European Economic Review 104 (2018) 185–219

Contents lists available at ScienceDirect

European Economic Review

journal homepage: www.elsevier.com/locate/euroecorev

Constructing strategies in the indefinitely repeated prisoner's dilemma game[☆]

Julian Romero[a,*], Yaroslav Rosokha[b]

[a] Eller College of Managment, Univeristy of Arizona, USA
[b] Krannert School of Management, Purdue University, USA

ARTICLE INFO

Article history:
Received 6 July 2016
Accepted 7 February 2018
Available online 17 March 2018

Keywords:
Indefinitely repeated games
Prisoner's dilemma
Experiments
Cooperation
Experimental design
Strategies

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 pure strategies that can be implemented. This gives us insight into what happens in long repeated games and into the types of strategies that subjects construct. We report results obtained from the indefinitely repeated prisoner's dilemma with a continuation probability of $\delta = .95$. We find that during such long repeated prisoner's dilemma games, cooperation drops from the first period of a supergame to the last period of a supergame. When analyzing strategies, we find that subjects rely on strategies similar to those found in the literature on shorter repeated games—specifically Tit-For-Tat, Grim Trigger, and Always Defect. However, we also identify features of strategies that depend on more than just the previous period that are responsible for the drop in cooperation within supergames, but that may be overlooked when using the common strategy frequency estimation approach.

1. Introduction

The repeated prisoner's dilemma has been used as a stylized setting to model a wide variety of situations across many disciplines (e.g., 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 "long" indefinitely repeated prisoner's dilemmas (continuation probability $\delta = 0.95$). We use this environment to gain a unique perspective on the strategies that subjects construct in the indefinitely repeated prisoner's dilemma and on the factors that make subjects cooperate.

☆ This paper benefited greatly from discussions with and comments from Tim Cason, Pedro Dal Bo, John Duffy, Guillaume Frechette, Ernan Haruvy, Christos Ioannou, Dale Stahl, Emanuel Vespa, Nathaniel Wilcox, Alistair Wilson, and Huanren Zhang, as well workshop participants at the 2014 North American ESA meetings in Ft. Lauderdale, the 2015 World ESA meetings in Sydney, ESI Theory/Experiments Workshop at Chapman University, the 2017 SAET meetings in Faro and seminar participants at EPFL, Florida State University, New York University, Penn State University, Purdue University, Texas A&M University, the University of Pittsburgh, the University of Southern California, the University of Texas at Dallas.

* Corresponding author.
E-mail addresses: jnromero@email.arizona.edu (J. Romero), yrosokha@purdue.edu (Y. Rosokha).

https://doi.org/10.1016/j.euroecorev.2018.02.008
0014-2921/

Our experimental interface implements 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 approach allows players' actions to be chosen automatically using the rules in the rule set.

Following Dal Bó and Fréchette (2017), the experiment is divided into three stages: the direct-response stage, the non-binding stage, and the locked-response stage. In the direct-response stage, players play the repeated game directly. In this way, players are able to learn the nature of the game and the trade-offs involved in such indefinitely repeated interactions. In the non-binding stage, players create strategies which select actions for them, though players are not required to play the action prescribed by the strategy. This stage allows players to gain experience with how constructed strategies make choices and allows them to construct a strategy that matches their desired behavior. In the locked-response stage, players cannot make any changes to their strategies, and their strategies play for them automatically. This stage provides incentives for strategy construction.

This experimental design offers several benefits over standard indefinitely repeated games experiments. First, we can directly view players' strategies. A growing body of literature aims to better understand the strategies played in repeated prisoner's dilemma games. The literature takes three approaches to identifying the strategies that subjects play. In the first approach, actions directly chosen by players are then used to make inferences about these players' actual strategies (Bigoni et al., 2015; Camera et al., 2012; Dal Bó and Fréchette, 2011; Fudenberg et al., 2012; Stahl, 2013). This inference requires the researcher to specify a predefined set of strategies to be used in the estimation. While commonly studied strategies work well in shorter repeated games, it is not clear if this same set of strategies is appropriate for longer repeated games. In the second approach, players select from a set of strategies (Cason and Mui, 2017; Dal Bó and Fréchette, 2017). This approach also requires researchers to specify a set of strategies to begin with. Although the strategies are now directly observable, subjects' behavior may be influenced by the strategies presented in the set. In the third approach, which we take, players construct strategies from scratch (Bruttel and Kamecke, 2012; Dal Bó and Fréchette, 2017; Embrey et al., 2016; Romero and Rosokha, 2016). An advantage of our interface is that there are minimal restrictions on the types and lengths of pure strategies.[1] We then can determine the extent to which the typically assumed sets of strategies are appropriate for long repeated prisoner's dilemma games and provide a foundation for using a particular set of strategies.

In addition to being able to observe subjects' strategies, our experimental interface 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 (Roth and Murnighan, 1978). One difficulty with this standard approach is that a single repeated game can last a very long time. Therefore, indefinitely repeated games in the laboratory have typically focused on situations with relatively low continuation probabilities. Since subjects are constructing complete strategies with our interface, choices can be semi-automated (i.e., actions played by the strategy are confirmed by the subject) or fully automated (i.e., actions are played by the strategy automatically) for a large part of the experiment. This feature is useful for running long repeated games experiments in the lab. Long repeated games may reveal important aspects of behavior that are not evident in shorter repeated games. Furthermore, 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).

Our design allows us to test whether our interface impacts subjects' behavior. In our experiment, subjects' ability to construct strategies did not significantly impact levels of cooperation, and these levels were similar to those found in previous studies that used similar experimental parameters. These findings suggest that our experimental interface does not affect subjects' behavior. Using this interface, we ran indefinitely repeated prisoner's dilemma experiments with continuation probability $\delta = 0.95$ and find three main results. First, in long repeated games, levels of cooperation decrease from the beginning of the supergame to the end of the supergame (*Result 1*). Second, by directly viewing strategies that subjects created using our interface, we find strong evidence that subjects construct strategies longer than memory-1 (*Result 2*). Finally, using a clustering algorithm that allows endogenously determining groups of similar strategies, we find that subjects use strategies that behave similarly to memory-1 strategies such as Tit-for-Tat, Grim Trigger, and Always Defect (*Result 3*).

Though many of the subjects played strategies that behave similarly to memory-1 strategies, a significant proportion of the strategies had a common feature that differentiated them from memory-1 strategies. This common feature, which we refer to as a *CsToD* rule, causes a strategy to defect after a sequence of multiple periods of mutual cooperation. These *CsToD* rules provide an explanation for *Result 1* and the seemingly contradictory *Result 2* and *Result 3*. Specifically, even though subjects may be playing according to a memory-1 strategy after most histories, they may have more complex components to their strategies, such as the *CsToD* rules. These complex components are played rarely, but could cause cooperation to break down from the beginning to the end of the supergame. Using simulations, we show that the *CsToD* rules impact cooperation rates but may not affect strategy estimates when standard maximum likelihood procedures are used.

[1] A shortcoming of this interface is that it doesn't allow subjects to play mixed strategies, which may play a role in the indefinitely repeated prisoner's dilemma (e.g., Breitmoser, 2015).

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.

We contribute to the recent experimental literature that investigates strategy elicitation in the indefinitely repeated prisoner's dilemma.[2] Bruttel and Kamecke (2012) run finitely and indefinitely repeated prisoner's dilemma games with a continuation probability of $\delta = 0.8$. They elicit strategies by asking subjects to construct memory-1 strategies.[3] Though they elicit strategies, they don't examine the elicited strategies, as the main focus is on eliminating endgame effects. Dal Bó and Fréchette (2017) ask subjects to design memory-1 strategies that will play in their place and play in games with probability of continuation of $\delta \in \{0.50, 0.75, 0.90, 0.95\}$. Additionally, they implement a treatment whereby subjects choose from a list of more complex predefined strategies. They find no effect of the strategy method on levels of cooperation, and find evidence of Tit-For-Tat, Grim Trigger and Always Defect. Cason and Mui (2017) run noisy indefinitely repeated prisoner's dilemma games with continuation probability $\delta = 0.875$. They examine the difference between group play and individual play when players (groups) use the strategy method by selecting (voting for) a strategy from a predefined list of 20 strategies. They find that the most commonly observed strategies are Always Defect, Lenient Grim-2, and Tit-for-2-Tats. Our work builds on this literature by designing a novel experimental interface that allows us to implement a strategy method for indefinitely repeated games with high continuation probability. Our experiment is the first that allows subjects to construct longer strategies (memory-2+) from scratch, which gives a unique perspective on how subjects behave in long repeated games.

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

2. Experimental design

The goal of this experiment is to better understand subjects' strategies in long repeated prisoner's dilemma games by using an intuitive interface that allows them to easily create complex strategies. Because the interface is different from previous experiments, we also want to test the robustness of our results both internally and externally. Internal validation comes from comparing subjects within our experiment, while external validation comes from comparing our results to results from previous studies.

2.1. Parameters

The experiment consists of two treatments (described in Section 2.6) that are used to test whether the interface impacts subjects' choices by comparing their behavior with the interface versus their behavior without it (internal validation). Each treatment has three sessions. In each session, subjects play the indefinitely repeated prisoner's dilemma for 60 supergames with continuation probability $\delta = 0.95$.[4] At the beginning of each supergame, subjects are randomly paired with one other subject and remain matched with that subject for the duration of the supergame. Since there are no participant identifiers within the game interface, participants remain anonymous throughout the experiment.

The payoff table used for all treatments is displayed in Fig. 1. The payoff table and the continuation probability were selected to ensure a direct comparison with previous results (external validation), as the same parameters were used in one of the treatments from Dal Bó and Fréchette (2017). Neutral action names, *W* and *Y*, are used throughout the experiment, though the corresponding action names *C* and *D* are used in this paper.

