

Constructing Strategies in the Indefinitely Repeated Prisoner's Dilemma Game*

March 2016

Julian Romero[†] Yaroslav Rosokha[‡]

Abstract:

We propose a new approach for running lab experiments on indefinitely repeated games with high continuation probability. This approach has two main advantages. First, it allows us to run multiple long repeated games per session. Second, it allows us to incorporate the strategy method with minimal restrictions on the set of strategies that can be implemented. This gives us insight into what happens in long repeated games and into the types of strategies that subjects use. We report results obtained from the indefinitely repeated prisoner's dilemma with a continuation probability of $\delta = .99$. We find that subjects respond to higher incentives to cooperation in both the cooperation level and the strategies used. When we analyze the constructed strategies, our results are largely similar to those found in the literature under shorter repeated games—specifically that the most common strategies are those similar to Tit-For-Tat and Always Defect.

Keywords: Indefinitely Repeated Games, Prisoners' Dilemma, Experiments, Cooperation, Experimental Design, Strategies

[†] Eller College of Management, University of Arizona • Email: jnromero@gmail.com

[‡] Krannert School of Management, Purdue University • Email: yrosokha@purdue.edu

* This paper benefited greatly from discussions with and comments from Pedro Dal Bo, Tim Cason, Guillaume Frechette, Ernan Haruvy, Christos Ioannou, Dale Stahl, Emanuel Vespa, Alistair Wilson, and Huanren Zhang, as well workshop participants at the 2014 North American ESA meetings in Ft. Lauderdale, the World ESA meetings in Sydney and seminar participants at Purdue University, the University of Pittsburgh, Penn State University and Texas A&M University.

1 Introduction

The repeated prisoner’s dilemma has been used as a stylized setting to model a wide variety of situations across many disciplines (Cournot competition, advertising, public good provision, arms races, evolution of organisms, etc.). Because of this breadth, the repeated prisoner’s dilemma is one of the most commonly studied games in all of game theory, as researchers try to gain a better understanding of how and when cooperation emerges. In this paper, we run experiments on the indefinitely repeated prisoner’s dilemma game using an innovative experimental interface that allows subjects to directly construct their strategies in an intuitive manner and to participate in “very long” indefinitely repeated prisoner’s dilemmas (continuation probability $\delta = 0.99$). We use this environment to gain a unique perspective on the strategies that subjects use in the indefinitely repeated prisoner’s dilemma and on the factors that make subjects cooperate.

Our experimental design uses the strategy method (Selten 1967), and allows subjects to construct strategies in an intuitive manner. A player constructs a strategy by developing a set of rules. Each rule is an “if this, then that” statement, which contains an input and an output. The input to a rule is a list of $n \geq 0$ action profiles, while the output of a rule is the action that will be played after the input has occurred. Our design ensures that in any period of a repeated game, the set of rules for a player prescribes a unique action to be played in that period. In contrast to standard indefinitely repeated games experiments, in which players directly choose an action in each period, our design allows players’ actions to be chosen automatically using the rules in the rule set. The repeated game is then divided into two stages: the free stage and the lock stage. In the free stage, each period lasts two seconds, and players are able to learn and adjust their rule set as the repeated game progresses. In every period of the free stage, there is a certain probability that the repeated game will enter the lock stage. In the lock stage, players’ rule sets are locked, and each period lasts less than two seconds. Subjects then play a number of repeated games, so that they are able to learn both within a repeated game and across repeated games.

This experimental design offers several benefits over standard indefinitely repeated games experiments. First, we can directly view players’ strategies. There is a growing literature that aims to better understand the strategies played in repeated prisoner’s dilemma games (Dal Bó and Fréchette, 2011; Fudenberg, Rand, and Dreber, 2012; Stahl, 2013). Typically, players choose actions that are then used to make inferences about the strategy that the players are actually using. These inferences can be difficult since multiple strategies can lead to the same realization of actions. Additionally, these inferences require researchers to specify a set of strategies to begin with. The set of memory-1 and memory-2 strategies is manageable, and, arguably, is appropriate for short repeated games. However, this may be too restrictive when studying longer repeated games. In our design, players directly construct strategies via the rule sets with minimal restrictions on the types and lengths of strategies. This allows us to directly view each player’s complete strategy. We then can determine the extent to which the typically assumed set of strategies is appropriate for our setting.

Second, this experimental design allows us to run long indefinitely repeated games. Indefinitely

repeated games are implemented in the lab by imposing a termination probability at the end of each period. One difficulty with this standard approach is that a single repeated game can last a very long time, even if the continuation probability is relatively small. A second difficulty is that subjects must know the length of time for which the experiment was scheduled, which means that they know that there must be some bound on the length of their interaction. For example, if subjects know that the experiment was scheduled for two hours, then they know that they cannot play a repeated game that lasts for more than two hours. Because of these difficulties, indefinitely repeated games in the laboratory have typically focused on situations with relatively low continuation probabilities. Our experimental design alleviates these problems because play proceeds automatically. In the free stage, each period lasts two seconds, and in the lock stage, each period get progressively faster until it lasts only one tenth of a second. This means that the lock stage can have hundreds of periods but still last less than one minute. Therefore, our design allows us to run long repeated games. Long repeated games give subjects the chance to learn within a repeated game and develop their strategies as play progresses. In addition, long repeated games are important for a broad class of macroeconomics experiments in which the underlying models rely on sufficiently high discount factors (Duffy, 2008).

We use this experimental design to conduct a detailed analysis of strategies used in the indefinitely repeated prisoner's dilemma game. The standard approach to estimating strategies from the actions played is to use a predefined set of strategies and run a maximum likelihood estimation to determine the frequency with which these strategies are used in the population. We find that the maximum likelihood estimates of strategies in these long repeated games are consistent with previous estimates for shorter versions of the game. Specifically, the most common strategies are memory-1 strategies such as Tit-For-Tat, Grim Trigger, and Always Defect. In addition to the standard strategy estimation, our experimental design allows us to view strategies directly. We run a cluster analysis on the strategies and find that the vast majority of subjects construct strategies that fall into three broad categories: those that almost always cooperate; those that almost always defect; and those that play Tit-For-Tat. One benefit of this cluster analysis is that it allows us to analyze the strategies without making any a priori assumptions on the set of strategies that will be used for the estimation.

We also find evidence of more-sophisticated strategies that may be overlooked when using standard estimation procedures with a predefined set of strategies. The most popular of these more-sophisticated strategies are of two types. The first type defects after a long sequence of mutual cooperation. This type of strategy may be used to check whether the other player's strategy can be exploited during a long interaction. The second type cooperates after a long sequence of mutual defections. This type of strategy may be used to signal to the other player to start cooperating after cooperation has broken down. We find that these more-sophisticated strategies play an important role in the sustainability of cooperation during a long interaction.

We run three main treatments of the experiment. These treatments allow us to compare our findings to the results in the existing literature. Specifically, we address the following questions:

1. Does the benefit for cooperation affect the types of strategies that arise and the level of cooperation in long interactions?
2. Does the strategy method matter for the inferred strategies and the observed cooperation levels (as compared to the direct-response method)?

We find that: 1) subjects respond to incentives for cooperation, as is evident from both their strategies and higher rates of cooperation; and 2) the strategy method approach results in reduced levels of cooperation.

The idea of asking participants to construct strategies is not new. Selten (1967) asked participants to design strategies based on their experience with the repeated oligopoly investment game. Axelrod (1980a, 1980b) ran two tournaments in which scholars were invited to submit computer programs to be played in a repeated prisoner’s dilemma. More recently, Selten et al. (1997) asked experienced subjects to program strategies in PASCAL to compete in a 20-period asymmetric Cournot duopoly. One of the contributions of our paper is to develop an interface that allows non-experienced subjects to create complex strategies in an intuitive manner.

Research on whether the strategy method itself yields a different behavior than the direct-response method has yielded mixed results. Brandts and Charness (2011) survey 29 papers that investigate similarities and differences in the outcomes of the strategy method versus the direct-response method for different games. The authors find that among 29 papers, 16 find no difference, four find differences, and nine find mixed evidence. The most relevant of these are Brandts and Charness (2000), Reuben and Suetens (2012), and Casari and Cason (2009). Brandts and Charness (2000) find no difference in cooperation rates between the direct-response and the strategy method in a single-shot prisoner’s dilemma game. Reuben and Suetens (2012) find no difference between the direct-response and strategy methods for a repeated sequential prisoner’s dilemma with a known probabilistic end when strategies are limited to memory-1 and condition on whether the period is the last one. Both of these studies, however, consider either a one-shot or a relatively short indefinitely repeated game ($\delta = 0.6$, 2.5 periods in expectation) with strategies limited to memory-1. Casari and Cason (2009) find that use of the strategy method lowers cooperation in the context of a trust game. The strategy method may cause subjects to think more deliberately. Rand, Greene, and Nowak (2012) and Rand, Peysakhovich, Kraft-Todd, Newman, Wurzbacher, Nowak, and Greene (2014) find evidence that making people think deliberately lowers their cooperative behavior in the prisoner’s dilemma, repeated prisoner’s dilemma, and public-goods provision games. Our experimental results suggest that the strategy method leads to lower levels of cooperation in long repeated games.

Our work builds on recent experimental literature that investigates cooperation in the indefinitely repeated prisoner’s dilemma (Dal Bó and Fréchette, 2011; Fudenberg, Rand, and Dreber, 2012; Dal Bó and Fréchette, 2015) and the continuous-time prisoner’s dilemma (Friedman and Oprea, 2012; Bigoni, Casari, Skrzypacz, and Spagnolo, 2015). Dal Bó and Fréchette (2011) study repeated games ranging from $\delta = 0.5$ to $\delta = 0.75$ (two to four periods in expectation). They find that cooperation increases as both the benefit for cooperation and the continuation probability are

increased. Fudenberg, Rand, and Dreber (2012) study repeated prisoner’s dilemma when intended actions are implemented with noise, and the continuation probability is $\delta = 0.875$ (eight periods in expectation). They find that while no strategy has a probability of greater than 30% in any treatment, a substantial number of players seem to use the “Always Defect” (ALLD) strategy. The authors also find that many subjects are conditioning on more than just the last round. Dal Bó and Fréchette (2015) ask subjects to design memory-1 strategies that will play in their place and play in games with probability of continuation of $\delta \in \{0.50, 0.75, 0.90\}$. Additionally, they implement a treatment whereby subjects choose from a list of more complex strategies. They find no effect of the strategy method on levels of cooperation, and evidence of Tit-For-Tat, Grim-Trigger and AllD. Our work differs in several important dimensions. First and foremost, we consider very long repeated interactions ($\delta = 0.99$, 100 periods in expectation), for which the behavior and the types of strategies might be vastly different from those in shorter games. Second, we do not limit the strategy construction to memory-1 strategies or a predefined set of strategies. Third, instead of explicitly asking subjects to design the full strategy, we incentivize them to build their strategies through implementation of the lock stage.

