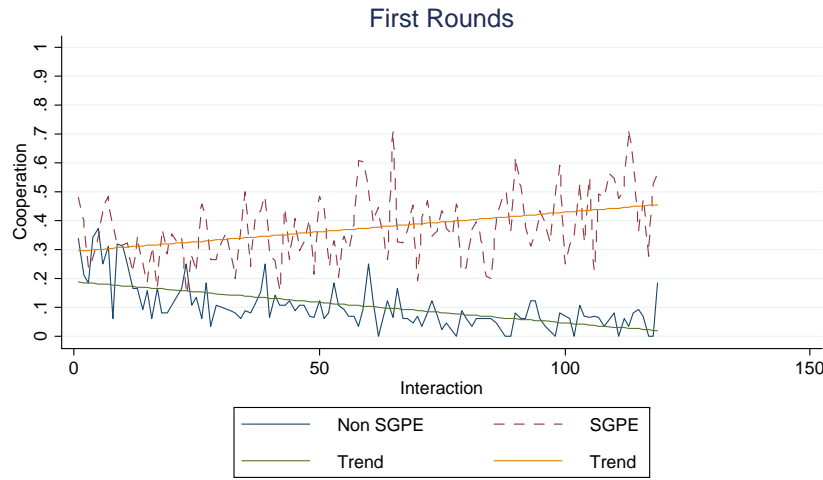
⁸ The random baseline is useful since many sequences of decisions could coincide with AD and TFT by chance.

increasing in the continuation probability and the payoff from cooperation. Finally, the popularity of TFT is decreasing with the size of the basin of attraction of AD.

These transition matrices allow us to compute the limit distribution over the three types of strategies (see Table A2.12). We find that these limit distributions are generally consistent with the observed prevalence of cooperation at the end of the sessions (see Table A2.5) and they are consistent with the long run simulation using the learning model but with a completely different approach. While AD is the most prevalent strategy in the limit distribution when cooperation cannot be supported in equilibrium, this is also the case in some of the treatments under which cooperation can be supported in equilibrium (similarly for risk dominance). It is also interesting to note that limit distributions may include a high share of both AD and TFT (see $\delta=1/2$ & $R=48$). This suggests that behavior may stabilize away from full coordination under some conditions.

Figure A2.1: Evolution of Cooperation and Linear Trends

By Equilibrium Condition



By Risk Dominance

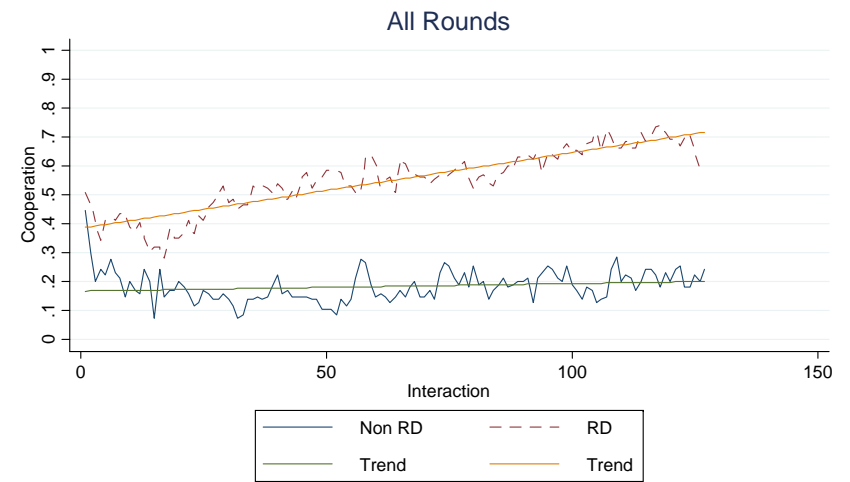
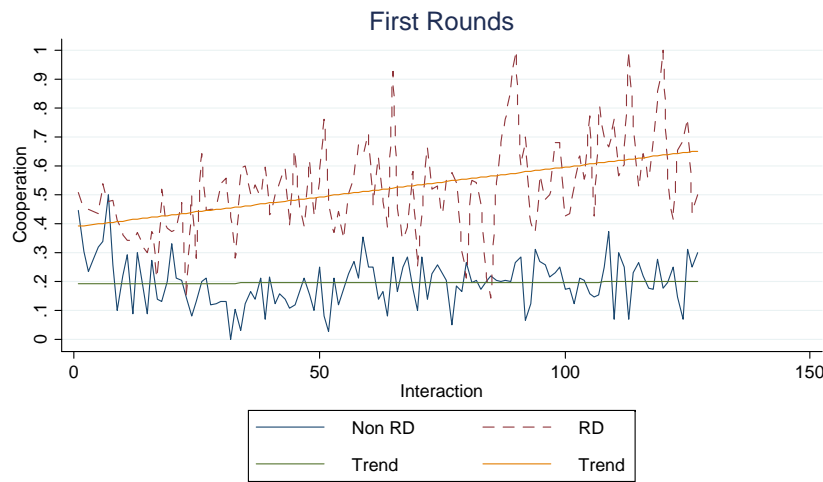