2.2. Rules

At certain points during the experiment, subjects will construct strategies that will be used to play the repeated prisoner's dilemma. The key feature in our experiment is the intuitive interface which allows subjects to easily construct strategies through a collection of "if-this, then-that" statements, referred to as *rules*.

[2] Embrey et al. (2016) perform strategy elicitation in a game other than the prisoner's dilemma by studying 4x4 games modeled after Bertrand and Cournot models.

[3] Bruttel and Kamecke (2012) also implement an automatic strategy elicitation (referred to as the Moore procedure) in which the computer determines aspects of the strategy played by the subject and eventually takes over (based on a procedure from Engle-Warnick and Slonim (2004)). This automatic elicitation is only a partial elicitation as it does not determine how the strategy plays off of the equilibrium path, and therefore cannot differentiate strategies such as Tit-for-Tat and Grim Trigger when matched against each other.

[4] Three sets of 60 supergame length realizations were pre-drawn before the experiment (each was used for one session in each treatment). Similar realizations across treatments allows for a cleaner comparison between the treatments, while different realizations within a treatment ensure that the results are not dependent on one specific realization. The exact realizations are displayed in the appendix in Table D2.

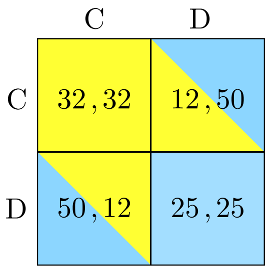


Fig. 1. Stage game payoffs.

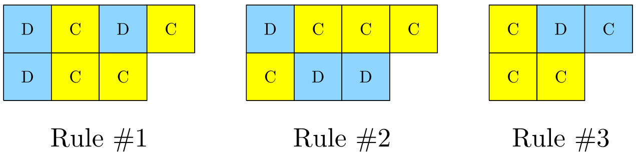


Fig. 2. Examples of rules.

Period	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45
My Choice	C	D	C	C	D	D	D	D	D	C	D	C	C	C	D	D	C	D	?
Other's Choice	C	C	C	C	C	C	D	C	D	C	C	C	D	D	D	D	C	C	?
My Payoff	32	50	32	32	50	50	25	50	25	32	50	32	12	12	25	25	32	50	

Fig. 3. Example of game history.

A 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. Fig. 2 displays some examples of rules. For example, Rule #1 has an input sequence of (*D, D*), (*C, C*), (*D, C*) and an output of *C* (to simplify notation, we denote such a sequence as $DDCCDC \rightarrow C$). Thus, if the subject plays *D, C*, and then *D* in the last three periods, and the participant with whom he is paired plays *D, C*, and then *C*, this rule will play *C* 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.

As the supergame progresses, subjects see the history of play across the top of the screen (Fig. 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) were (*D, D*), (*C, C*) and (*D, C*), 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.

Subjects create a collection of rules which is referred to as the *rule set*. If more than one rule from the rule set fits the history, then the rule with the longest length determines the choice. For example, given the history in Fig. 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 Fig. 2, the choice next period will be *C*, as prescribed by rule #1. We choose to have the longest-length rule determine the action when more than one rule fits the history because 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.

To ensure that the rule set is a well-defined strategy and always makes a unique choice, subjects are always required to have two memory-0 rules: the first-period rule and the default rule. The first-period rule is used to make the choice in the first period of a supergame, while the default rule is used to make the choice if no other rules fit the history.

Notice that the interface gives subjects flexibility to create rule sets that implement a wide variety of strategies. Although some strategies may require an infinite rule set (for an example of one such strategy see Stahl, 2011), commonly studied strategies can be constructed with a small number of rules. Fig. 4 shows several examples of rule sets that implement commonly studied strategies. For example, panel (a) shows a rule set that implements the Tit-for-Tat strategy. This rule set contains four rules: a first period rule that plays *C* (denoted *C* (first)), a default rule that plays *C* (denoted $\rightarrow C$), and two memory-1 rules $DD \rightarrow D$ and $CD \rightarrow D$. In addition, more complex strategies (those longer than memory-1) can still be constructed with relatively simple rules sets: Lenient Grim 2 can be constructed with 5 rules, Tit-for-2-Tats can be constructed with 6 rules, and Tit-for-3-Tats can be constructed with 10 rules.

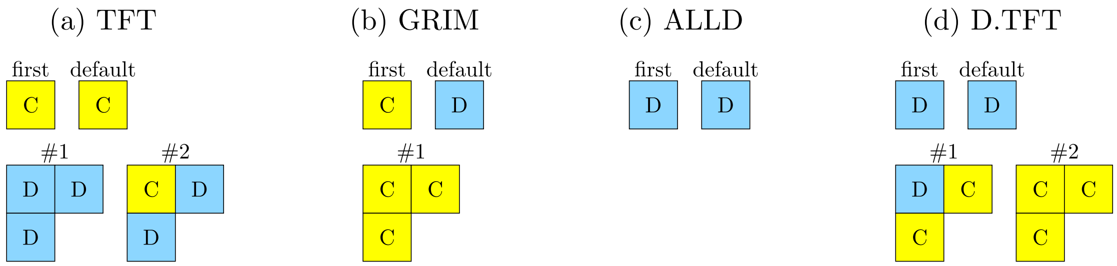


Fig. 4. Examples of rule sets. *Notes*: Commonly studied strategies: (a) TFT – Tit-for-Tat, (b) GRIM – Grim Trigger, (c) ALLD – Always Defect, (d) D.TFT – Exploitative Tit-for-Tat.

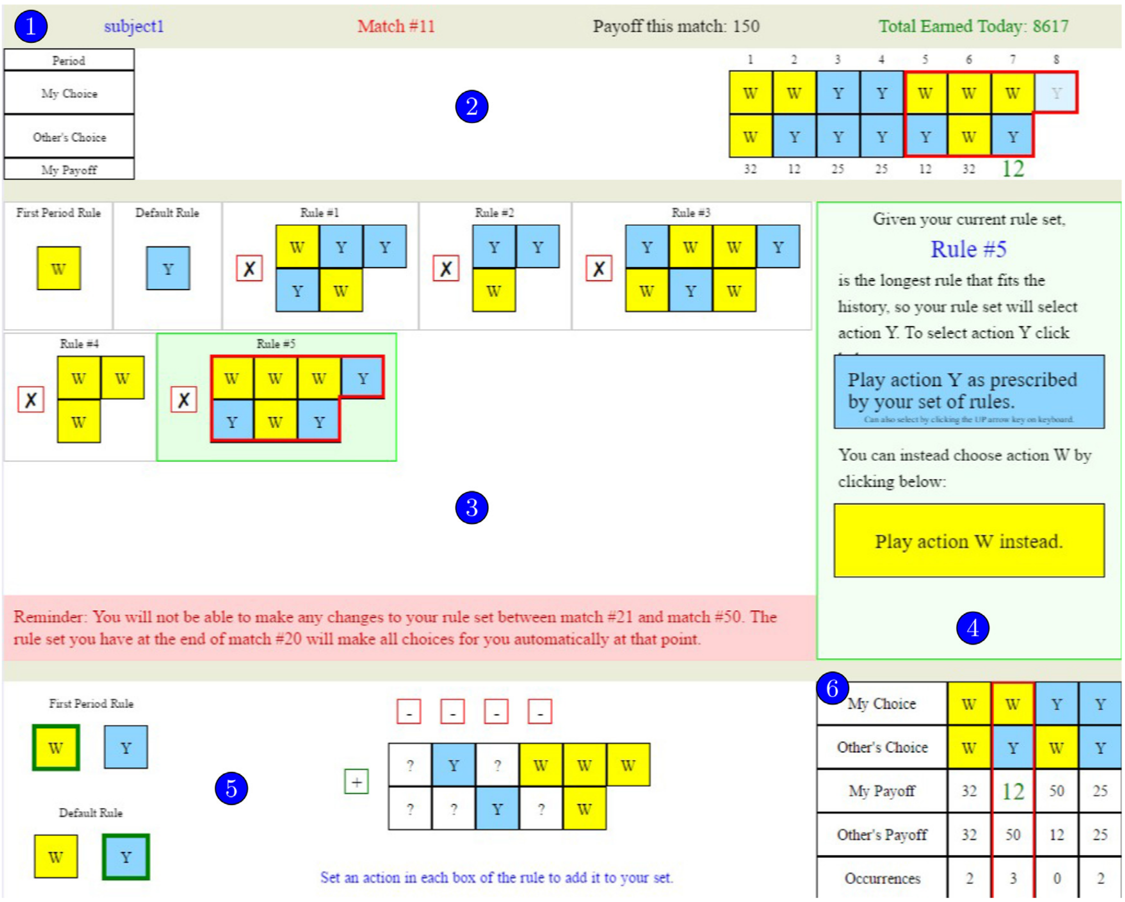


Fig. 5. Screenshot of the experimental interface. *Notes*: The screenshot shows: (1) General Information, (2) History, (3) Rule set, (4) Rule summary, (5) Rule constructor, (6) Payoff table. The neutral action names *W* and *Y* correspond to the usual action names *C* and *D* from the prisoner's dilemma.

2.3. Experimental interface

Fig. 5 shows a screenshot of the experimental interface. We developed the experimental interface using a Python server and JavaScript clients.[5] Important aspects of the interface, denoted by numbers inside a circle in the figure, are described below. All payoffs are displayed in experimental Francs.

1 *General information* - The general information includes computer ID, supergame number, payoffs in the current supergame, and total payoffs earned.