The setting of our paper is approaching the continuous time prisoner’s dilemma studied by Friedman and Oprea (2012) and Bigoni, Casari, Skrzypacz, and Spagnolo (2015). Friedman and Oprea (2012) find high levels of cooperation in the continuous-time setting with a finite horizon. However, when the time horizon is stochastic, Bigoni, Casari, Skrzypacz, and Spagnolo (2015) find that cooperation is harder to sustain than when the game horizon is deterministic.

The rest of the paper is organized as follows: Section 2 presents details of our experimental design. Section 3 presents the results. Section 4 provides simulations which give context to some of the main results for the paper. Finally, in Section 5, we conclude.

2 Experimental Design

The experiment consists of fifteen matches. At the beginning of each match, participants are randomly paired and remain in the same pairs until the end of the match. The number of periods in each match is determined randomly using continuation probability $\delta = .99$. To ensure a valid comparison, the same sequence of supergame lengths is used in every session. Since there are no participant identifiers within the game interface, participants remain anonymous throughout the experiment. Neutral action names, Y and W , are used throughout the experiment instead of C and D .

2.1 Game Payoffs

	C	D
C	R R	12 50
D	50 12	25 25

Figure 1: Stage Game Payoffs

We use a parameterization of the stage game similar to that in Dal Bó and Fréchette (2011), which is displayed in Figure 1. We use $R = 38$ for our benchmark treatment and $R = 48$ for the high benefit to cooperation treatment.

2.2 Rules

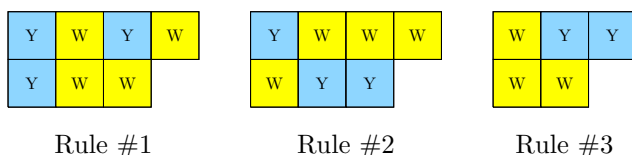


Figure 2: Examples of Rules

Rather than directly making choices, subjects develop a set of rules that automatically makes choices for them. Each rule consists of two parts: i) *Input Sequence* - a sequence of action profiles; and ii) *Output* - an action to be played by the subject after the input sequence has occurred. Figure 2 displays some examples of rules. For example, Rule #1 has an input sequence of (Y,Y) , (W,W) , (Y,W) and an output of W (to simplify notation, we denote such a sequence as $YYWWYW \rightarrow W$). This means that if the subject plays Y , W , and then Y in the last three periods, and the participant that he is paired with plays Y , W , and then W , this rule will play W in the next period. The length of the rule is measured by the length of the input sequence. Thus, rule #1 and rule #2 have a length of 3, and rule #3 has a length of 2.

At the beginning of the experiment, subjects have no rules in their set. They are given the opportunity to add new rules to the set using the rule constructor.

The rule constructor is displayed in the screenshot in Figure 12 at the bottom center of the screen. Subjects can construct rules of any length by clicking the boxes in the rule constructor. Once they have constructed a complete rule, a button will appear that says “Add Rule,” which they can click to add the rule to the set. Note that it is not possible to have two rules with the same input sequence but different outputs. If subjects create a rule that has the same input sequence but a different output, then they get a message that says “Conflicting rule in set,” and a button that says “Switch Rule” appears. If subjects press this button, it will delete the conflicting rule

from the rule set, and add the rule from the constructor. Subjects can also delete any rule from the set by clicking the red delete button next to the rule. See Appendix A for more details on the experimental instructions and the experimental interface.

2.3 Match-play

23	Period	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45
	My Choice		W	Y	Y	Y	W	W	Y	Y	Y	W	Y	W	Y	W	Y	Y	Y	W	Y	?
	Other's Choice		W	Y	W	W	W	W	Y	W	Y	Y	Y	Y	W	Y	Y	Y	Y	W	W	?

Figure 3: Example of Game History

As the match progresses, subjects see the history of play across the top of the screen (Figure 3). A rule of length n is said to *fit* the history if the input sequence matches the last n periods of the history. For example, since the last three periods of play in the above history (periods 42-44) have been (Y,Y), (W,W) and (Y,W), and that sequence is also the input for rule #1, then rule #1 is said to fit the history. Similarly, given the above history, rule #3 fits the history, but rule #2 does not.

If more than one rule fits the history, then the rule with the longest length will determine the choice. For example, given the history in Figure 3, since both rule #1 and rule #3 fit the history, rule #1 will be used to make the choice since it is longer. Therefore, given the history, and the three rules in Figure 2, the choice next period will be W, as prescribed by rule #1. If no rules fit the history, then the “default rule” will be selected. The default rule is a memory-0 rule, which is set prior to start of the first match and can be switched at any time during the free stage of a supergame or between supergames. It is important to note that since the default rule will always be defined, the rule set makes a unique choice in each period.

We choose to have the longest-length rule determine the action when more than one rule fits the history for two reasons. The first reason is to encourage the use of rules as opposed to direct-responding. For example, if the shorter rule determined the choice, then the memory-0 rule would always be selected, in which case subjects would not need to construct any rules. The second reason is that any procedure that selects a shorter rule over a longer rule precludes the longer rule from ever being played, which is equivalent to not having the longer rule in the set.

During each supergame, a choice is made from the rule set automatically in each period (every two seconds). When the choice is made, the rule that was used to make the choice, the corresponding sequence in the history, and the corresponding choice on the payoff table are all highlighted.

2.4 Lock Stage

Instead of relying on rules, subjects may choose to use the default rule to make their choices directly. In such a case, our environment would be equivalent to a (near) continuous-time version of the standard repeated prisoner’s dilemma experiment. This means that without an additional

incentive to use the rules, subjects might not use the rules at all. In order to incentivize subjects to use rules, we distinguish between two stages of the match: the free stage and the lock stage. Subjects are free to edit their set of rules during the free stage but not during the lock stage. In each period of a supergame, there is a 1% chance that the supergame enters the lock stage (if it already hasn't), a 1% chance that the supergame ends, and a 98% chance that the supergame continues for another period.

Although subjects cannot edit the set of rules in the lock stage, they can still observe all of the actions taken by their rule set and their opponent's rule set. Since neither player can change his rule set in the lock stage, play converges to a deterministic sequence of action profiles that is repeatedly played until the lock stage finishes. Since this sequence is deterministic, players do not need to watch it get played repeatedly. Therefore, to expedite the lock stage, we gradually increase the speed of each period from two seconds per period at the beginning to 0.1 seconds per period at the end. This allows us to run longer interactions in a shorter period of time. This is important, from a practical perspective, for carrying out long repeated games in the lab.

2.5 Administration and Data

One hundred and six students were recruited for the experiment using ORSEE software (Greiner, 2004) on the campus of Purdue University. Six sessions of the experiment were administered between September and December of 2014, with the number of participants varying between 16 and 18. The experimental interface was developed by the authors using a Python server and JavaScript clients.¹

Upon entering the lab, the subjects were randomly assigned to a computer and given a handout containing the instructions (see Appendix A for instructions). After all of the subjects had been seated, recorded video and audio instructions were played on a projector at the front of the lab and also displayed on each computer terminal.² Then, subjects completed a ten-question quiz to make sure that they understood the format of the game and the environment (see Appendix B for quiz details). The questions on the quiz were monetarily incentivized as follows: each participant would earn \$0.50 for answering a question on the first attempt, \$0.25 for answering a question on the second attempt, and \$0.00 for answering a question on the third or later attempt. After the second incorrect attempt, the subjects were given an explanation to help them answer the question.

The experiment did not start until all of the subjects had correctly answered all of the quiz questions. Each session, including the instructions portion and the quiz, took about one hour and 20 minutes to complete, with an average earnings from the experiment of \$17.50. The breakdown of the three treatments are presented in Table 1. To summarize: the *R38* treatment is our baseline; the *R48* treatment is used to identify the effect of increased cooperation payoffs; finally, the *D38* treatment explores the effect of the strategy method in a long indefinitely repeated prisoners' dilemma.

¹A demo of the interface is available at <http://jnromero.com/StrategyChoice/game.html?viewType=demo>

²Link to video instructions: <http://jnromero.com/StrategyChoice/instructions.mp4>

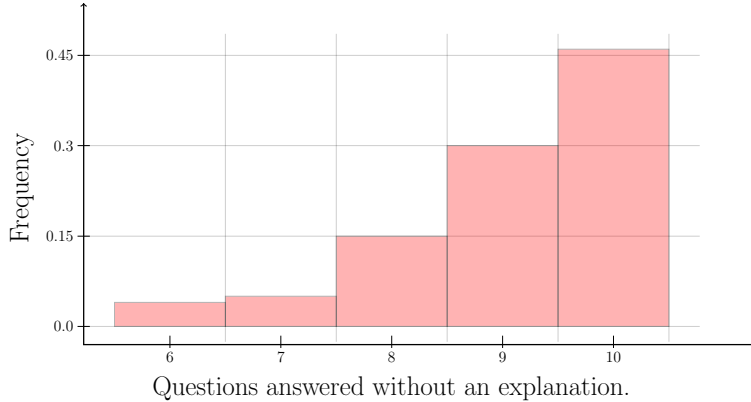


Figure 4: Quiz Performance. *Notes:* There were ten questions in total. Subjects were provided an explanation of the correct answer if they incorrectly answered a question more than two times.

Treatm.	R	Rules	N	Sessions	Description
<i>R38</i>	38	Yes	36	2	Baseline
<i>R48</i>	<u>48</u>	Yes	36	2	Increased payoff to cooperation
<i>D38</i>	38	<u>No</u>	34	2	Direct-response (no rules)

Table 1: Treatment summary.

3 Experimental Results

This section is organized as follows: in Section 3.1, we present the results on aggregate cooperation observed during our experiment. In Section 3.2, we analyze individual rules that were constructed during our experiment. In Section 3.3, we focus on rule sets and the resulting strategies. Specifically, we use cluster analysis to investigate which strategies participants constructed. In Section 3.4, we estimate strategies from the observed actions using the maximum likelihood approach. And finally, in Section 4.2, we evaluate the performance of the commonly used strategies against the strategies used in the experiment.

3.1 Cooperation

The fundamental question in the repeated prisoner’s dilemma literature is: what factors lead to higher cooperation rates? One of our contributions is to study this question in a setting of an indefinitely repeated game with a very high continuation probability ($\delta = .99$). Figure 5 presents the average cooperation during the free stage over the last five supergames within each of the three treatments.

Several results are worth noting. First, the cooperation in the *R48* treatment is higher than in the *R38* treatment when we look at all periods within a supergame. Second, the cooperation in

Treatm.	Periods				
	First	First.4		Last.4	All
<i>R48</i>	0.578 (0.039)	0.496 (0.036)	**	0.428 (0.041) **	0.448 (0.037) **
<i>R38</i>	0.594 (0.033) **	0.461 (0.035) ***	***	0.317 (0.04) ***	0.352 (0.038) ***
<i>D38</i>	0.682 (0.032)	0.65 (0.038)	**	0.568 (0.049)	0.584 (0.04)

Figure 5: Average Cooperation in the Last Five Supergames. *Notes:* The unit of observation is a matched pair of subjects per supergame. Bootstrapped standard errors are in parentheses. Black stars denote significance using the permutation test. Red stars denote significance using the matched-pair randomization test. ** denotes significance at 5%. *** denotes significance at 1%.

the last four periods of the supergame is significantly lower than in the first four periods. Finally, the cooperation in the *D38* treatment is higher than in the *R38* treatment. The above differences are found to be significant using non-parametric permutation and randomization tests.^{3,4} In what follows, we elaborate on each of these results.