Figure A2.2: Evolution of Cooperation and Linear Trends by Treatment (Rounds 2-4)

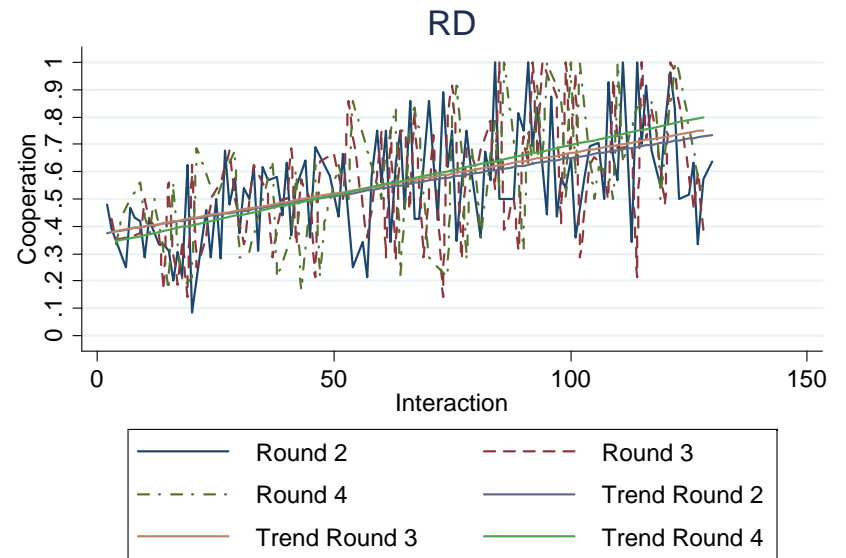
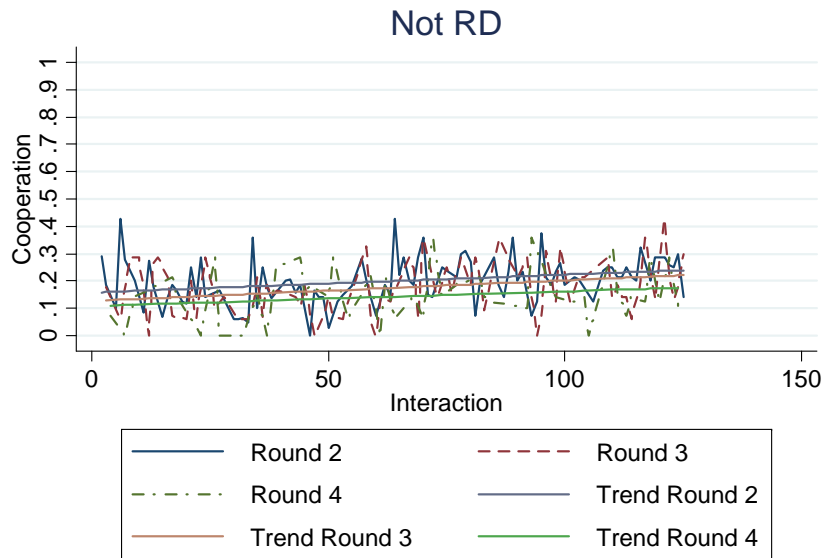
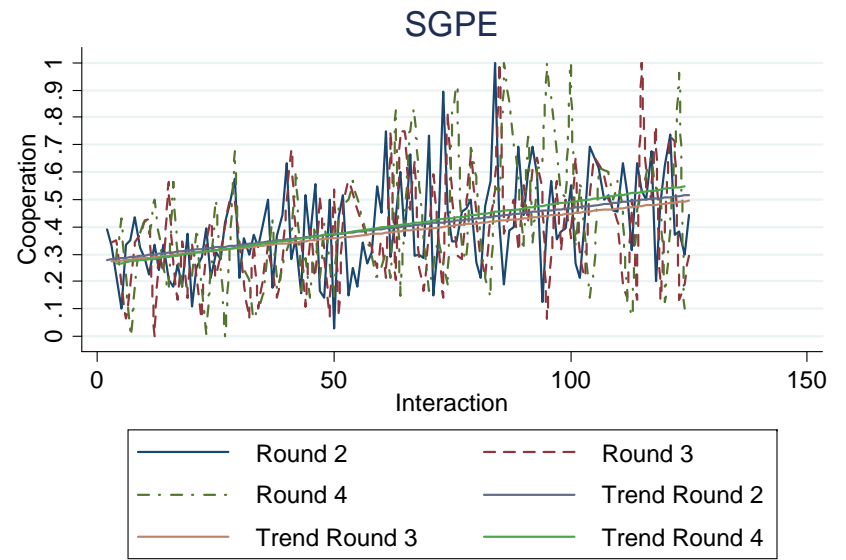
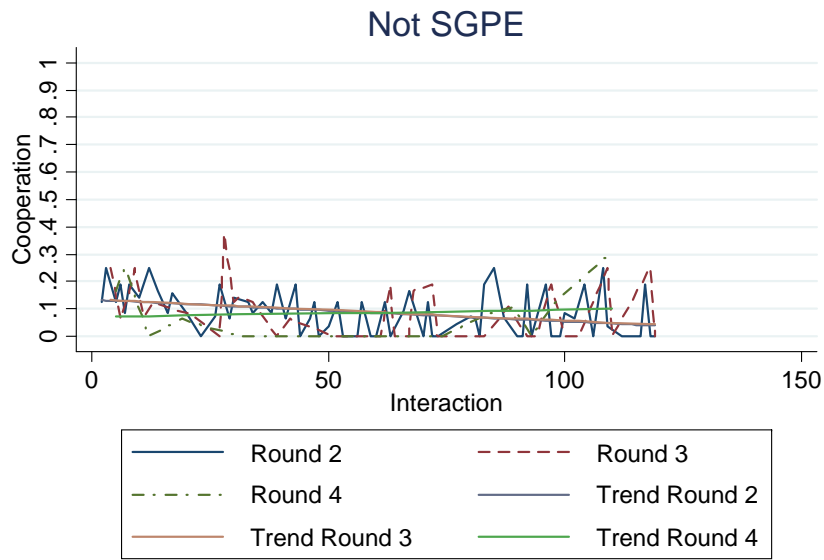