[5] We will open-source the software upon publication. Further information about the interface can be found at http://jnromero.com/strategyChoiceExperiments/.

2 *History* - The history shows the period number (labeled Period), the choice of both the subject (labeled My Choice) and the subject they are matched with (labeled Other's Choice), and the payoff earned for the corresponding period (labeled My Payoff). An outline of the longest rule that currently fits the history is overlaid on top of the history, and a preview of that rule's output is displayed under the label for the next period.

3 *Rule set* - The rule set shows all of the rules that subjects currently have in their set. If subjects create a large number of rules, they are able to scroll down to see all of their rules. The longest rule that currently fits the history is highlighted. Subjects can delete a rule from the rule set by clicking on the ✘ to the left of the corresponding rule.

4 *Rule summary* - The rule summary states which of the rules will be selected, and what action the selected rule will choose. In addition, subjects have the option to play the other action directly if they do not want to play the action selected by their current rule set and do not want to change their rule set.

5 *Rule constructor* - The rule constructor allows subjects to add rules to their rule sets. On the left side of the rule constructor, subjects can select their first-period rule and default rule. On the right side of the rule constructor, subjects can construct rules of any length by clicking the boxes. The **+** button adds a new column to the rule, while the **—** button subtracts the corresponding column. Subjects can click on the rule to set the action in each box of the rule. If a subject clicks a box with a *C*, it will switch to a *D*, and vice versa. If a subject clicks a box with a "**?**", then it will switch to a *C* or *D* randomly (neutral names *Y* and *W* are shown in the screenshot). Once subjects have constructed a complete rule (no question marks remaining), 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 as one of the rules currently in their rule set but has a different output, then they get a message that says "Conflicting rule in set," and will not be able to add that rule to the set.

6 *Payoff table* - The payoff table shows the choice of the subject (labeled My Choice), the choice of the subject that they are matched with (labeled Other's Choice), the payoffs corresponding to that action profile (labeled My Payoff and Other's Payoff), and the number of times that this action profile has occurred during this supergame (labeled Occurrences).

2.4. Instructions and quiz

Subjects were given interactive instructions equivalent to the information presented in Sections 2.1 to 2.3 above. Subjects were able to proceed through these instructions at their own pace and were able to go back and forth at will. Twenty incentivized quiz questions were given throughout the instructions. Subjects were able to attempt each question only once, after which they were given the correct answer. Subjects were told that they had 20 min to read the instructions and complete the quiz. Prior to the beginning of the instructions, subjects were provided details concerning the number of questions, incentives, and timing.

The purpose of the quiz was to ensure that subjects understood the complex experimental interface. Subjects received \$5 if they answered at least 18 out of the 20 questions correctly and received \$0 otherwise. In the experiment, all subjects that received \$5 on the quiz were put into a high-quiz group. All subjects in the high-quiz group were told the number of participants in their group, and that everyone else in their group received \$5 on the quiz.[6] All subjects that received \$0 on the quiz were put into a separate group. The analysis in the paper focuses only on the high-quiz group.[7]

The quiz covered a number of topics. There were two questions regarding the number of supergames and the number of periods in each supergame. There were six questions ensuring that subjects understood the payoff table and the history. Finally, there were a total of 12 questions regarding the rules, ensuring that subjects understood the concept of rule length, determining which rule fits the history, determining which rule would be selected from a given rule set, and editing a rule set by adding and deleting rules. A set of screenshots of the instructions and quiz questions is given in Appendix A.

2.5. Experimental stages

The experiment consists of three different stages: the direct-response stage, the non-binding stage, and the locked-response stage. In the direct-response stage, subjects play the repeated prisoner's dilemma by choosing *C* or *D* directly. In the non-binding stage, subjects create rule sets, but are able to directly choose any desired action. In the locked-response stage, subjects' rule sets are locked, and choices are made automatically. These three stages are similar to the three phases from the experiments in Dal Bó and Fréchette (2017). In all stages, subjects see the general information, history, and payoff table. We describe the three stages in detail below.

Direct-response stage

In the direct-response stage, subjects choose *C* or *D* directly. Since subjects do not construct rules during this stage, they do not see the rule set, the rule summary or the rule constructor. A screenshot of the direct-response stage is displayed in

[6] If there were an odd number of subjects that received \$5 on the quiz, then one random subject was grouped with those that received \$0 on the quiz.

[7] Subjects in the low-quiz group did not understand certain aspects of the interface, and therefore did not construct rules during the experiment, but nonetheless played a regular indefinitely repeated prisoner's dilemma game.

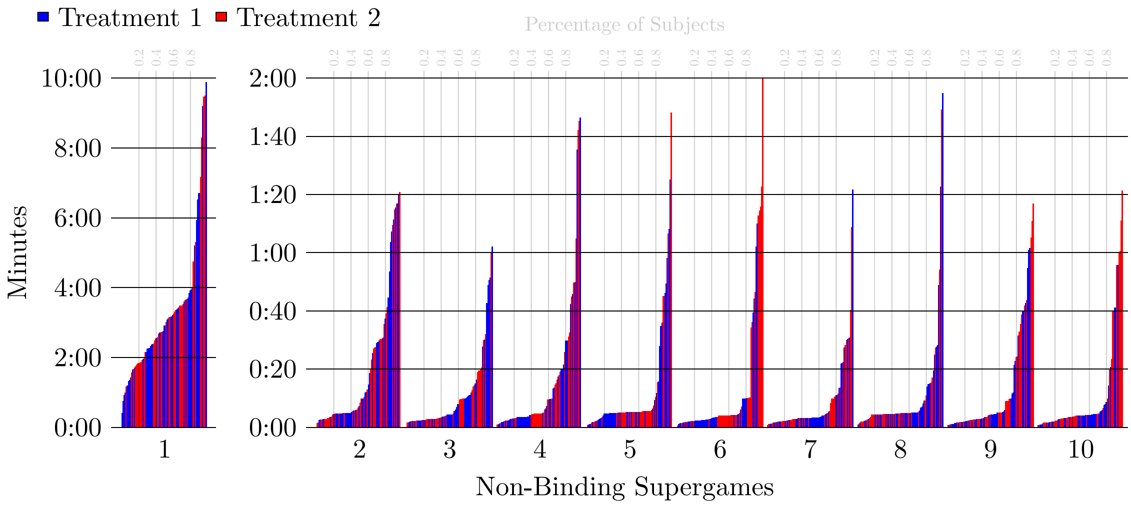


Fig. 6. Elapsed time before clicking "Start Match" button. *Notes*: Subjects are given ten minutes before the first supergame and two minutes before every other supergame of the non-binding stage.

the Appendix in Fig. C1. Subjects make their choice by either clicking one of two large buttons on the screen or pressing buttons on the keyboard.[8]

Non-binding stage

In the non-binding stage, subjects create rule sets but can choose to go against their rule set at any point. In this stage, subjects see all parts of the interface, as displayed in Fig. 5.

Before each supergame in the non-binding stage, subjects are given time to edit their rule sets. Specifically, subjects are given ten minutes before the first supergame and two minutes before every other supergame. If subjects do not need the entire time to edit their rule sets, they can click the "Start Match" button. After clicking this button, subjects are still able to make edits to their rule sets until either all other subjects have clicked the "Start Match" button or the allotted time runs out. Fig. 6 shows the distribution of length of time until the "Start Match" button is pressed before each supergame of the non-binding stage. This figure shows that a large majority of subjects clicked the "Start Match" button less than half-way through their allotted time. In addition, only one subject took the entire allotted time (prior to supergame 6), suggesting that subjects had sufficient time to edit their rule sets prior to each supergame.

After subjects have edited their rules, the supergame begins. Subjects have unlimited time to make a choice in each period. Subjects can also continue to edit their rule sets during the supergame. The rule summary tells the the rule that will be selected (the longest rule that fits the history) and the prescribed action given the current rule set. When subjects edit their rule set, the rule summary is automatically updated. The rule summary contains two buttons: one which confirms the action prescribed by the rule set and another which selects the other action.[9] Thus, subjects can make a choice in three different ways: by confirming the action prescribed by their rule set; by choosing the action not prescribed by their rule set; or by changing their rule set to play a different action, and confirming that action. The ability to view the action prescribed by their rule set and edit their rule set while they are playing the game allows subjects to construct a rule set that matches the strategy underlying their play.[10]

Locked-response stage

In the locked-response stage, subjects' rule sets make choices for them automatically. Subjects are not able to edit their rule sets at any point during the locked-response stage (neither between supergames nor during supergames). Since subjects are not able to edit their rules, the rule constructor is not displayed during this stage. A screenshot of the locked-response stage is displayed in the Appendix in Fig. C4.

Although subjects cannot edit their rule sets during the locked-response stage, they can still observe rules chosen and actions taken by their 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 supergame finishes. Since this sequence is deterministic,

[8] Subjects could select the left box by clicking the left arrow and the right box by clicking the right arrow. In order to make a choice, subjects had to press and release the button, which ensured that subjects were not able to make choices for multiple periods with one press of the button.

[9] Subjects can confirm the action prescribed by their rule set by clicking the top button or pressing and releasing the up arrow on the keyboard. Subjects can choose the other action only by clicking the bottom button.

[10] We find that subjects choose the action that is not prescribed by their rule set 4% of the time across all non-binding supergames, and 2% of the time in the last non-binding supergame.

Table 1
Treatments summary.