Result 1 *Cooperation increases with an increase in payoff to mutual cooperation.*

Consistent with the results on the shorter repeated prisoner’s dilemma ((Dal Bó and Fréchette, 2011), we find that increasing payoff to mutual cooperation increases the average cooperation within a supergame. Note that the cooperation early on is approximately the same between the *R38* and *R48* treatments, but the drop-off in the cooperation rate is much greater for the *R38* treatment (24%), than for the *R48* treatment (7%). Thus, we find that cooperation is easier to sustain in the *R48* treatment.

Result 2 *Cooperation decreases as a supergame progresses.*

Previous results suggest that higher δ would lead to more cooperation. However, we see that in long repeated games, cooperation often breaks down. We attribute this break down in cooperation

³Permutation tests do not rely on any assumptions regarding the underlying distribution of the data (Good, 2013). For the permutation tests carried out in this paper, we took all observations for the two considered cells and constructed a distribution of the difference in cooperation rates under the null hypothesis; we then used this distribution to test where the original difference lies. The unit of observation is a matched pair of subjects per supergame.

⁴The null hypothesis in the randomization test is that there is no difference in participants’ behavior (see, e.g., Good, 2006), which in our case is the difference between the average cooperation rate between the first four and the last four periods of the supergame. If this is true, then the observed cooperation across the four periods are equally likely to have come from either the first four or the last four periods. That is, there are two cases for each pair of participants: (i) their responses at the beginning and the end of each supergame are as selected; and (ii) their responses are flipped from first to last and from last to first. We sampled each pair of participants and selected either (i) or (ii) with equal probability. Then, we determined the difference in the cooperation rate between what are now labeled as the “first” four and the “last” four periods and found the average difference across all participants. We repeated this process 10,000 times and obtained the histogram of the average differences. To determine the p-value we considered where the actual realized average difference falls within this distribution.

to subjects intentionally defecting after a long sequence of mutual cooperation. This strategy may be useful to check whether the other player’s strategy can be exploited. This is especially useful in long repeated games when the potential gain from exploitation is high. This breakdown in cooperation is most pronounced in the *R38* treatment, in which these strategies are most prevalent.⁵ These rules are further analyzed in the Section 3.2.

Result 3 *Cooperation decreases with the strategy method.*

We find that cooperation rates in the *R38* treatment are lower than in the *D38* treatment.⁶ This is in contrast to results from the shorter repeated prisoner’s dilemma studies of Reuben and Suetens (2012) and Dal Bó and Fréchette (2015). There are three potential sources of differences in our experiment: i) long interactions; ii) available strategies; and iii) procedural factors. In particular, the expected number of periods in our repeated games is 100. In terms of strategies, we allow strategies that are longer than memory-1, while not restricting subjects to a set of predefined strategies. With regard to the procedure, the game proceeds automatically and has the lock stage to incentivize strategy construction.

The difference in cooperation rates between *R38* and *D38* is in line with the lower cooperation found by Casari and Cason (2009) when using the strategy method in a trust game. In addition, Rand et al. (2012) and Rand et al. (2014) provide evidence that making people think deliberately lowers cooperation in the prisoner’s dilemma, the repeated prisoner’s dilemma, and public-good provision games. Our interface provides a framework for subjects to construct strategies, which forces them to think deliberately, and, therefore, may lower cooperation as well.

3.2 Constructed Rules

In this section, we examine the rules that subjects construct and use. We look at the rule length distribution among subjects and also describe a commonly used class of long rules (memory 3+).

It is important to look not only at the rules that are used, but also at the other rules in the set. Even if some rules are frequently used, and others are rarely used, the presence of the latter can be important in enforcing the desired behavior of the former. For example, consider a subject with the rule set that implements the TFT strategy: $\{\rightarrow C, DD \rightarrow D, CD \rightarrow D\}$. If this set is matched against another cooperative strategy, then the default rule ($\rightarrow C$) is played all the time, but the longer rules ensure that it is not susceptible to exploitation.

In addition, it can be difficult to draw conclusions from specific rules. For example, the rule sets $\{\rightarrow C, DD \rightarrow D\}$ and $\{\rightarrow C, DD \rightarrow C\}$ both implement ALLC, even though they have the conflicting rules $DD \rightarrow D$ and $DD \rightarrow C$. Therefore, we do not provide a thorough analysis of shorter rules in this section. Instead, we defer the analysis of full strategies to Section 3.3.

⁵In Appendix C, we provide results from a robustness treatment, which show that subjects that do not understand the interface do not affect the level of cooperation.

⁶As a benchmark for comparison with the *D38* treatment, Bigoni, Casari, Skrzypacz, and Spagnolo (2015) find average cooperation of .669 in the long stochastic treatment of their continuous time repeated prisoner’s dilemma experiments.

Figure 6 presents the data on the distribution of rules constructed and used during our experiment. Panel (a) describes the distribution of the length of the rules that were used in different supergames. Panel (b) shows the average number of rules of a given length in the rule set at the beginning of each supergame. Panel (c) shows the distribution of the longest rule that each subject used. Finally, panel (d) show the distribution of the longest rule that each subject used.

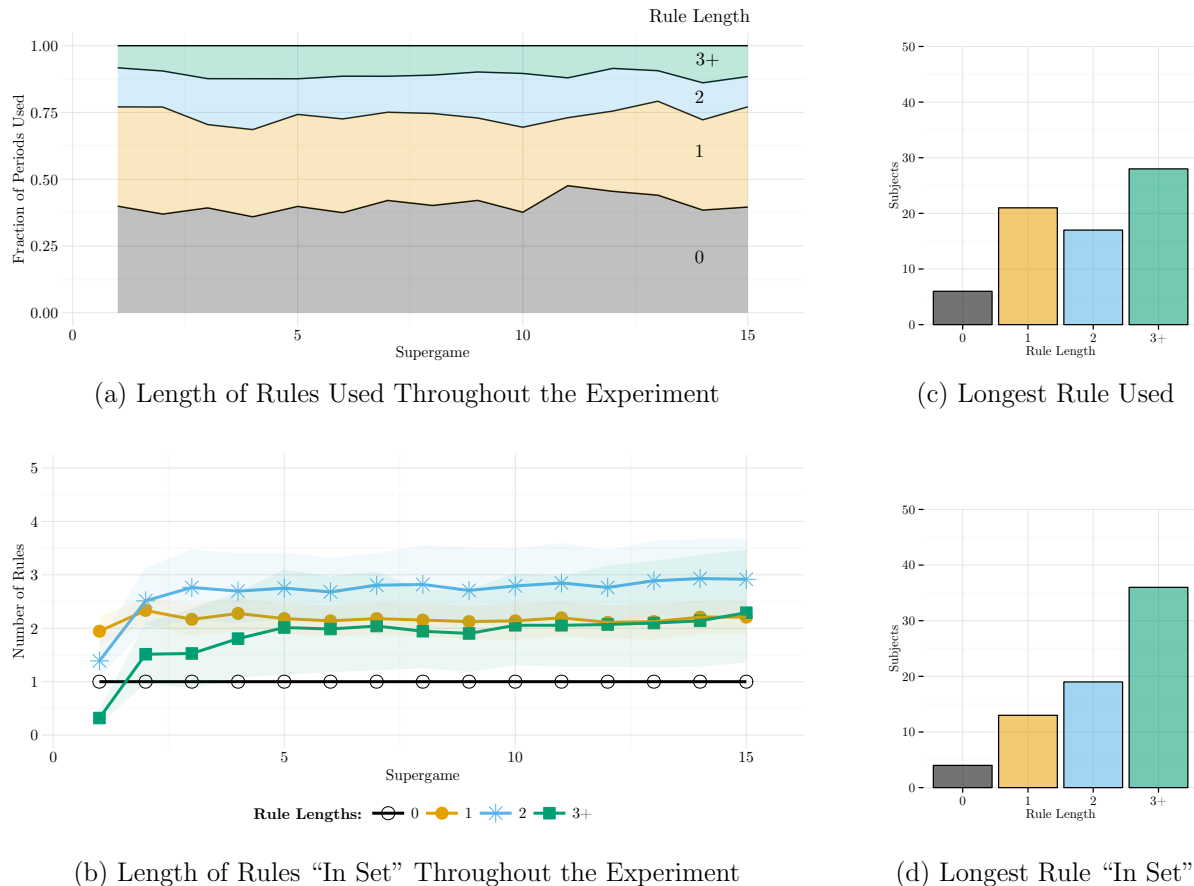


Figure 6: Rule Lengths. *Notes:* (a) Fraction of periods used within a supergame. (b) Average number of rules in the rule set at the beginning of a supergame. Bootstrapped 95% confidence intervals are superimposed. (c) Longest rule of length $l \in \{0, 1, 2, 3+\}$ used in supergame 15. For example, seven subjects (out of 72) relied only on rules of length 0 in supergame 15. (d) Longest rule of length $l \in \{0, 1, 2, 3+\}$ “in set” at the beginning of supergame 15.

Result 4 *Subjects construct and use rules longer than memory-1.*

We find that while the most-used rules are memory-0 and memory-1 rules, there is strong evidence of longer rules during the experiment. As mentioned above, even if the longer rules aren’t used, they may still play an important role in enforcing the behavior of the memory-0 and memory-1 rules that are used frequently.

Panel (a) shows that the subjects used memory-0 and memory-1 rules roughly 75% of the time and longer rules the other 25% of the time. Panel (b) shows that, on average, participants had one memory-0 rule,⁷ two memory-1 rules, three memory-2 rules, and two memory-3+ rules in any given period. Panel (b) also shows that the subjects’ rule sets converged quickly in terms of number of rules (though specific rules in the set may still have been changing). This indicates that time constraint was not an issue for rule construction in the experiment. Panels (c) and (d) show that several subjects used only the default rule, but the large majority of the subjects used longer rules.

As alluded to in the previous section, we find evidence of two types of long (memory-3+) rules. The first type is when a defection follows a long sequence of mutual cooperation. We will refer to this type as the *CsToD* rule. The second type is when cooperation follows a long sequence of mutual defections. We refer to this type as the *DsToC* rule. These rules do not match strategies that are currently studied in the literature (Dal Bó and Fréchette (2011), Fudenberg, Rand, and Dreber (2012)). Figure 7 presents information on the prevalence of this type of behavior. Column (a) shows the number of subjects that had each type of rule in their set at least once in the experiment. Column (b) shows the number of times that each sequence was observed during game play.