Figure A2.3: Evolution of Cooperation and Linear Trends by Treatment and Session (first rounds)

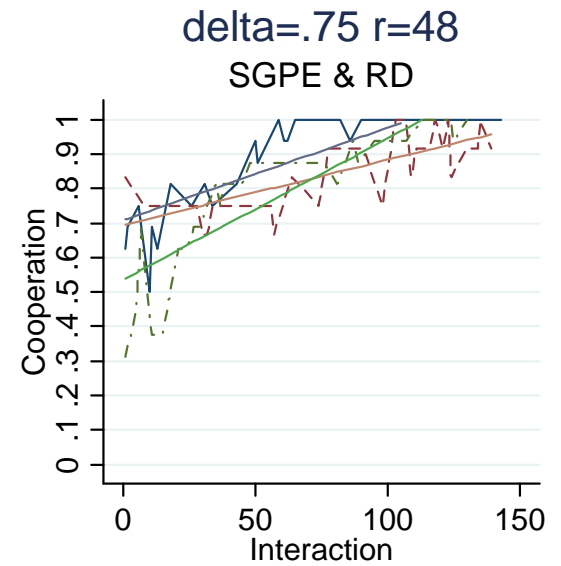
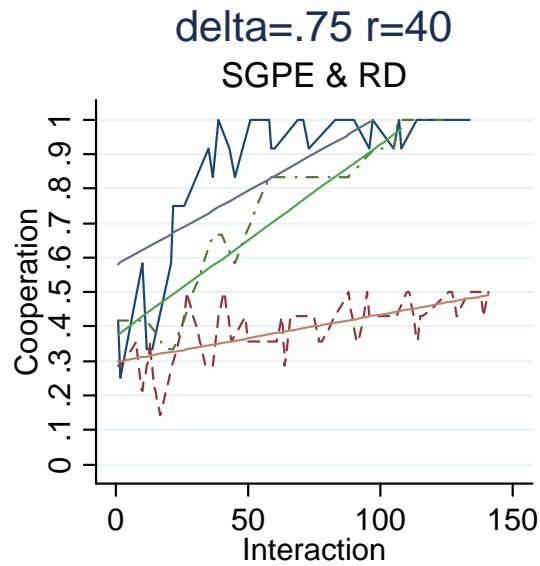
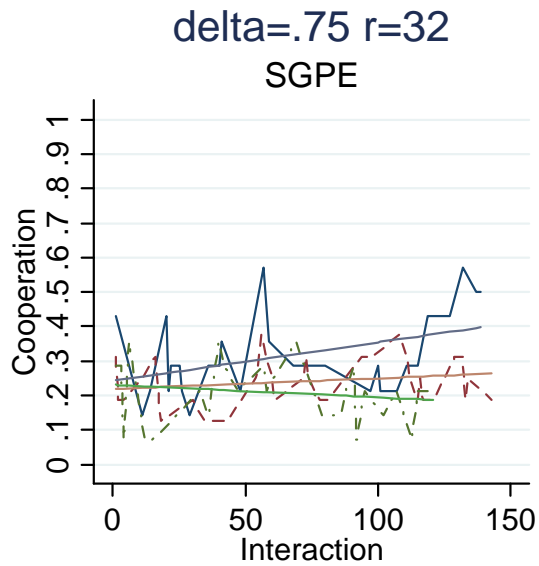
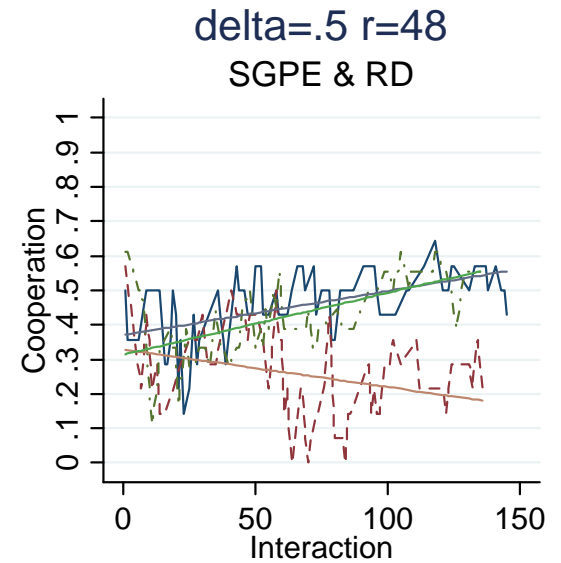