Treatment	Sessions	Subjects			Supergames				
		HQ	LQ	Total	1–10	11–20	21–30	31–50	51–60
1	3	44	14	58	DR	NB	LR	LR	DR
2	3	38	20	58	DR	DR	NB	LR	LR

Notes: HQ – subjects in the high-quiz group. LQ – subjects in the low-quiz group. DR – direct-response stage. NB – non-binding stage. LR – locked-response stage.

players do not need to watch it be played repeatedly. Therefore, to expedite the lock stage, we gradually increase the speed of each period from one second per period at the beginning to 0.1 seconds per period at the end, thus allowing 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.

2.6. Treatments

The experiment consists of two treatments, allowing us to test the impact of our interface on subjects' choices. The two treatments are presented in Table 1.

Treatment 1 begins with 10 supergames of the direct-response stage, followed by 10 supergames of the non-binding stage, then 30 supergames of the locked-response stage, and ends with 10 additional supergames of the direct-response stage. Treatment 2 begins with 10 supergames of the direct-response stage, followed by 10 additional supergames of the direct-response stage, followed by 10 supergames of the non-binding stage, and ends with 30 supergames of the locked-response stage.

Subjects knew that they were going to play 60 supergames but were not told about the three stages during the instructions. Prior to each ten supergames of the direct-response stage, subjects received a message saying that they will be making choices directly for the next ten supergames. For example, prior to the first supergame, subjects received the following message: "In matches 1–10 you will make choices directly without constructing rules." Prior to the non-binding stage, subjects were told that they would play 10 supergames of the non-binding stage followed by 30 supergames of the locked-response stage. For example, prior to supergame 11 in treatment 1, subjects received the following message: "In matches 11–20 you can construct rules and edit your rule set. You will be told which action your rule set will make, but can still choose the other action if you want. In matches 21–50 you will NOT be able to edit your rule set. The rule set that you have at the end of match #20 will make all choices for you automatically." The second half of this message remained on the screen throughout the non-binding stage and was reiterated with an alert when subjects played the action not prescribed by their rule set (a screenshot of the alert is given in the Appendix in Fig. C3).

This sequence of stages was selected for several reasons. First, the two treatments start out identically for the first ten supergames, allowing subjects to learn how to play the indefinitely repeated prisoner's dilemma. Second, the two treatments diverge in supergame 11, allowing a comparison between how subjects play when they are constructing rules versus how they play regularly via direct response (i.e., comparing supergames 11–20 of treatment 1 and treatment 2). Third, the 30 supergames of the locked-response stage account for half of the possible payoffs in the experiment (excluding the quiz), thereby giving subjects strong incentives to create rule sets in the non-binding stage. Finally, we can check to see if play has converged by comparing supergames where one treatment is in the locked-response stage and the other is not (supergames 21–30 and 51–60).

2.7. Administration

We recruited 116 students for the experiment using ORSEE software (Greiner, 2015) on the campus of Purdue University. Six sessions of the experiment were administered in March and April 2017, with the number of participants varying between 18 and 22. Upon entering the lab, the subjects were randomly assigned to a computer. After all of the subjects had been seated, interactive instructions began. Each session, including the instructions portion and the quiz, took about one hour and 30 min to complete,[11] with an average earnings from the experiment of $16.68. During the experiment, all payoffs were presented in experimental Francs with an exchange rate of 2500 experimental Francs to 1.00 USD.

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. And, finally, in

[11] One of the sessions in treatment 1 ran out of time, so subjects did not participate in the last direct-response stage and therefore only participated in 50 supergames.

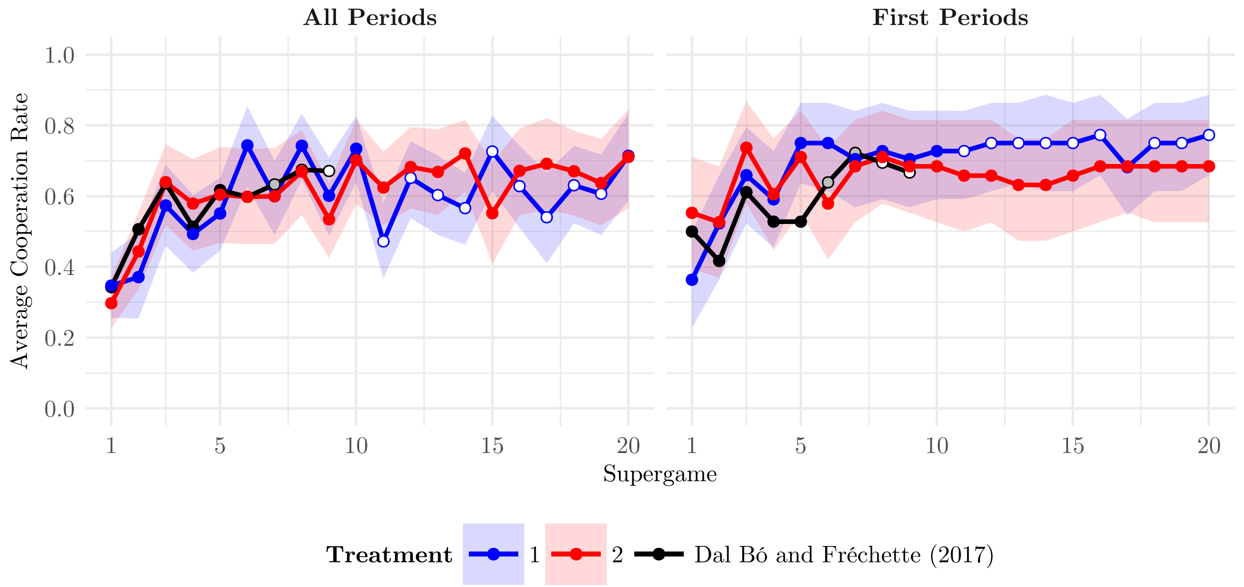


Fig. 7. Evolution of cooperation. *Notes:* The cooperation rate is the fraction of periods that a subject cooperated in the given range of periods. Solid points represent direct-response stage. White points represent non-binding stage. 95% bootstrapped confidence intervals are superimposed. Dal Bó and Fréchette (2017) run two sessions with different numbers of supergames in each stage. Gray points for (supergames 6 and 7) represent one session being in the direct-response stage while the other being in the non-binding stage.

Table 2
Average cooperation rate across supergames. *Notes:* The unit of observation is the average cooperation rate by a subject over the corresponding range of supergames. The cooperation rate is the fraction of periods that a subject cooperated in the given range of periods. Bootstrapped standard errors are in parentheses. If the supergame is less than four periods, then the cooperation rate for the first four and the last four is set to the cooperation rate for all periods. Tests between treatments are carried out using non-parametric permutation tests (Good, 2013). Tests within treatments are carried out using non-parametric randomization tests (Good, 2006). *p*-values are provided for each test. $<$, $\ll$, $\lll$ denote significant difference at the 0.10, 0.05, and 0.01 levels, respectively. $\sim$ denotes no significant difference. * As noted in footnote 11, one session in treatment 1 ran out of time and completed only 50 supergames, so treatment 1 has only 26 observations for supergames 51–60.

Treatment	1		2	1		2	1		2	1		2
Supergames		11–20			21–30			31–50			51–60	
Type	NB		DR	LR		NB	LR		LR	DR*		LR
First periods	0.75 (0.06)	$\overset{0.38}{\sim}$	0.67 (0.07)	0.77 (0.06)	$\overset{0.66}{\sim}$	0.73 (0.07)	0.77 (0.06)	$\overset{0.8}{\sim}$	0.74 (0.07)	0.86 (0.06)	$\overset{0.22}{\sim}$	0.74 (0.07)
First 4 periods	0.67 (0.04)	$\overset{0.55}{\sim}$	0.71 (0.05)	0.66 (0.05)	$\overset{0.42}{\sim}$	0.72 (0.05)	0.65 (0.05)	$\overset{0.38}{\sim}$	0.71 (0.05)	0.825 (0.05)	$\overset{0.2}{\sim}$	0.72 (0.05)
	$\overset{0.0}{\ggg}$		$\overset{0.0}{\ggg}$	$\overset{0.0}{\ggg}$		$\overset{0.0}{\ggg}$	$\overset{0.0}{\ggg}$		$\overset{0.0}{\ggg}$	$\overset{0.01}{\ggg}$		$\overset{0.0}{\ggg}$
Last 4 periods	0.58 (0.04)	$\overset{0.44}{\sim}$	0.63 (0.05)	0.55 (0.05)	$\overset{0.46}{\sim}$	0.6 (0.04)	0.51 (0.04)	$\overset{0.08}{<}$	0.62 (0.05)	0.71 (0.05)	$\overset{0.28}{\sim}$	0.63 (0.05)
All periods	0.61 (0.04)	$\overset{0.41}{\sim}$	0.66 (0.05)	0.58 (0.05)	$\overset{0.31}{\sim}$	0.65 (0.04)	0.56 (0.04)	$\overset{0.14}{\sim}$	0.65 (0.05)	0.76 (0.05)	$\overset{0.21}{\sim}$	0.67 (0.05)

Section 3.3, we focus on rule sets and the resulting strategies. Specifically, we use a cluster analysis to investigate which strategies participants constructed.

3.1. Cooperation

We first examine the evolution of cooperation across supergames. Fig. 7 presents the average cooperation rates over the first 20 supergames in each of the two treatments. The figure also contains experimental results from a treatment in Dal Bó and Fréchette (2017) that used the same parameters as in our experiment. In addition, Table 2 shows the average cooperation rates in the first period, the first four periods, the last four periods, and all periods in supergames 11–20, 21–30, 31–50, and 51–60 in each of the two treatments. The cooperation rate is the fraction of periods that a subject cooperated in the given range of periods.