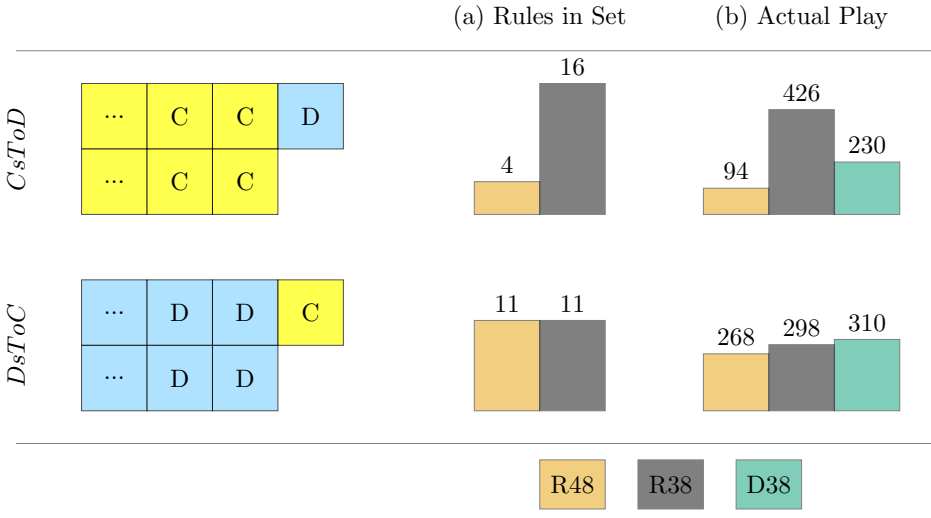


Figure 7: Memory 3+ Rules *Notes:* (a) Number of subjects that had each type of rule “in set” at least once throughout the experiment. (b) The number of times that D (C) is followed by at least three consecutive CCs (DDs) in the actual play. This uses only the actual play, and not information about the rules.

There are several important observations about these rules. First, the *CsToD* rule and behavior are much more common in the *R38* treatment than in the *R48* treatment. It is common in the *R38* treatment because it is an effective exploitative tool during long interactions. It is not as common

⁷Recall that subjects always had exactly one memory-0 (default) rule present in all treatments – that is why the average number of rules of length zero is exactly 1.

in the *R48* treatment, because the benefit from switching from C to D when the other subject plays C is relatively small (2 vs 12), which makes exploitative rules like this less desirable. Second, the *DsToC* rule and behavior are equally likely in both treatments. The *DsToC* rule could be used to ignite cooperation, which may be especially useful in long repeated games. This type of signaling is equally costly in both treatments, because the cost for switching from D to C when the other subject plays D is the same in both treatments (25-12=13). Finally, while we could make these observations from the actual game play, the presence of the rules highlights that this behavior is premeditated. Such premeditated behavior – in particular, *CsToD* – may hinder cooperation, which is consistent with Rand, Greene, and Nowak (2012).

The fact that there are more *CsToD* in the *R38* treatment aligns well with the results reported in Section 3.1. Specifically, the greater presence of *CsToD* in the *R38* treatment results in lower rates of cooperation, which consistent with **Result 1**. In addition, the presence of these rules in both treatments causes cooperation to decrease as the supergame progresses, which is in line with **Result 2**. Lastly, the greater presence of *CsToD* behavior in the *R38* treatment relative to the *D38* treatment leads to lower cooperation in *R38*, which is consistent with **Result 3**.

3.3 Constructed Strategies

As reported in the previous section, we find evidence that subjects use rules that do not fit within the typical set of strategies that are used in the literature on repeated games. In this section, we map the subjects’ complete rule sets to strategies. This analysis can be used to compare strategies that are observed in our experiments with those that are commonly studied in the literature.

Our interface allows subjects to construct a wide variety of rule sets. Because of this great variation, two rule sets may be identical on almost all histories, but not equivalent. For example, consider the rule sets $\{\rightarrow D\}$ and $\{\rightarrow D, DCDCDCDCDCDCDCDC \rightarrow C\}$. They are identical after every history that does not end with seven periods of (D, C) . So, though these rule sets are not equivalent, they lead to very similar behavior. Therefore, we need a measure of similarity of rule sets that would help to classify them into groups of similar strategies.

We determine how similar any two strategies are by comparing how similar their behavior is against a wide variety of opponents. If two strategies are very similar, then they should behave similarly against many different opponents. If two strategies are not very similar, then their behavior should be different against some opponents. More precisely, for each rule set, we generate a vector of actions played against a fixed set of opponents. The similarity of two rule sets is then determined by looking at the Manhattan distance, which, in this case, is the number of periods that the two rule sets give different outputs.

The “opponents” we consider are random sequences that are generated using a Markov transition probability matrix, P :

$$P = \begin{matrix} & C & D \\ \begin{matrix} C \\ D \end{matrix} & \begin{bmatrix} a & 1 - a \\ 1 - b & b \end{bmatrix} \end{matrix} \quad (1)$$

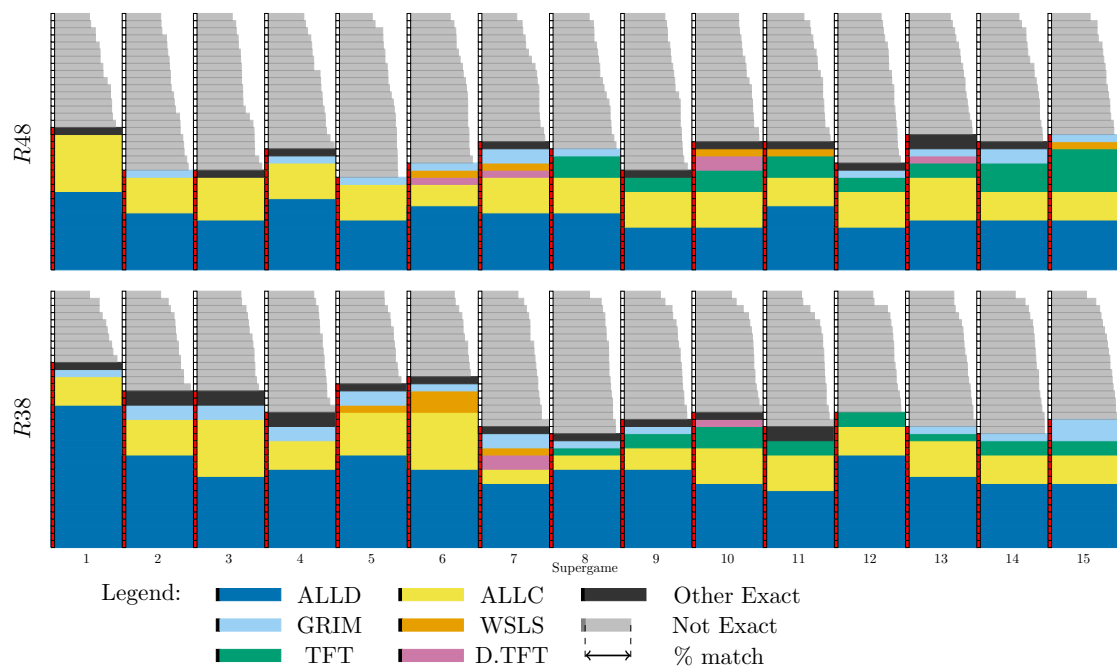


Figure 8: Exact Strategies At The Beginning of Each Supergame. *Notes:* After simulating a strategy against a fixed set of sequences we compare the resulting action sequence to an action sequence obtained from simulating each of the 20 predetermined strategies commonly used in MLE analysis. Percent match for “not exact” strategies is relative to the closest strategy among the 20.

where a and b are drawn randomly from a uniform distribution on $[0,1]$. For each a and b , a random sequence is generated, for a total of 100 sequences. Each sequence is of length 20 and starts in a state determined randomly, given the stationary distribution of P . We then simulate each of the participants’ strategies against these sequences, which generates 72 (36 for each treatment) vectors in $\{0,1\}^{20 \times 100}$. This process generates a wide variety of behaviors for the opponents.^{8,9}

We compare each of the participants’ strategies to the 20 strategies typically used in the MLE estimation (see, e.g., Fudenberg, Rand, and Dreber, 2012) in two ways. First, we determine whether any of the participants’ strategies match the play of one of the 20 strategies exactly (Figure 8). Second, we classify strategies into similar groups, and then compare those groups to the 20 strategies (Figure 9).

Result 5 *Half of the participants construct strategies that exactly match one of the 20 strategies typically used in MLE estimation for short repeated games.*

⁸If both a and b are low, then the sequence alternates between C and D . If both a and b are high, then the sequence is persistent, playing long sequences of C s and long sequences of D s. If a is high and b is low, then the sequence plays mostly C s with occasional D s. If a is low and b is high, then the sequence plays mostly D s with occasional C s. If both a and b are in the middle, then the sequence is playing C and D with approximately the same probability. This process allows us to consider a wide variety of behaviors without going through all possible histories.

⁹Possible alternatives for the opponents could be: i) to use realized sequences from the experiment, ii) use the subjects’ rule sets from the experiment. These alternatives don’t change the qualitative results of the paper.

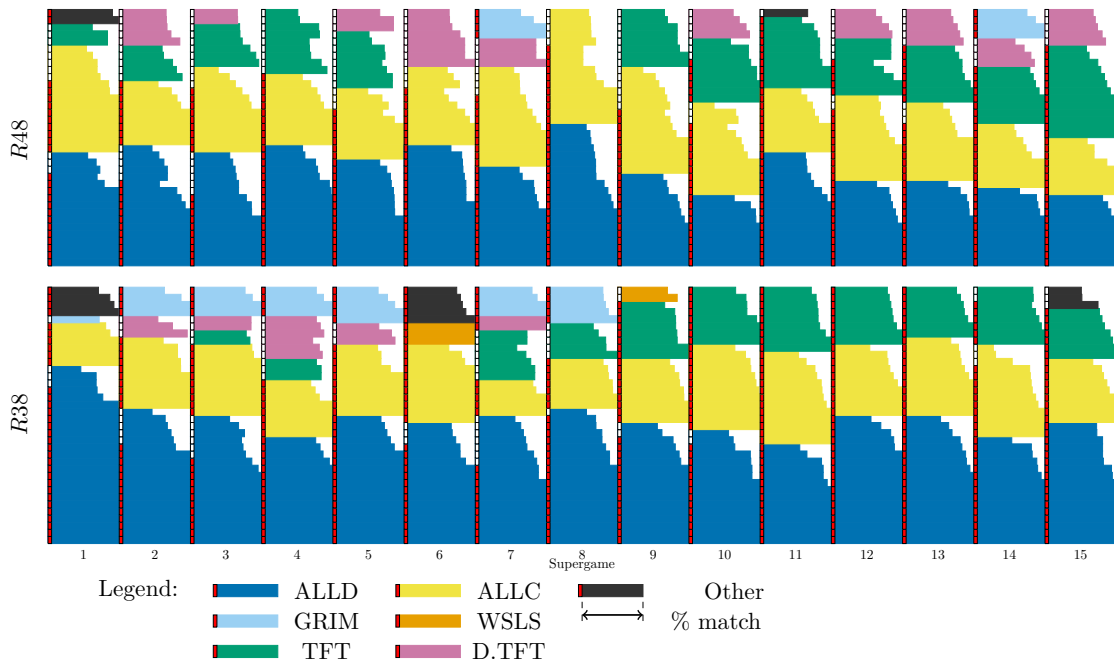


Figure 9: Clusters At The Beginning of Each Supergame. *Notes:* % match is relative to the strategy that is closest to the cluster exemplar. Red(white) line at the base of each bar denotes that the cluster exemplar matches(does not match) the closest strategy exactly.