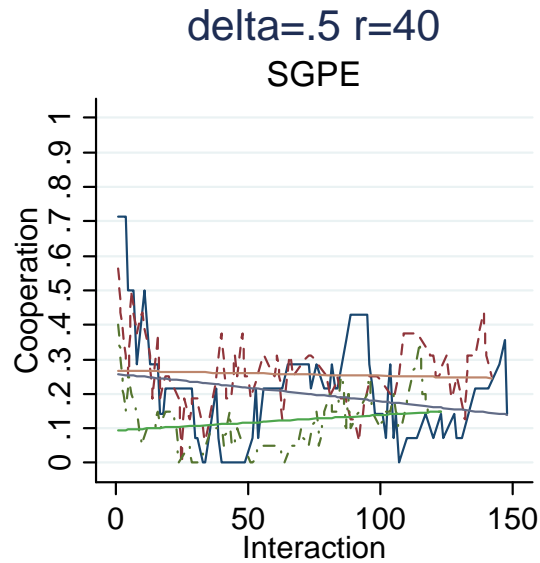
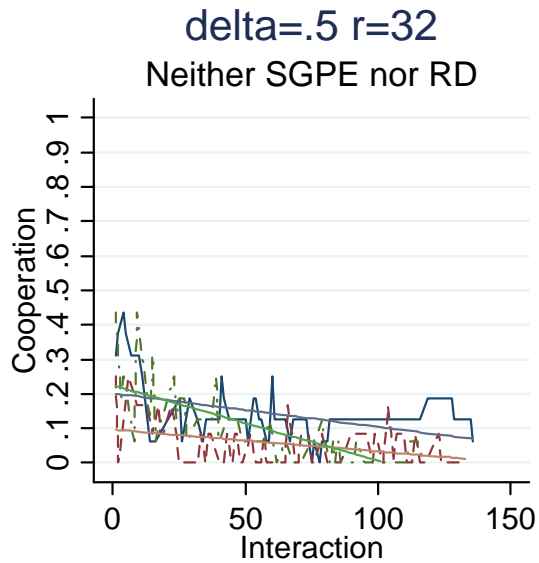


Figure A2.4: Distribution of Outcomes in Simulated Sessions

Repeated Game 1 (dashed) vs. Repeated Game 1000 (solid). Vertical lines denotes limit of basins of attraction.