Several findings are worth noting. First, both Fig. 7 and Table 2 suggest that there is no difference in cooperation rate between the two treatments. Second, Fig. 7 shows no difference between the cooperation rates in our experiments and those from a similar treatment in Dal Bó and Fréchette (2017). Finally, Table 2 shows that the cooperation rate is decreasing from the beginning to the end of the supergame. Table 2 also suggests that subjects may still be learning throughout the locked-response stage. In what follows, we elaborate on each of these findings and summarize the main results.

Remark 1. The experimental interface has both internal and external validity.

The experimental design allows us to check for internal validity of the experimental interface by comparing how subjects play when they are allowed to construct rules versus how they play when they make choices directly. In treatment 1, subjects make choices directly in supergames 1–10 and then construct rules in supergames 11–20, while in treatment 2, subjects make choices directly in supergames 1–20. If we see no difference in the cooperation rates in supergames 11–20 of the two treatments, then this suggests that rule construction does not affect how subjects behave. Comparing the columns corresponding to supergames 11–20 in Table 2, we find no significant difference in cooperation rates between the two treatments, using the non-parametric permutation test.[12,13]

The choice of parameters allows us to check for external validity of the experimental interface by comparing cooperation rates from our experiment with those from Dal Bó and Fréchette (2017). Fig. 7 shows a close match between cooperation rates from our experiment and the corresponding treatment in Dal Bó and Fréchette (2017). To further confirm this, we run a permutation test and find no difference in the average cooperation rates in the first period (p-value 0.44) and the average cooperation rates over all periods (p-value 0.66). These results suggest that our experimental interface does not change how subjects behave at the aggregate level.

In the Appendix, we also investigate average cooperation rates after each of the four possible memory-1 histories in supergames 11–20 (Table E4). We find no difference in average cooperation rates between the two treatments after the CC, DD, and DC histories, but a significant difference in average cooperation rate after the CD history. In particular, subjects cooperate 47% of the time in the direct response supergames and 28% of the time in the non-binding supergames, which leads to a significant difference at $p = 0.01$. This difference could be due to the fact that it is easier to make an immediate, intended response to the opponent's action if you have a rule that can do it automatically. However, we are somewhat concerned about making conclusions with such fine partitions of the data, especially for the CD history which accounts for only 7.5% of memory-1 histories.[14]

Result 1. Cooperation decreases from the beginning to the end of a long, indefinitely repeated prisoner's dilemma game.

Using a non-parametric randomization test, we compare the rows corresponding to cooperation rates in the first four periods and the last four periods in Table 2 and find that cooperation significantly decreases from the beginning of the supergame to the end of the supergame.[15] This decrease has been noted in non-noisy (Bruttel and Kamecke, 2012; Lugovskyy et al., 2017) and noisy (Aoyagi et al., 2015; Cason and Mui, 2017; Fudenberg et al., 2012) indefinitely repeated prisoner's dilemma games, although these previous experiments have used lower continuation probabilities, which lead to shorter games in expectation.[16] We find that in long, non-noisy, indefinitely repeated prisoner's games, cooperation often breaks down. We attribute this breakdown in cooperation to subjects using a specific class of rules that intentionally defects after a long sequence of mutual cooperation. These rules may be useful to check whether the other player's strategy can be exploited, which is especially useful in long repeated games because the potential gain from exploitation is high. These rules are further analyzed in Section 3.2, and the impact of these rules on cooperation will be further analyzed in Section 4.

Our experiment was designed to allow us to check whether the cooperation rates converged or if the learning persists across all 60 supergames. In supergames 21–30, the subjects in treatment 1 enter the locked-response stage and are not able to change their rule sets while the subjects in treatment 2 play the non-binding stage and still are able to make changes. Similarly, in supergames 51–60, the subjects in treatment 1 have a direct-response stage (after finishing the locked-response

[12] Permutation tests do not rely on any assumptions regarding the underlying distribution of the data (Good, 2013). The null hypothesis in the permutation test is that there is no difference in participants' behavior across the two treatments and, therefore, the treatment labels are interchangeable. The distribution of the test statistic (which, in our case, is difference in the average cooperation rate between the two treatments) is obtained through random permutation of treatment labels among observations (Phipson and Smyth, 2010). The p-value for a two-sided test is then determined by finding the fraction of permutations that have the test statistic with a greater absolute value than that of the original realization. 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 using 10,000 permutations of the labels; we then used this distribution to test where the original difference lies. The unit of observation is the average cooperation rate of a single subject across the given range of supergames.

[13] In addition to the results of permutation tests, we provide the results of Wilcoxon rank-sum test in Appendix E.

[14] The distribution of CD histories is not uniform across supergames, so the overall percentage may overestimate the how often a given subject experiences this history. For example, the median number of times a subject experiences CD history in a supergame is 0, and 20% of supergames account for 75% of occurrences of CD.

[15] The 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 participant: (i) the responses at the beginning and the end of each supergame are as selected; and (ii) the responses are flipped from first to last and from last to first. We sampled each participant 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.

[16] Lugovskyy et al. (2017) studies two- and four-person indefinitely repeated public goods game.

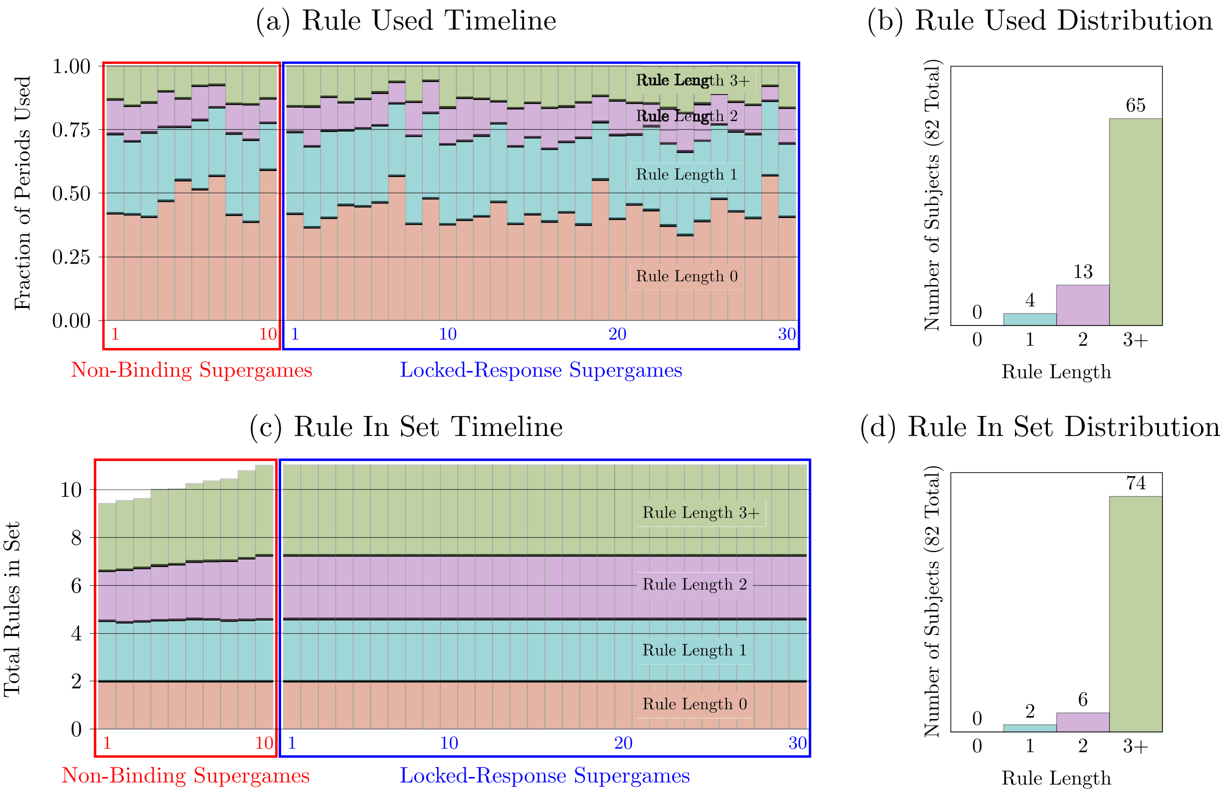


Fig. 8. Rule lengths. *Notes*: **(a)** Average fraction of periods that a rule of a given length was used to make a choice within a supergame. For example, in the first supergame of the non-binding treatment, 42 percent of choices were made with rules of length 0, 31 percent of the choices were made with rules of length 1, 14 percent of the choices were made with rules of length 2, and the remaining 13 percent of the choices were made with rules of length 3 or more. **(b)** Longest rule of length $l \in \{0, 1, 2, 3+\}$ used in the experiment. For example, four subjects (out of 82) relied on at most memory-1 rules in the experiment. **(c)** Average number of rules in the rule set at the beginning of non-binding and locked-response supergames. **(d)** Longest rule of length $l \in \{0, 1, 2, 3+\}$ "in set" at the beginning of non-binding and locked-response supergames.

stage), while the subjects in treatment 2 remain in the locked-response stage. Table 2 shows that there is no significant difference in play between the two treatments in both supergames 21–30 and supergames 51–60 (though the DR data is limited in supergames 51–60, as mentioned in footnote 11). However, in treatment 1, the cooperation rate in all periods increases significantly from supergames 31–50 to supergames 51–60 (0.56 to 0.76, p-value 0.00). This suggests that subjects may still be learning during the locked-response stage. This learning is not studied further in this paper, but could be an interesting avenue for future research.

3.2. Constructed rules

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

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: $\{C$ (first), $\rightarrow C$, $DD \rightarrow D$, $CD \rightarrow D\}$. If this set is matched against another cooperative strategy, then the default rule ($\rightarrow C$) is played every period after the first period, but the longer rules ensure that it is not susceptible to exploitation.