We find that approximately 50% of participants' strategies exactly match one of the 20 strategies commonly used in the MLE estimation. The most common strategies are ALLD, ALLC, and TFT. Interestingly, while ALLD and ALLC are present from the beginning, TFT arises about halfway through the experiment. Among less commonly observed strategies are GRIM and WSLS, which appear earlier than TFT in both treatments.

The above analysis is somewhat restrictive. For example, an ALLC strategy with an additional *CsToD* rule would not be considered an exact match, even though it is different by as little as one period. In order to provide a broader classification of strategies, we run a cluster analysis on the vectors of actions obtained from the simulations. The cluster analysis seeks to classify items into groups with other similar items, based on some distance criterion. While a number of clustering methods are available, we opt for *affinity propagation* - a relatively recent clustering approach that has been shown to find clusters with much fewer errors and in much less time (Frey and Dueck, 2007). A useful feature of affinity propagation is that the optimal number of clusters is computed within the algorithm. Affinity propagation picks one of the participants' strategies from within a cluster to be an *exemplar* that is representative of that cluster. We classify each cluster based on which of the 20 MLE strategies is closest to the exemplar strategy. Figure 9 presents the results.

Result 6 *Behavior converges to three main clusters with exemplars at ALLD, ALLC, and TFT.*

There are several important results to point out in Figure 9. First, in the last five supergames,

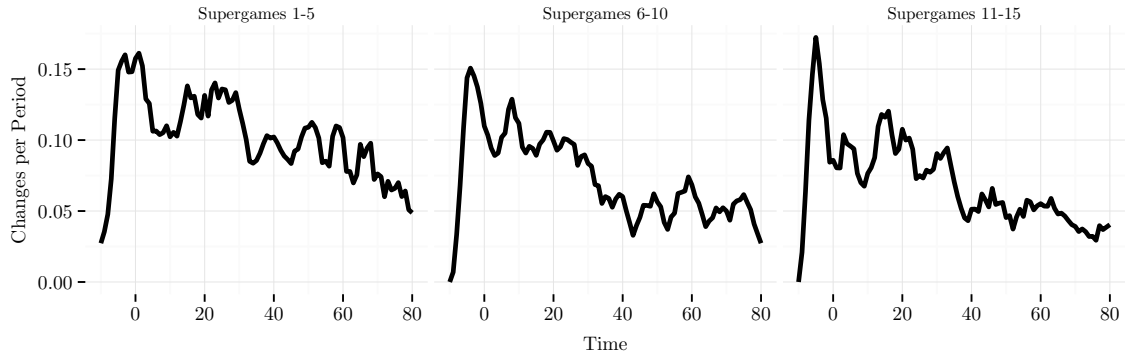


Figure 10: Average Number of Changes to Strategies Per Period. *Notes:* [-10,0] is the time prior to the start of a supergame when subjects can make changes to their rule sets. Each period lasts two seconds.

between 80% and 100% of subjects' strategies are in clusters with exemplars that match ALLD, ALLC, or TFT exactly. Second, clusters with exemplars exactly at ALLC and ALLD are present from the beginning, while the cluster with an exemplar exactly at TFT does not emerge until the second half of the experiment.¹⁰ Finally, TFT is most prominent in the *R48* treatment, while ALLD is most prominent in the *R38* treatment, which is consistent with the difference in levels of cooperation from **Result 1**.

3.4 Inferred Strategies

Although we give participants the ability to construct strategies, some may still choose to direct-respond within our interface. On the one hand, Figures 6(b) and 9 suggest that rule sets are relatively stable across the last five supergames. On the other hand, Figure 10 presents evidence that subjects are still making changes throughout the match. If these changes were planned in advance, but not implemented via rules, then analyzing behavior using the maximum likelihood estimation approach (Dal Bó and Fréchette, 2011) would be appropriate. We use this method to find strategies that are the most likely among the population, given the observed actions (see, Dal Bó and Fréchette (2011), Fudenberg, Rand, and Dreber (2012)).¹¹

The method works on the history of play as follows. First, fix the opponent's action sequence and compare the subject's actual play against that sequence to play generated by a given strategy, s^k , against that sequence. Then, strategy s^k correctly matches the subject's play in C periods and does not match the subject's play in E periods. Thus, a given strategy s^k has a certain number of correct plays C and errors E , and the probability that player i plays strategy k is

$$P_i(s^k) = \prod_{\text{Matches}} \prod_{\text{Periods}} \beta^C (1 - \beta)^E$$

¹⁰Note that there is a cluster with an exemplar that is closest to TFT in the beginning, but the exemplar doesn't match TFT exactly.

¹¹More papers that use this approach are cited in Dal Bó and Fréchette (2015): Vespa (2011), Camera, Casari, and Bigoni (2012), Fréchette and Yuksel (2013).

And the likelihood function is:

$$\mathcal{L}(\beta, \phi) = \sum_{i \in \text{Subjects}} \ln \left(\sum_{k \in \text{Strategies}} \phi^k P_i(s^k) \right)$$

Table 2 presents the estimation results for the full set of 20 strategies used in Fudenberg, Rand, and Dreber (2012). While there exist strategies that require an infinite rule set in our setting (one is described in Stahl (2011)), all 20 strategies in Fudenberg, Rand, and Dreber (2012) can be constructed with a finite number of rules in our setting. Furthermore, as is evident from the previous section, these strategies account for the majority of subjects' behavior. For each treatment, bootstrapped standard errors are calculated by drawing 100 random samples of the appropriate size (with replacement), estimating the MLE estimates of the strategy frequencies corresponding to each of these samples, and then calculating the standard deviation of the sampling distribution (Efron and Tibshirani, 1986).

Session	ALLD	Grim	WSLS	Tit For Tat	Lenient Grim 3	Exp. Tit For Tat	2 Tits for 2 Tats	False Cooperator	Tit For 3 Tats	Exp. Grim 2	Lenient Grim 2	Exp. Tit For 3 Tats	ALLC	β	Cooperation	# of Subjects
R48	0.28*** (0.07)	0.22*** (0.08)	0.09*** (0.04)	0.09*** (0.04)	0.06 (0.05)	0.08** (0.05)		0.04 (0.04)	0.09* (0.06)		0.02 (0.03)	0.03 (0.03)		0.87	0.45	36
R38	0.29*** (0.07)	0.14** (0.07)	0.16*** (0.04)	0.13*** (0.04)	0.06* (0.04)	0.03 (0.03)		0.07* (0.05)		0.03* (0.02)	0.04 (0.04)		0.06* (0.04)	0.90	0.37	36
D38	0.07*** (0.04)	0.07* (0.05)	0.14*** (0.05)	0.17*** (0.05)	0.12* (0.08)	0.06 (0.05)	0.15** (0.08)	0.03 (0.03)		0.06* (0.04)		0.03 (0.03)		0.91	0.60	34

Table 2: Maximum Likelihood Estimates *Notes:* Estimates use 20 initial periods of the last five supergames. Bootstrapped standard errors are shown in parentheses. Cooperation rates are reported for first 20 periods of interaction. Values of 0.00 are dropped for ease of reading.

Result 7 *Strategies obtained with the MLE procedure are ALLD, GRIM, TFT, and WSLS, with not much evidence of ALLC.*

We find strong evidence of ALLD, TFT (and exploitative TFT), which were also prominent in the cluster analysis in Section 3.3. Additionally, we find evidence of GRIM and WSLS, which were among the exactly constructed strategies that were identified in Section 3.3. The main difference between the results in Section 3.3 and the MLE is the prominence of the ALLC strategy. Specifically, while a significant fraction of participants specify the exact ALLC strategy at the beginning of each supergame, in the actual game-play, very few participants end up following this strategy. It appears that the subjects who were in the ALLC cluster may have been implementing GRIM or WSLS via direct response. Next, we investigate this further by considering which strategies are inferred from each of the three clusters.

Cluster	ALLD	Grim	Tit For Tat	WSLS	Lenient Grim 3	Tit For 3 Tats	Exp. Tit For Tat	Lenient Grim 2	ALLC	Tit For 2 Tats	False Cooperator	Exp. Tit For 3 Tats	β	Cooperation	# of Subjects
ALLC			0.08** (0.04)	0.06* (0.05)	0.22* (0.14)	0.20* (0.14)	0.06 (0.06)	0.14* (0.10)	0.12* (0.08)	0.07 (0.08)	0.06 (0.05)		0.87	0.65	17
ALLD	0.67*** (0.09)	0.17* (0.10)	0.06* (0.03)	0.06** (0.03)			0.04 (0.04)						0.89	0.20	27
TFT		0.31*** (0.11)	0.27*** (0.06)	0.27*** (0.05)			0.05 (0.04)					0.05 (0.05)	0.90	0.57	20

Table 3: Maximum Likelihood Estimates for strategy clusters. *Notes:* Not all subjects are included: 3 subjects were in a cluster for False Cooperator and 5 subjects were in a cluster for Exploitative Tit For Tat. Estimates use 20 initial periods of the last five supergames. Bootstrapped standard errors are shown in parentheses. Cooperation rates are reported for first 20 periods of interaction. Strategies that aren't above 0.05 for any cluster aren't included. Values of 0.00 are dropped for ease of reading.

We run an MLE on three separate subsets of the data, with each subset limited to one of the three main clusters obtained in Section 3.3. Table 3 presents the results. We find that the estimated strategies and realized cooperation rates within the three clusters are substantially different. Specifically, subjects who classified into the ALLC cluster cooperate the most, but rarely follow the exact ALLC strategy. Instead, their behavior is most often corresponds to a cooperative strategies like Lenient Grim 2, a Lenient Grim 3, or a Tit for 3 Tats. The delayed reaction to a defection and the fact that the constructed strategy prior to the supergame is ALLC suggests that subjects manually switch to D after observing that the other participant is taking advantage of their cooperative behavior.

The MLE estimates and cooperation rates from the ALLD cluster provide a sharp contrast to the estimates from the ALLC cluster. The differences in between the estimated strategies is both in terms of the types of strategies obtained and the extent to which those strategies match the exemplar strategy from the cluster. In particular, the vast majority of subjects that are classified into the ALLD cluster do follow the ALLD strategy.

The estimates from the TFT cluster indicate that subjects which are classified into the TFT cluster play either TFT, GRIM, or WSLS. There are two observations from Section 3.3 to support this. First, all three of these strategies were the exact strategies constructed in our experiment (see Figure 8). Second, the three strategies have common features and can be obtained from each other with minimal changes to the rule sets.¹² Additionally, the immediate reaction to defection, which is also common to these three strategies, means that subjects are likely to have constructed the strategies via rules instead of direct responding.