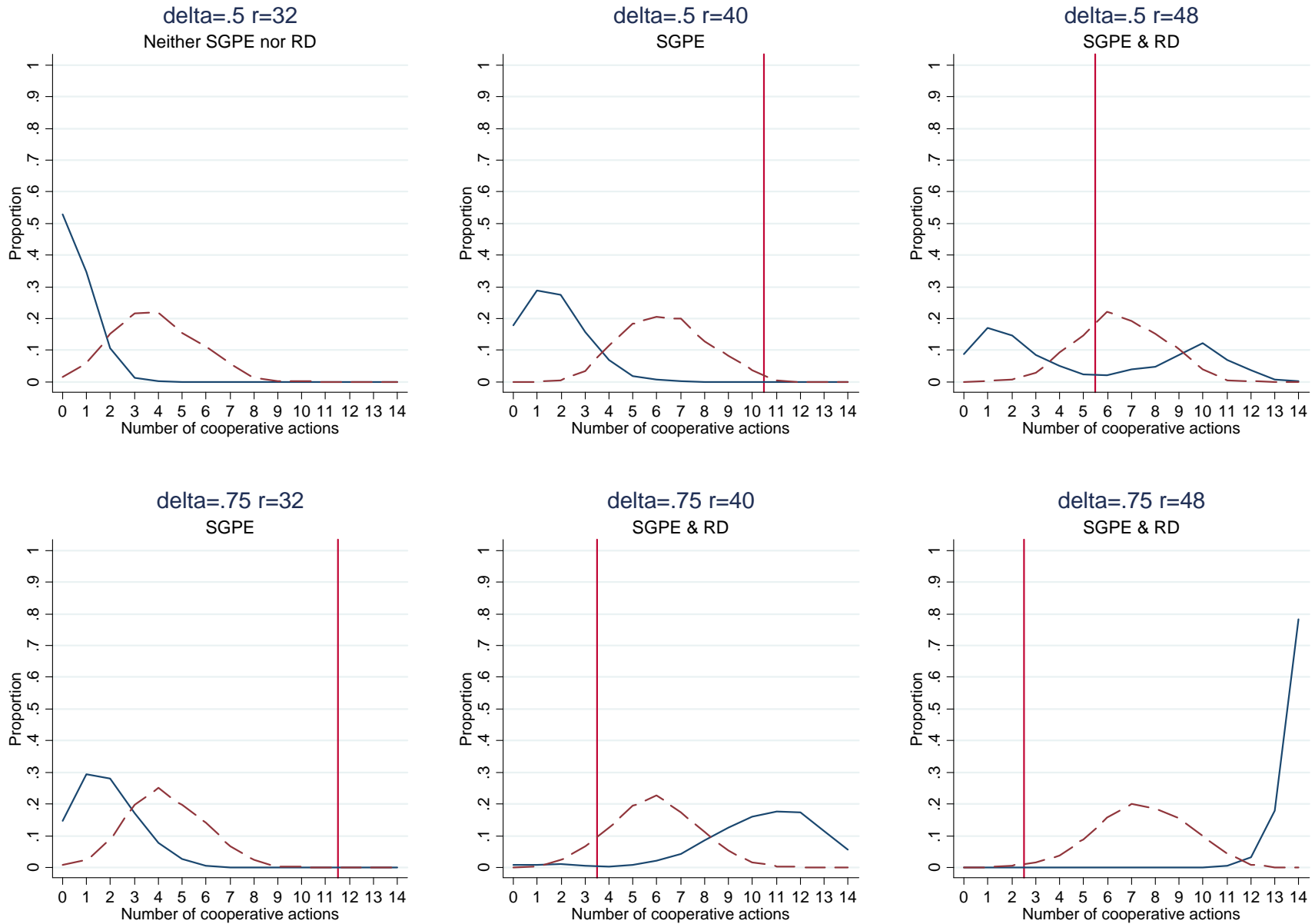


Figure A2.5: Simulated Evolution of Cooperation Implied by the Learning Estimates from other Treatments