Fig. 8 presents the data on the distribution of rules constructed and used during our experiment. Panel (a) shows the average fraction of periods that a rule of a given length was used to make a choice within a supergame. Panel (b) shows the distribution of the longest rule that each subject used. Panel (c) shows the average number of rules of a given length in the rule set at the beginning of each supergame. Finally, panel (d) shows the distribution of the longest rule that each subject had in her rule set at any point in the experiment.

Result 2. 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. Panel (a) of Fig. 8 shows that the subjects used memory-0 and memory-1 rules roughly 75% of the

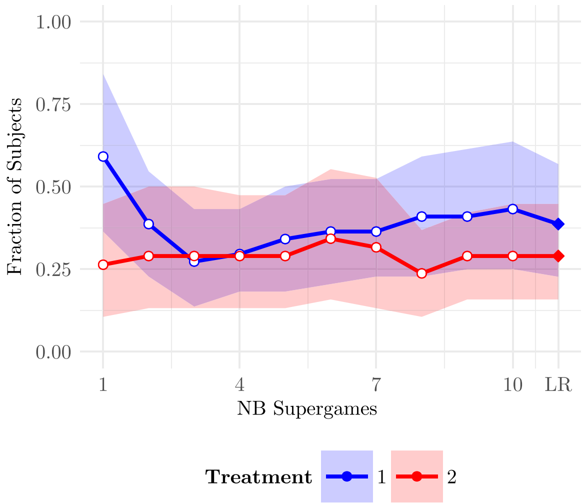


Fig. 9. Subjects with *CsToD* rules. *Notes:* Fraction of subjects that had a *CsToD* rule "in set" at least once during each of the non-binding supergames. *LR* denotes first supergame of the locked-response stage, at which point the rule sets become fixed.

time and longer rules the other 25% of the time. Panel (b) shows that none of the subjects used only memory-0 rules; four subjects used, at most, a memory-1 rule; 13 subjects used, at most, a memory-2 rule; while 65 out of 82 subjects used a rule longer than memory-2 at least once during the experiment. Panel (c) shows that the total number of rules is increasing from 9.48 in the first supergame of the non-binding stage to 11.02 in the last supergame of the non-binding stage. By the first supergame of the locked-response stage, participants had two memory-0 rules,[17] 2.60 memory-1 rules, 2.66 memory-2 rules, and 3.78 memory-3+ rules in any given period. Panel (d) shows that 74 out of the 82 subjects had a rule of at least memory-3 in their rule set at some point during the experiment.

As alluded to in the previous section, we find evidence of a specific class of long rules. These rules prescribe a defection following a sequence of at least two periods of mutual cooperation. We will refer to this type as the *CsToD* rule. These rules do not match strategies that are currently studied in the literature on the indefinitely repeated prisoner's dilemma games (Dal Bó and Fréchette (2011); Fudenberg et al. (2012)).[18] Fig. 9 presents information on the prevalence of these rules throughout the experiment. Specifically, it shows the fractions of subjects with a *CsToD* rule in their rule set in each supergame during the non-binding stage and into the first supergame of the locked-response stage.

We find that 40 out of 82 subjects had a *CsToD* rule in their rule set at some point during the experiment, and 33 out of 82 subjects used it at some point in the experiment. Furthermore, we find that the fraction of subjects with a *CsToD* rule in their rule set consistently hovered above 0.25 (see Fig. 9). The presence of these rules could potentially cause cooperation to break down during a supergame. For example, if both players are playing Grim Trigger, and one has the *CsToD* rule, then they would start by both cooperating at the beginning of the supergame; eventually, the *CsToD* rule would be used, causing a *D* to be played, which would cause both players to eventually play *D* for the remainder of the supergame. Therefore, the presence of these rules is consistent with the cooperation breakdown found in *Result 1*.

Note that we do not provide any further analysis of specific rules because it can be difficult to draw conclusions from them for two reasons. First, rule sets with common rules may lead to vastly different outcomes. For example, the rule sets {C (first), $\rightarrow C$, $CC\rightarrow C$} and {D (first), $\rightarrow D$, $CC\rightarrow C$} both have the rule $CC\rightarrow C$, but one implements the ALLC strategy, while the other implements the ALLD strategy. Second, rule sets with different rules may lead to the same outcome. For example, the rule sets {C (first), $\rightarrow C$, $DD\rightarrow D$} and {C (first), $\rightarrow C$, $DD\rightarrow C$} both implement the ALLC strategy, even though they have the conflicting rules $DD\rightarrow C$ and $DD\rightarrow D$. Therefore, the presence of a specific rule in the rule set does not necessarily tell us much about the strategy that the subject is using. Because of these difficulties, we focus on the entire rule set and provide an analysis of strategies in the next section.

3.3. Constructed strategies

As reported in the previous section, we find evidence that subjects use rules that do not fit within the set of strategies commonly studied in the literature on indefinitely repeated games. In this section, we map the subjects' complete rule

[17] Recall that subjects always had exactly two memory-0 (first period and default) rules present in all treatments – that is why the average number of rules of length zero is exactly two.

[18] Embrey et al. (2015) studies finitely repeated prisoner's dilemma games, and observes threshold strategies that conditionally cooperate until a threshold period. A *CsToD* rule is needed in order for a rule set to implement these strategies.

sets to strategies. This analysis can be used to compare strategies that are observed in our experiment 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 {*D* (first), $\rightarrow D$} and {*D* (first), $\rightarrow D$, *DCDCDCDCDCDCDC* $\rightarrow C$}. They are identical after every history that does not end with seven periods of (*D, C*). So, although these rule sets are not equivalent, they lead to very similar behavior. Therefore, we need a measure of the similarity of rule sets that would help us 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 the vector of actions (referred to as an output vector) that is played when the rule set is matched against a fixed set of opponents. The similarity of two rule sets is then determined by looking at the Manhattan distance between the two output vectors. In our case, the Manhattan distance is the number of periods in which the two rule sets play different actions.[19]

The "opponents" we consider are fixed sequences of actions. If we could use all possible sequences of length *L*, then we could perfectly differentiate any two strategies that may play differently against the same sequence of actions in the first *L* periods. However, since we are dealing with indefinitely repeated games, there are an infinite number of possible sequences, so it isn't feasible to use all sequences. Instead, we use a large number of sequences that have a variety of different properties. Specifically, we use sequences that are generated using a Markov transition probability matrix, *P*:

$$P = \begin{array}{c c} & \begin{array}{cc} C & D \end{array} \\ \begin{array}{c} C \\ D \end{array} & \begin{bmatrix} a & 1-a \\ 1-b & b \end{bmatrix} \end{array}, \tag{1}$$

where a and b are drawn randomly from a uniform distribution on [0,1]. We randomly draw 4000 pairs of a and b and generate a random sequence of actions for each pair. Each sequence starts in a state determined randomly, given the stationary distribution of *P*. The length of each sequence is determined randomly using a continuation probability of $\delta = 0.95$ (as in the experiment). Next, we simulate each of the participants' strategies against these sequences, which generates 82 (44 for treatment 1 and 38 for treatment 2) vectors in $\{0, 1\}^N$, where the expected length of these vectors is $\mathbb{E}[N] = 20 \times 4000$. This process generates a wide variety of behaviors for the opponents, which allows differentiation between subjects' rule sets.[20,21]

We compare each of the participants' strategies to 20 commonly studied strategies (see, e.g., Fudenberg et al., 2012, and Appendix H for a description of the strategies) in three ways. First, we determine whether any of the participants'

---

[19] An alternative approach would be to define a distance measure on the rule sets themselves (instead of the outputs of the rule sets). The flexibility of our interface allows subjects to create the same strategy in many different ways. For example, the rule set {*C* (first), $\rightarrow C$, $DD \rightarrow D$, $CD \rightarrow D$} and the rule set {*C* (first), $\rightarrow D$, $CC \rightarrow C$, $DC \rightarrow C$} both implement TFT. A distance measure on the rule sets would view these two sets as being different. The approach that we take will always view these two rule sets as being identical, because they lead to the same output for any possible input.

[20] If 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.

[21] Possible alternatives for inputs into the clustering algorithm could be: i) realized sequences from the experiment; or ii) subjects' rule sets from the experiment. We opt for the random Markov sequence generation process, because they provide more variability in the inputs, which allows for better differentiation of strategies by the clustering algorithm. For example, the eight most common strategies according to both the clustering and the maximum likelihood estimation are ALLD, TFT, D.TFT, GRIM, TF2T, 2TFT, ALLC, and GRIM2. These eight strategies account for 92% of the subjects using maximum likelihood estimation and 99% of the subjects using the clustering approach. When we create all of the 64 possible matchings using these eight strategies, we obtain only eight different sequences:

1. All Cs (TFT when matched with GRIM, among others)
2. All Ds (ALL D when matched with D.TFT, among others)
3. One C then all Ds after that. (GRIM when matched with ALLD, among others)
4. One D then all Cs after that. (D.TFT when matched with TF2T, among others)
5. Two Cs then all Ds after that. (TF2T when matched with ALLD, among others)
6. One D, one C, then all Ds after that (D.TFT when matched with GRIM, among others)
7. Alternate between C and D starting with C (TFT when matched with D.TFT).
8. Alternate between D and C starting with D (D.FTF when matched with TFT).