The fact that MLE estimates for the three clusters have very little overlap, provide further

¹²For example, going from TFT ($\rightarrow C, DD \rightarrow D, CD \rightarrow D$) to GRIM ($\rightarrow C, DD \rightarrow D, CD \rightarrow D, DC \rightarrow D$) requires an addition of one rule to the rule set.

evidence that participants’ behavior among the three clusters obtained in Section 3.3 is substantially different.

4 Simulation Analysis

In this section we conduct simulation exercises in order to i) better understand the implications of the *CsToD* rules on cooperation rates and MLE estimates; and ii) evaluate the performance of observed strategies in our experiment.

4.1 Implications of *CsToD* Rules

We simulate a population of 100 agents that play against each other for 20 periods. Each agent is assigned a rule set that specifies one of the five most popular strategies observed in our experiment: ALLC, ALLD, TFT, GRIM, and WSLS. We assume that all agents make mistakes 5% of the time. Specifically, 5% of the time the intended action is switched to the opposite action. For simplicity, we consider an assignment such that there are an equal number of subjects following each strategy. In what follows, we investigate what happens to the MLE estimates and cooperation rates when we add a memory-3 *CsToD* rule to a fraction, α , of each agent type. Figure 4 presents the results for $\alpha \in \{0, 0.25, .5, 0.75, 1\}$.

α	ALLC	ALLD	Grim	WSLS	Tit For Tat	False Cooperator	2 Tits for 2 Tats	T2	2 Tits for 1 Tat	β	Cooperation	# of Subjects
0.00	0.26 (0.05)	0.19 (0.04)	0.14 (0.05)	0.12 (0.03)	0.12 (0.03)	0.05 (0.03)	0.03 (0.02)	0.01 (0.02)	0.04 (0.02)	0.89	0.54	100
0.25	0.24 (0.05)	0.20 (0.04)	0.16 (0.04)	0.12 (0.03)	0.10 (0.03)	0.09 (0.02)	0.06 (0.03)		0.01 (0.01)	0.87	0.47	100
0.50	0.20 (0.04)	0.20 (0.04)	0.16 (0.04)	0.10 (0.03)	0.10 (0.03)	0.11 (0.03)	0.07 (0.03)	0.04 (0.02)		0.87	0.43	100
0.75	0.24 (0.04)	0.20 (0.04)	0.22 (0.04)	0.10 (0.03)	0.10 (0.03)	0.05 (0.02)	0.04 (0.02)	0.01 (0.01)		0.86	0.42	100
1.00	0.22 (0.04)	0.20 (0.04)	0.21 (0.04)	0.10 (0.03)	0.10 (0.03)	0.04 (0.02)	0.12 (0.04)			0.84	0.39	100

Table 4: Maximum Likelihood Estimates for simulated data. *Notes:* Estimates based on 20 periods of simulations. Bootstrapped standard errors are shown in parentheses. Cooperation rates are reported for first 20 periods of interaction. Strategies that aren’t above 0.05 for any cluster aren’t included. Values of 0.00 are dropped for ease of reading.

There are several points worth noting. First, as α increases, the overall cooperation decreases from 0.54 to 0.39. This result corroborates our discussion in Section 3.2 and **Results 1-3**. Second,

MLE estimates are largely unaffected by the increase in α . Third, the β parameter of the MLE, which captures the amount of noise, decreases from 0.89 to 0.84. Thus, we find that regularly observed aspects of strategies, specifically $-CsToD$ rules, are likely to be perceived as noise in the estimation procedure, but nevertheless are vital to the cooperation dynamics within supergames.

4.2 Strategy Performance

Since we have the full specification of the subjects' strategies, we can calculate the expected performance of a strategy against the population of subjects in the experiment. We simulate five of the commonly observed strategies against all of the participants' strategies taken at the beginning of each supergame. We then calculate the average earnings per period. Figure 11 presents the results.

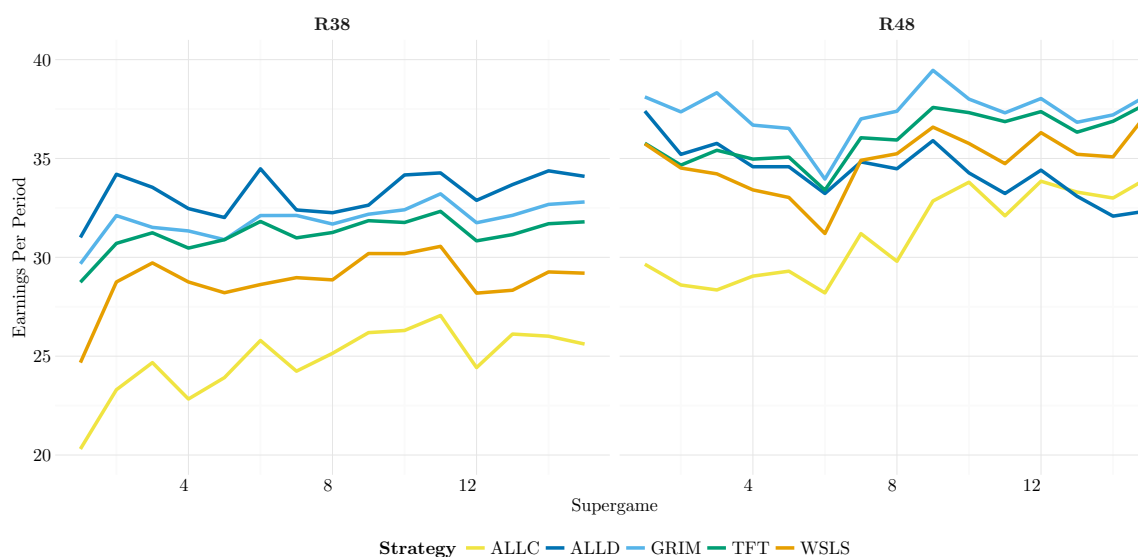


Figure 11: Strategy Performance Throughout The Experiment. *Notes:* Each of the five strategies was matched with all 36 participants' strategies and simulated for 20 periods.

Figure 11 shows that the performance of the ALLC strategy is increasing in the first part of the experiment, but is relatively stable in the second part of the experiment. The expected performance of the ALLC strategy is better than 25 (the D-D payoff), even in the *R38* treatment. This means that playing the ALLC strategy is not necessarily irrational across supergames. There is a stark difference between the two treatments in terms of the evolution of the strategy performance: In the *R38* treatment, for the duration of the experiment, the best-performing strategy is ALLD and the worst-performing strategy is ALLC. In the *R48* treatment, ALLD is one of the best-performing strategies in the first supergame, but is one of the worst-performing strategies in the last supergame. These findings are not surprising given **Result 1**.

5 Conclusions

The contribution of this paper is twofold. First, we develop the interface that allows us to run experiments on long repeated games. The interface combines features of the continuous-time experiments with the strategy method, allowing us to gain a unique insight into the strategies that participants develop. Second, we conduct experiments with our interface. In particular, we study the repeated prisoner’s dilemma with a continuation probability of $\delta = .99$.

We find that cooperation increases with the payoff to mutual cooperation, decreases as the supergame progresses, and increases when subjects do not use the strategy method. When analyzing the rules, we find that subjects consistently construct and use rules of length longer than memory-1. In particular, they regularly use a specific class of rules, *CsToD* and *DsToC*, which play an important role in determining the levels of cooperation in the different treatments. We then analyze the fully specified strategies and find that, while roughly 50% of the strategies are exact matches of those commonly used in the literature, about 80%-100% of the strategies are close to ALLD, ALLC, and TFT. Finally, we perform the standard maximum likelihood estimation procedures based on only the observed actions, and we find that results for long repeated games are consistent with those of previous studies on shorter repeated games.

Combined with the main results, the simulations presented in Section 4 suggest that the standard maximum likelihood procedure does an excellent job of uncovering strategies in fairly complex environments. However, they also suggest that the strategy estimates may miss some important aspects of play, such as the *CsToD* rules that subjects use. Though the addition of these rules has little impact on the strategy frequencies in the maximum likelihood estimates, they can have large impact on the level of cooperation.

There are many interesting avenues of future research. First, it would be interesting to gain a better understanding of what causes these long rules to be played. For example, are they only used in very long repeated games, like the one studied here, or would they also be present in shorter games (such as $\delta = 0.9$). In addition, it would be useful to try to better understand how subject learn. From the results, it is not clear whether subjects’ strategies are changing over the course of a supergame, or whether they have a fixed strategy in mind at the beginning of the supergame, and implement that strategy via direct response. Finally, it may be interesting to better understand the difference between short and long repeated games. For example, is there some level of continuation probabilities δ after which behavior doesn’t change? The proposed experimental interface will be a useful tool in these further investigations.

References

- BIGONI, M., M. CASARI, A. SKRZYPACZ, AND G. SPAGNOLO (2015): “Time horizon and cooperation in continuous time,” *Econometrica*, 83(2), 587–616.
- BRANDTS, J., AND G. CHARNES (2000): “Hot vs. cold: Sequential responses and preference stability in experimental games,” *Experimental Economics*, 2(3), 227–238.
- (2011): “The strategy versus the direct-response method: a first survey of experimental comparisons,” *Experimental Economics*, 14(3), 375–398.
- CAMERA, G., M. CASARI, AND M. BIGONI (2012): “Cooperative strategies in anonymous economies: An experiment,” *Games and Economic Behavior*, 75(2), 570–586.
- CASARI, M., AND T. N. CASON (2009): “The strategy method lowers measured trustworthy behavior,” *Economics Letters*, 103(3), 157–159.
- DAL BÓ, P., AND G. R. FRÉCHETTE (2011): “The evolution of cooperation in infinitely repeated games: Experimental evidence,” *The American Economic Review*, 101(1), 411–429.
- (2015): “Strategy choice in the infinitely repeated prisoners dilemma,” *Available at SSRN 2292390*.
- DUFFY, J. (2008): “Macroeconomics: a survey of laboratory research,” *Handbook of experimental economics*, 2.
- EFRON, B., AND R. TIBSHIRANI (1986): “Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy,” *Statistical science*, pp. 54–75.
- FRÉCHETTE, G. R., AND S. YUKSEL (2013): “Infinitely repeated games in the laboratory: four perspectives on discounting and random termination,” *Available at SSRN 2225331*.
- FREY, B. J., AND D. DUECK (2007): “Clustering by Passing Messages Between Data Points,” *Science*, 315(5814), 972–976.
- FRIEDMAN, D., AND R. OPREA (2012): “A continuous dilemma,” *The American Economic Review*, 102(1), 337–363.
- FUDENBERG, D., D. G. RAND, AND A. DREBER (2012): “Slow to Anger and Fast to Forgive: Cooperation in an Uncertain World,” *American Economic Review*, 102(2), 720–49.
- GOOD, P. (2013): *Permutation tests: a practical guide to resampling methods for testing hypotheses*. Springer Science & Business Media.
- GOOD, P. I. (2006): *Permutation, parametric, and bootstrap tests of hypotheses*. Springer Science & Business Media.
- GREINER, B. (2004): “An online recruitment system for economic experiments,” .
- RAND, D. G., J. D. GREENE, AND M. A. NOWAK (2012): “Spontaneous giving and calculated greed,” *Nature*, 489(7416), 427–430.