If we use either of the two alternative approaches, then a large percent of the inputs to the clustering algorithm are going to match these eight sequences. This will make it difficult to differentiate among certain types of strategies. For example, the output sequences for strategies such as GRIM3 and TF3T are going to be identical when the eight sequences are used as inputs. In order to differentiate between these two strategies, we need a sequence that plays C after playing D at least three times, so a sequence that is persistent in D, but randomly plays one C every once in a while (low value of a, but high values of b). Though there may be other sequences, a large majority of the sequences observed in the experiment will look like the eight above, which will make it difficult to differentiate among certain strategies. We believe that the random inputs proposed in this paper provide the greatest differentiation between strategies. Another benefit of the random sequences is that it allows the clustering approach to be used in theoretical settings in which there may not be any data. For example, if you want to group the set of four-state automata into similar behavioral types, then this algorithm could be used without any experimental data.

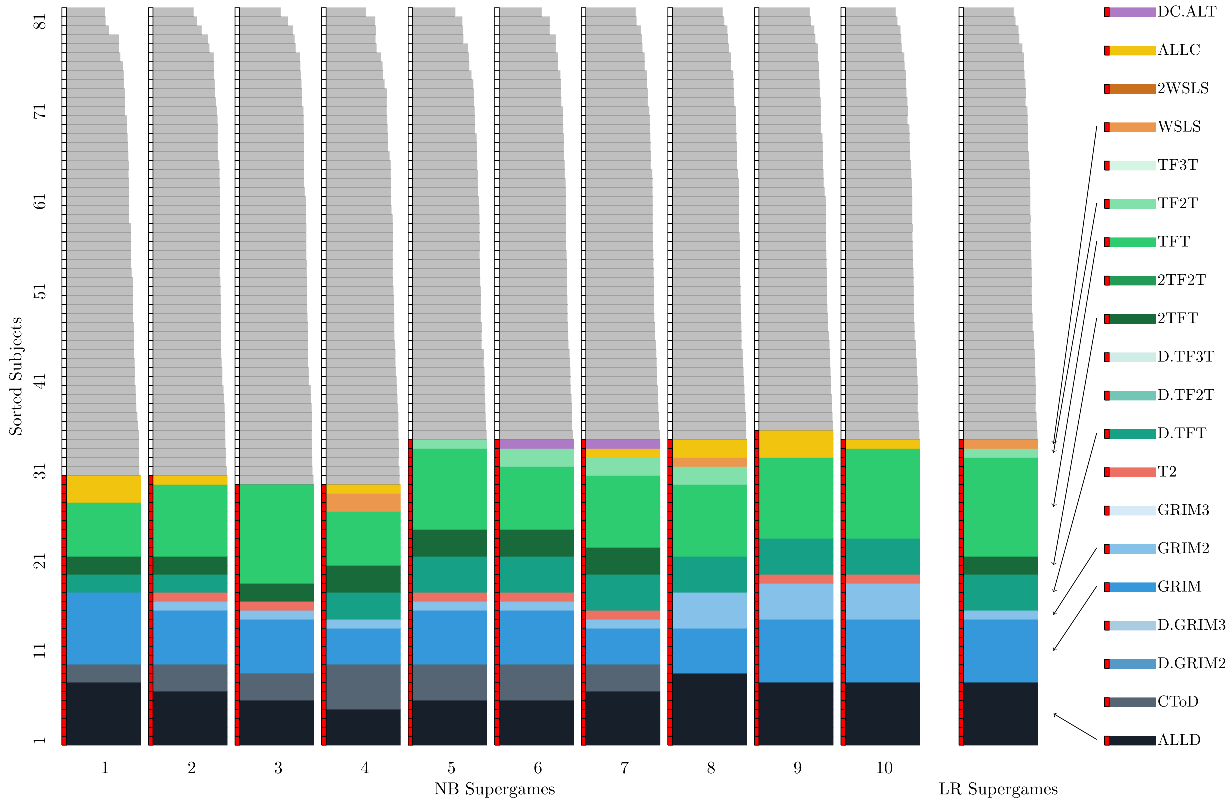


Fig. 10. Exact strategies at the beginning of each supergame. *Notes*: Rule sets are color coded based on the strategy that they match. After simulating a strategy against a fixed set of sequences, we compare the resulting action sequence to an action sequence obtained from simulating each of the 20 commonly studied strategies. Gray bars denote percent match for the "not exact" strategies relative to the closest strategy among the 20. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 3
Strategy frequencies. *Notes*: MLE estimates and cooperation rates are based on the first 20 periods of supergames 31–50 of the experiment. If a supergame had less than 20 periods then all period were used. Bootstrapped standard errors are shown in parentheses. Strategies with estimates below 0.04 in the MLE are not included.

Type	TFT	GRIM	ALLD	D.TFT	2TFT	TF2T	ALLC	β	Cooperation	# of Subjects
MLE	0.33^{***} (0.05)	0.14^{***} (0.05)	0.12^{***} (0.04)	0.10^{***} (0.03)	0.08^{***} (0.03)	0.07^{**} (0.03)	0.05^{**} (0.02)	0.96	0.61	82
Clustering	0.28	0.20	0.12	0.12	0.09	0.05	0.05	-	0.61	82

strategies match the play of one of the 20 strategies exactly (Fig. 10). Second, we classify strategies into similar groups (clusters), and then compare those groups to the 20 strategies (Fig. 11). Finally, as a check, we compare the results from the cluster analysis to the strategy frequency estimation (Dal Bó and Fréchette, 2011) commonly used in the literature (Table 3).

As a first step in better understanding the strategies that subjects construct, we determine whether any of these strategies match one the 20 commonly studies strategies. Fig. 10 shows the number of subjects that have a rule set that exactly matches one of the 20 strategies at the beginning of each of the ten supergames during the non-binding stage and the first supergame of the locked-response stage. Out of the 20 commonly studied strategies, subjects construct 13 at some point during the experiment. The most common exact strategies are TFT, GRIM, and ALLD. Despite these exact matches, the majority (approximately 60%) of subjects' constructed rule sets do not match any of the 20 strategies exactly. However, this analysis is somewhat restrictive. For example, an ALLC strategy with an additional *CsToD* rule would not be considered an

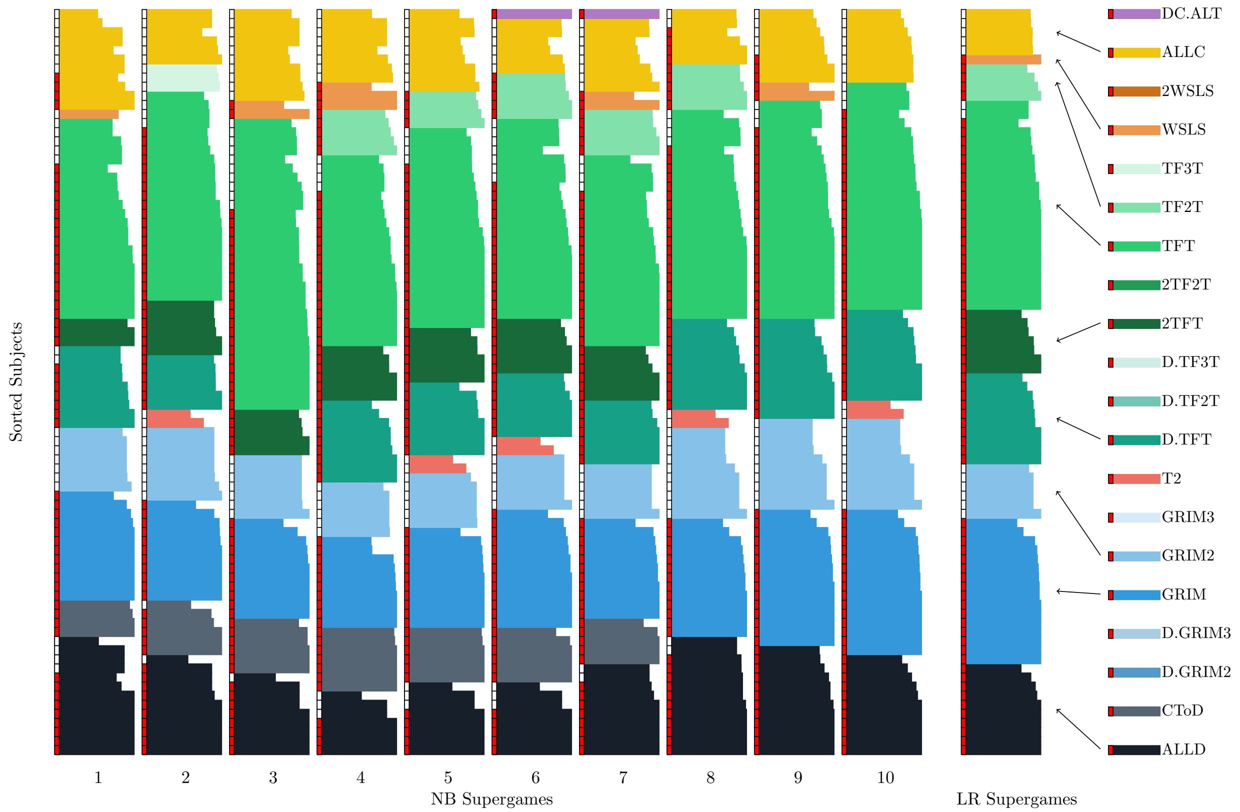


Fig. 11. Clusters at the beginning of each supergame. *Notes*: Clusters of rule sets are color coded based on the strategy that is most similar to the exemplar of the given cluster. Horizontal length of bars denotes % match relative to the strategy that is closest to the cluster exemplar. Red (white) line at the base of each bar denotes that the cluster exemplar matches (does not match) the closest strategy exactly. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

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 commonly studied strategies is closest to the exemplar strategy.[22] Fig. 11 presents the results.