- RAND, D. G., A. PEYSAKHOVICH, G. T. KRAFT-TODD, G. E. NEWMAN, O. WURZBACHER, M. A. NOWAK, AND J. D. GREENE (2014): “Social heuristics shape intuitive cooperation,” *Nature communications*, 5.
- REUBEN, E., AND S. SUETENS (2012): “Revisiting strategic versus non-strategic cooperation,” *Experimental Economics*, 15(1), 24–43.
- SELTEN, R. (1967): “Die Strategiemethode zur Erforschung des eingeschränkt rationalen Verhaltensim Rahmen eines Oligopolexperiments,” *Beitrge zur Experimentellen Wirtschaftsforschung*, pp. 136–168, ed.by H. Sauermann. Tiubingen: J. C. B. Mohr.
- STAHL, D. O. (2011): “Cooperation in the sporadically repeated prisoners dilemma via reputation mechanisms,” *Journal of Evolutionary Economics*, 21(4), 687–702.
- (2013): “An experimental test of the efficacy of a simple reputation mechanism to solve social dilemmas,” *Journal of Economic Behavior & Organization*, 94, 116–124.
- VESPA, E. (2011): “Cooperation in dynamic games: an experimental investigation,” *Available at SSRN 1961450*.

Appendix A: Experimental Instructions

You are about to participate in an experiment in the economics of decision-making. If you listen carefully, you could earn a large amount of money that will be paid to you in cash in private at the end of the experiment.

It is important that you remain silent and do not look at other people's work. If you have any questions, or need any assistance of any kind, please raise your hand and an experimenter will come to you. During the experiment, **do not talk, laugh or exclaim out loud** and be sure to keep your **eyes on your screen only**. In addition please **turn off your cell phones, etc.** and put them away during the experiment. Anybody that violates these rules will be asked to leave and will **not** be paid. We expect and appreciate your cooperation.

Agenda

1. We will first go over the instructions.
2. Then we will have a practice match to learn the interface.
3. Next, there will be a quiz with 10 questions to make sure everyone understands the instructions. You will be able to earn up to \$0.50 for each of 10 quiz questions. If you answer the question correctly on your first attempt, you earn \$0.50. If you answer the question incorrectly on your first attempt and correctly on your second attempt, then you earn \$0.25. If you don't answer the question correctly on your first two attempts, you will earn \$0 and the answer will be explained to you. You **must** still answer question correctly to move on to the next question.
4. After the quiz, the experiment will begin. In the experiment you will be working with a fictitious currency called Francs. You will be paid in US Dollars at the end of the experiment. **The exchange rate today is: 5,000.00 Francs = 1.00 USD.**

Experiment Details

- This experiment consists of **fifteen matches**.
- Each match consists of **two stages**, called the **Free Stage** and the **Lock Stage**.
- Each stage has a **different number of periods**.
- At the beginning of each match you will be paired randomly with one other participant. You will remain matched with this same participant until the **end of the match**, but then will be paired with another randomly selected participant in the following match.
- Each match will have the same structure, but may contain different numbers of periods.
- You will remain anonymous throughout the experiment. You will not know the identity of the participant that you are paired with, and they will not know your identity. The choices made by you and the participant your are paired with have no effect on the payoffs of participants in other pairs and vice versa. Therefore, your payoff in a given match is based solely on the choices made by you and the participant that you are paired with.

Specific Instructions for Each Period

- Your payoff in each period will depend on your choice and the choice of the participant that you are paired with.
- You will choose one of two options, either W or Y .
- You will be able to see the payoffs for each combination of choices for you and the participant that you are paired with.
- These payoffs will remain the same throughout the entire experiment (all matches).
- The payoffs will be displayed in a table like this:

My Choice	W	W	Y	Y
Other's Choice	W	Y	W	Y
My Payoff	1	2	3	4
Other's Payoff	5	6	7	8
Times Occurred	10	12	13	9

In this example above, the rows are the following:

- Row #1 - Your choice (either W or Y in this example)
 - Row #2 - Other's choice (either W or Y in this example)
 - Row #3 - Your payoff
 - Row #4 - Other's payoff
 - Row #5 - Total number of times that that combination has been played **this match**.
- For example, in the table above, if you choose W and the participant you are paired with chooses Y , then you receive a payoff of 2 and the other participant receives a payoff of 6. This combination has occurred 12 times so far this match.

History

- As the match progresses, you will see the **history** of play across the top of the screen, displayed like this:

23	Period	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45
My Choice:	W	Y	Y	Y	W	W	Y	Y	Y	W	Y	W	Y	W	Y	W	Y	Y	Y	W	Y	?
Other's Choice	W	Y	W	W	W	W	Y	W	Y	Y	Y	Y	Y	W	Y	Y	Y	Y	Y	W	W	?

- This history tells your choice (labeled My Choice), the choice of the participant you are paired with (labeled Other's Choice), and the current period.
- For example, in the above example, you played *W* and the participant you are paired with played *Y* in period 39.

Rules

- Rather than directly making choices of *W* or *Y*, you will develop a set of rules which will automatically make choices for you.
- The set of rules will appear in the middle of the screen.
- You will be able to construct rules using the rule constructor at the bottom of the screen.
- A rule consists of two parts:
 - Input Sequence** - A sequence of choices by you and the participant you are paired with.
 - Output** - A choice to be made by you after the input sequence occurs.
- Some example rules are displayed below:

Y	W	Y	W
Y	W	W	

Rule #2

Y	W	W	W
W	Y	Y	

Rule #3

W	Y	Y
W	W	

Rule #4

- For example, Rule #2 has an input sequence of $(Y,Y),(W,W),(Y,W)$ and an output of *W*. This means that if you have played *Y*, *W*, and then *Y* in the last three periods and the participant that you are paired with has played *Y*, *W*, and then *W*, this rule will lead you to play *W* in the next period.
- The length of the rule is measured by the length of the input sequence.
- So rule #2 and rule #3 have a length of 3, and rule #4 has a length of 2.
- There are several ways that you can modify your rules during the experiment.
 - First, you can use the rule constructor in the bottom center of the screen. Press the plus button to add more columns or one of the minus buttons to subtract columns. Click on the question marks to fill in the boxes. When you click on a question mark, either a *W* or a *Y* will appear. If you click on a *W* it will switch to a *Y* and if you click on a *Y* it will switch to a *W*. Once you have completely filled in the rule (leaving no question marks), the add rule button will appear, and the rule can be added to your set.

2. Second, if you look at a rule in the set, you will notice a green copy button and a red delete button. If you press the **copy button**, it will copy the rule down to the constructor and you will be able to create a similar rule. If you press the red **delete button** it will delete the rule from the set.
 3. Since your rule set needs to make a single choice each period, it is not possible to have two rules with the same input sequence, but with different outputs. If you create a rule that has the same input sequence but a different output, you will get an error that says "**Conflicting rule in set**" and a button that says "**Switch Rule**" will appear. If you press this button, it will delete the conflicting rule from the rule set, and add the rule from the constructor.
- A rule of length n is said to **fit the history** if the input sequence matches the last n periods of the history.
 - For example, since the last three periods of play in the above history (periods 42-44) have been (Y, Y) , (W, W) , and (Y, W) and that sequence is also the input for rule #2, then rule #2 is said to fit the history. Similarly, given the above history, we can see that rule #4 fits the history but rule #3 does not fit the history.
 - As play progresses you will develop a set of rules that will be used to make your choices.
 - If more than one rule fits the history, then the **rule with the longer length will determine the choice**.
 - For example, given the above history since both rule #2 and rule #4 fit the history, rule #2 will be used to make your choice since it is longer. Therefore, given the history, and the three rules, your choice next period will be W , as prescribed by rule #2.
 - If no rules fit the history, then your **Default Rule** rule will be selected. The default rule will only be used when no rules fit the history. To select your default rule select either W or Y in the bottom left of your screen
 - Play will proceed automatically. Each period when the choice is made, the rule that was used to make the choice will be highlighted, the corresponding sequence in the history will be highlighted, and the corresponding choice on the payoff table will also be highlighted.

Number of Periods Per Match

- Each match will consist of two stages:
 - **Free Stage** - you are free to edit your set of rules.
 - **Lock Stage** - you are not able to edit your set of rules.
- The number of periods in each stage will be determined randomly using the following procedure.
 - At the end of each period, a number will be chosen randomly from the set of numbers $\{1, 2, 3, \dots, 98, 99, 100\}$, where **each number is equally likely**.
 - If the number is 1, then the match will end.
 - If the number is 2, then the Free Stage will end, and you will transition into the Lock Stage for the rest of the match.
 - If the number is not 1, then there will be an additional period.

- The number will always be placed back into the set after it is drawn.
 - Thus, in any period there is a 1% CHANCE that the match will end and a 99% CHANCE that the match will have another period.
 - Therefore, the expected number of periods in each match will be 100.
 - You will not see the number selected from $\{1, 2, 3, \dots, 98, 99, 100\}$.
 - To ensure that the length of the match is not dependent on your play, the number of periods for each match has been written on the board before the experiment, and will be uncovered at the end of the experiment.
- Since your choices will be determined automatically from your rule set, each period will last **2 seconds** during the Free Stage and will last less than 2 seconds during the Lock Stage. Therefore, the game will progress automatically even if you don't do anything.
 - It is possible that the lock stage will have zero periods.

Additional Information about Matches

- All of the rules in your set at the end of one match will remain in your set at the start of the next match.
- Before the first match you will have 1 minute to look over the payoffs followed by an additional minute to edit your set of rules. Before you can make any changes to your rule set however, you must set your default rule in the bottom left.
- Before every other match, you will have 10 seconds before play starts to edit your set of rules.

Payoffs

- At the end of the experiment, you will be paid in **cash**.
- Your payoff at the end of the experiment will be the sum of the payoffs for each period.

Practice Match

- Next you will be given the chance to have a practice match to make you comfortable with the interface.
- It is important to note that in this practice match,
 - you will NOT be paid for the choices made in the practice round.
 - the payoffs are BLURRED, so that you can focus on getting comfortable with the interface rather than focusing on the payoffs.
 - You are NOT matched with another participant. Your opponent for the practice round is a computer that is playing RANDOMLY.

Interface Screen-shot

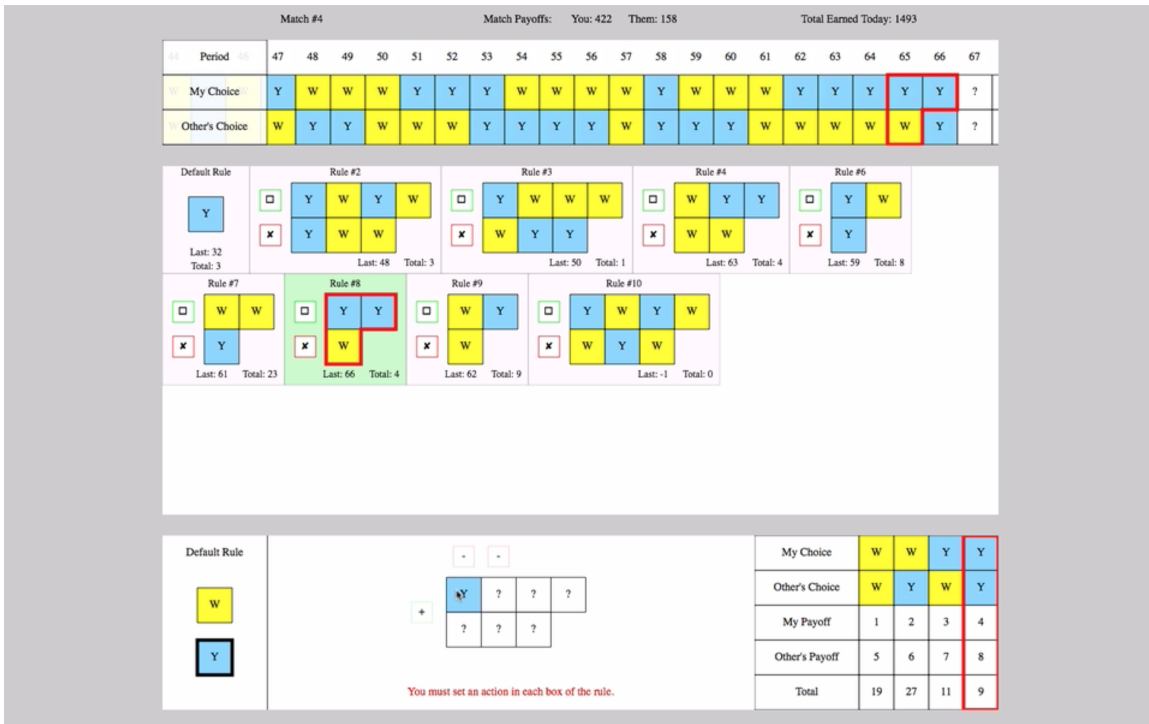


Figure 12: Interface Screenshot.

Appendix B: Quiz

History and rules for questions 1-3:



- New Screen -

2. (Try #1 - \$0.50) Last problem you learned that you will play W next period. If the other's choice is W, what will YOUR payoff be?

1 2 3 4 5 6 7 8

Quiz Earnings: \$0.50

My Choice	W	W	Y	Y
Other's Choice	W	Y	W	Y
My Payoff	8	6	4	2
Other's Payoff	7	5	3	1
Total	0	0	0	0

– New Screen –

3. (Try #1 - \$0.50)
ADD a rule to the set to ensure that Y will be played next period.

Quiz Earnings: \$1.00



You must set an action in each box of the rule.

My Choice	W	W	Y	Y
Other's Choice	W	Y	W	Y
My Payoff	8	6	4	2
Other's Payoff	7	5	3	1
Total	0	0	0	0

History and rules for questions 4-6:

Period	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
My Choice	W	W	W	Y	W	Y	W	Y	Y	W	Y	W	W	W	Y	W	W	W	W	?	?	?	?
Other's Choice	W	W	W	Y	W	Y	W	Y	W	Y	W	Y	W	W	Y	W	W	Y	W	?	?	?	?

Default Rule	Rule #4	Rule #7	Rule #9
<input type="checkbox"/> Y	<input type="checkbox"/> W <input type="checkbox"/> Y	<input type="checkbox"/> Y <input type="checkbox"/> Y	<input type="checkbox"/> W <input type="checkbox"/> W <input type="checkbox"/> W
Last: 5 Total: 9	<input checked="" type="checkbox"/> W <input checked="" type="checkbox"/> Y	<input checked="" type="checkbox"/> Y <input checked="" type="checkbox"/> Y	<input checked="" type="checkbox"/> Y <input checked="" type="checkbox"/> W
	Last: 3 Total: 3	Last: 7 Total: 5	Last: 2 Total: 4

4. (Try #1 - \$0.50) Given the current history, what action will be played in the next period?

W Y

Quiz Earnings: \$1.50

My Choice	W	W	Y	Y
Other's Choice	W	Y	W	Y
My Payoff	8	6	4	2
Other's Payoff	7	5	3	1
Total	0	0	0	0

– New Screen –

5. (Try #1 - \$0.50) Last problem you learned that you will play Y next period. If the other's choice is Y, what will YOUR payoff be?

1 2 3 4 5 6 7 8

Quiz Earnings: \$2.00

My Choice	W	W	Y	Y
Other's Choice	W	Y	W	Y
My Payoff	8	6	4	2
Other's Payoff	7	5	3	1
Total	0	0	0	0

– New Screen –

6. (Try #1 - \$0.50) DELETE a rule from the set to ensure that W will be played next period.

Quiz Earnings: \$2.50

My Choice	W	W	Y	Y
Other's Choice	W	Y	W	Y
My Payoff	8	6	4	2
Other's Payoff	7	5	3	1
Total	0	0	0	0

History and rules for questions 7-9:

Period	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
My Choice	Y	W	W	W	W	W	W	W	Y	W	W	W	Y	W	W	Y	Y	W	W	?	?	?	?
Other's Choice	Y	Y	Y	Y	Y	W	W	W	Y	W	Y	Y	Y	W	Y	Y	W	Y	Y	?	?	?	?

Rule	Grid	Last	Total
Default Rule	<input type="checkbox"/> Y	5	9
Rule #4	<input type="checkbox"/> W Y <input checked="" type="checkbox"/> W	3	3
Rule #7	<input type="checkbox"/> Y Y <input checked="" type="checkbox"/> Y	7	5
Rule #9	<input type="checkbox"/> W W W <input checked="" type="checkbox"/> Y Y	2	4
Rule #13	<input type="checkbox"/> W Y Y <input checked="" type="checkbox"/> Y Y	4	7

7. (Try #1 - \$0.50) Given the current history, what action will be played in the next period?

W

Y

Quiz Earnings: \$3.00

My Choice	W	W	Y	Y
Other's Choice	W	Y	W	Y
My Payoff	8	6	4	2
Other's Payoff	7	5	3	1
Total	0	0	0	0

- New Screen -

8. (Try #1 - \$0.50) Last problem you learned that you will play W next period. If the other's choice is W, what will the OTHER'S payoff be?

1

2

3

4

5

6

7

8

Quiz Earnings: \$3.50

My Choice	W	W	Y	Y
Other's Choice	W	Y	W	Y
My Payoff	8	6	4	2
Other's Payoff	7	5	3	1
Total	0	0	0	0

- New Screen -

9. (Try #1 - \$0.50) COPY and SWITCH a rule from the set to ensure that Y will be played next period.

+

?	?
?	

You must set an action in each box of the rule.

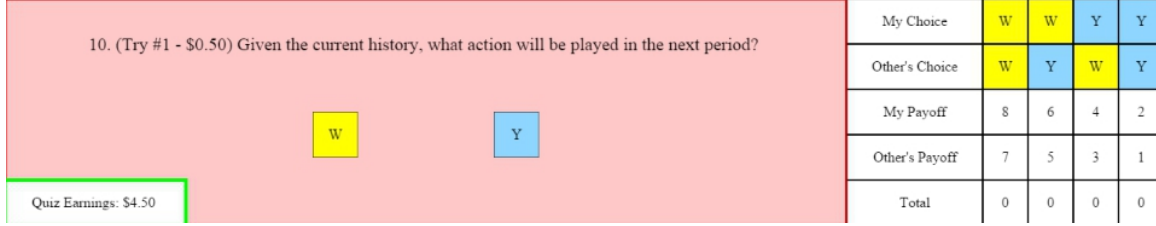
Quiz Earnings: \$4.00

My Choice	W	W	Y	Y
Other's Choice	W	Y	W	Y
My Payoff	8	6	4	2
Other's Payoff	7	5	3	1
Total	0	0	0	0

History and rules for question 10:

Period	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
My Choice	Y	W	W	Y	W	W	Y	W	W	W	Y	W	W	W	W	Y	Y	Y	W	?	?	?	?
Other's Choice	Y	W	W	W	Y	W	Y	Y	Y	W	W	Y	W	W	Y	Y	W	Y	Y	?	?	?	?

Rule	Grid	Last	Total
Default Rule	<input type="checkbox"/> W	5	9
Rule #4	<input type="checkbox"/> W W <input checked="" type="checkbox"/> W	3	3
Rule #7	<input type="checkbox"/> Y Y <input checked="" type="checkbox"/> Y	7	5
Rule #9	<input type="checkbox"/> W W Y <input checked="" type="checkbox"/> Y Y	2	4
Rule #13	<input type="checkbox"/> W W Y W <input checked="" type="checkbox"/> Y Y W	4	7



Appendix C: Robustness Check

Given the relative complexity of the environment, the level of cooperation around 50% could be a result of noise that stems from confusion about the rule construction process. A natural question, therefore, is: can a higher level of cooperation be achieved if we separate the subjects who best understand the interface from the other subjects? To address this question, we recruited 32 additional participants (two sessions of 16) to participate in an additional (robustness) treatment of our experiment. The only modification that was made was that each session was split in two subgroups (eight subjects each) based on their quiz score: High-Quiz and Low-Quiz. It turned out that the median quiz earnings in both sessions was \$4.50, which means that the participants in the High-Quiz group answered at least nine out of ten questions without an explanation, which indicates an excellent level of understanding of the interface. To keep the information about the treatments exactly the same compared to the original *R38* treatment, we did not tell the participants about the split.

Treatm.	Periods				
	First	First.4		Last.4	All
<i>R38</i>	0.594 (0.033)	0.461 (0.036)	***	0.317 (0.04)	0.352 (0.037)
<i>R38.H</i>	0.512 (0.06) **	0.441 (0.057) **	***	0.297 (0.057)	0.356 (0.049)
<i>R38.L</i>	0.338 (0.052)	0.294 (0.045)		0.272 (0.054)	0.274 (0.041)

Figure 13: Cooperation and Quiz Performance. *Notes:* *R38* - original treatment reported in Figure 5. *R38.H* - new treatment with only High-Quiz participants. *R38.L* - new treatment with only Low-Quiz participants. The unit of observation is a matched pair of subjects per supergame. Bootstrapped standard errors are in parentheses. Black stars denote significance using permutation test. Red stars denote significance using a matched pairs randomization test.

We find that cooperation rates in the High-Quiz group follow the pattern observed in the original *R38* treatment: higher cooperation early on and lower cooperation later on. Furthermore, there are no significant differences between the original *R38* treatment and the new *R38.H* treatment whether we look at the first period, the first four periods, the last four periods, or all periods. Therefore, given the new treatment, we find no evidence that separating potentially “noisy” Low-Quiz subjects would improve cooperation in our experiment.