Result 3. The four most common clusters of strategies are represented by memory-1 strategies TFT, GRIM, ALLD, and D.TFT.

Fig. 11 shows the number of subjects in each cluster and which of the 20 commonly studied strategies is closest to each cluster. In addition, a red (white) line at the base of each bar denotes whether the cluster exemplar matches (does not match) the closest strategy exactly. For example, in the first supergame of the locked-response stage, there are a total of 17 subjects in the cluster with an exemplar at GRIM, and the exemplar exactly matches GRIM.

The first row of Table 3 shows the maximum likelihood estimates of the proportion of strategies in the population. The second row of Table 3 shows the frequencies of strategies identified by the cluster analysis for comparison to the MLE estimates. Further description and discussion of the maximum likelihood estimation procedure can be found in Appendix G.

[22] We implement affinity propagation in R using APCluster package (Bodenhofer et al., 2011). In the simplest form the function takes two arguments: input data, and similarity function. For the input data we take the vectors generated from the simulation procedure. For the similarity function we take the Manhattan distance. There are other parameters that could be provided by the user. In particular, the parameter considered to be most important other than data and similarity function, is the likelihood of a data sample to become an exemplar. We use the default value which is the median of the input similarities.

Table 4
Simulated cooperation rates. *Notes:* ALL – all rules in the subjects' rule sets are kept. M1 – all memory 2+ rules are removed from subjects' rule sets. The data for the ALL column is taken from the LR stage of the experiment, while the data from the M1 column is simulated (denoted SIM) using the exact same matchings and supergame lengths as in the experiment. The unit of observation is the average cooperation rate by a subject. The cooperation rate is the fraction of periods that a subject cooperated in the given range of periods. Bootstrapped standard errors are in parentheses. If the supergame is less than four periods, then the cooperation rate for the first four and the last four is set to the cooperation rate for all periods. Tests between "ALL" and "M1" are carried out using non-parametric permutation tests (Good, 2013). Tests within each of "ALL" and "M1" are carried out using non-parametric randomization tests (Good, 2006). p-values are provided for each test. $<$, $\ll$, $\lll$ denote significant difference at the 0.10, 0.05, and 0.01 levels, respectively. $\sim$ denotes no significant difference.

Rules	ALL		M1
Supergames		31–50	
Type	LR		SIM
First periods	0.76 (0.05)	$\overset{1.0}{\sim}$	0.76 (0.05)
First 4 periods	0.68	$\overset{0.88}{\sim}$	0.69
	$\overset{0.0}{\ggg}$		$\overset{0.0}{\ggg}$
Last 4 periods	0.56 (0.03)	$\overset{0.05}{<}$	0.66 (0.04)
All periods	0.6 (0.03)	$\overset{0.2}{\sim}$	0.67 (0.04)

Fig. 11 shows that approximately 72% of subjects' rule sets are in four clusters, with exemplars that match TFT, GRIM, ALLD, or D.TFT exactly.[23] This suggests that even though subjects are not necessarily playing one of the 20 commonly studied strategies exactly, they are playing strategies that are similar to one of these 20 strategies. As a check, we compare the results from the cluster analysis to those from the standard maximum likelihood strategy estimation procedure using Table 3, and we find that the estimates match closely. Furthermore, the four largest clusters all correspond to memory-1 strategies. Thus we find support for the majority of subjects relying on strategies that are close to memory-1 strategies commonly found in the shorter games.[24]

Result 2 suggests that subjects use rules longer than memory-1. *Result 3* suggests that most of the strategies are similar to memory-1 strategies. These two results seem to contradict each other. However, even though subjects' strategies may be similar to the commonly studied memory-1 strategies, they still may have longer rules that play an important role in the resulting levels of cooperation. We further explore the impact that longer rules may have on these memory-1 strategies using simulations in Section 4 and connect it to *Result 1*.

4. Simulation analysis

In this section, we conduct two simulation exercises in order to better understand the implications of the memory-2+ rules on cooperation rates and the MLE estimates obtained in the experiment.[25] Specifically, we use the first set of simulations to compare levels of cooperation when agents are restricted to only memory-1 rules versus the unrestricted case. For the unrestricted agents, we consider the behavior from the locked-response stage of the experiment (supergame 31–50 in Table 2). As a proxy for the restricted agents, we remove all memory 2+ rules from the subjects' rule sets and then run the simulation using the exact same matchings and supergame lengths as in the locked-response stage of the experiment. The resulting levels of cooperation are displayed in Table 4.

In accordance with *Result 1*, Table 4 shows that cooperation breaks down from the beginning of the supergame to the end of the supergame (0.68 to 0.56) when the full rule sets are used. However, when the longer rules are removed from the subjects' strategies, cooperation rates are sustained from the beginning to the end of the supergame (0.69 to 0.66).[26] The difference between the drop in the unrestricted case (0.12) and the drop in the restricted case (0.03) is highly significant (p-value of 0.00). This provides some evidence that the presence of longer rules causes cooperation to break down more than it would when subjects are restricted to only memory 1 rules.

[23] Note that there is a cluster in the first supergame of the locked-response stage with two subjects, which has an exemplar that is closest to TFT, but the exemplar does not match TFT exactly.

[24] Dal Bó and Fréchette (2016) review results of five papers that run indefinitely repeated prisoner's dilemma games over a wide range of parameters. They report maximum likelihood estimates and elicited strategies for a total of 17 different treatments, and find that TFT, GRIM, ALLD, D.TFT and ALLC account for the majority of behavior in 16 of them. In the treatment with the same parameters as ours, $\delta = 0.95$ and the payoffs from Fig. 1, they report elicited strategies of 25% TFT, 6% GRIM, 22% ALLD, and 0% D.TFT.

[25] Appendix I contains the results of an additional simulation exercise which evaluates strategy performance.

[26] Even though the difference between the first 4 and last 4 in M1 (0.69 to 0.66) is small, it is still significant because the cooperation rate increases for only 3 out of 82 subjects but decreases by a small amount for 47 out of the 82 subjects. The drop in cooperation observed when subjects are restricted to only memory 1 rules can be attributed to strategies such as GRIM.

Table 5
Maximum likelihood estimates for simulated data. *Notes*: Estimates based on the first 20 periods of simulations and all 20 supergames. If the number of periods in a given supergame is less than 20, then all periods were used. Bootstrapped standard errors are shown in parentheses. Cooperation rates are reported for the first 20 periods of interaction. Strategies that are not above 0.04 for any α are not included.

α	TFT	GRIM	D.TFT	ALLD	2TFT	GRIM2	ALLC	TF2T	CToD	β	Cooperation	# of Subjects
0.00	0.28*** (0.12)	0.24*** (0.09)	0.12*** (0.05)	0.12*** (0.04)	0.04** (0.02)	0.07*** (0.03)	0.06** (0.03)	0.05*** (0.02)		0.94	0.52	82
0.25	0.28*** (0.10)	0.19*** (0.08)	0.12** (0.05)	0.12*** (0.05)	0.11** (0.05)	0.05** (0.02)	0.06** (0.03)	0.05** (0.02)	0.01** (0.01)	0.93	0.49	82
0.50	0.28 (0.43)	0.18 (0.51)	0.12** (0.06)	0.12*** (0.05)	0.10** (0.04)	0.07** (0.03)	0.06 (0.18)	0.05*** (0.02)		0.91	0.42	82
0.75	0.28*** (0.11)	0.16*** (0.06)	0.12*** (0.04)	0.12** (0.05)	0.09** (0.04)	0.07*** (0.03)	0.06** (0.03)	0.05** (0.02)	0.04*** (0.01)	0.91	0.40	82
1.00	0.28*** (0.11)	0.19*** (0.07)	0.12** (0.05)	0.12*** (0.05)	0.08** (0.03)	0.07*** (0.03)	0.06** (0.03)	0.05** (0.02)	0.01* (0.01)	0.91	0.38	82

We use the second set of simulations to focus on a particular type of longer rules — *CsToD* rules. Specifically, we investigate the impact of *CsToD* rules on the levels of cooperation and MLE estimates. We simulate a population of 82 agents that are randomly matched for twenty supergames with continuation probability $\delta = 0.95$. Each agent is assigned a rule set that specifies one of the nine strategies obtained from the cluster analysis in Fig. 11. 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 the number of subjects following each strategy is equal to the number of subjects in a corresponding cluster. Thus, to match the strategy frequencies from the clustering in Table 3, we have 23 TFT agents, 16 GRIM agents, 10 ALLD agents, etc. In what follows, we investigate what happens to the cooperation rates and strategy estimates when a fraction α of agents with each strategy have a memory-3 *CsToD* rule in their set. Table 5 presents the results for $\alpha \in \{0, 0.25, .5, 0.75, 1\}$.

There are several points worth noting. First, as α increases, the overall cooperation decreases from 0.52 to 0.38. This result corroborates our discussion in Section 3.2. Second, MLE estimates are unaffected by the increase in α. Third, the β parameter of the MLE, which captures the amount of noise, decreases from 0.94 to 0.91. Thus, we find that regularly observed aspects of strategies — specifically *CsToD* rules — do not have an affect on the strategy estimates and are perceived as noise in the estimation procedure; nevertheless, these rules are vital to the cooperation dynamics within supergames.

To summarize, we find that the MLE does an excellent job at uncovering key aspects of subject's strategies (as evidenced by the estimated frequencies that match the frequencies of strategies used for the simulations), but may miss other aspects of subject's strategies (such as the *CsToD* rules). These other aspects of subject's strategies are an important factor in the level of cooperation, especially in long repeated games. Developing new techniques for better uncovering more complex aspects of subject's strategies is an interesting avenue for future research.

5. Conclusion

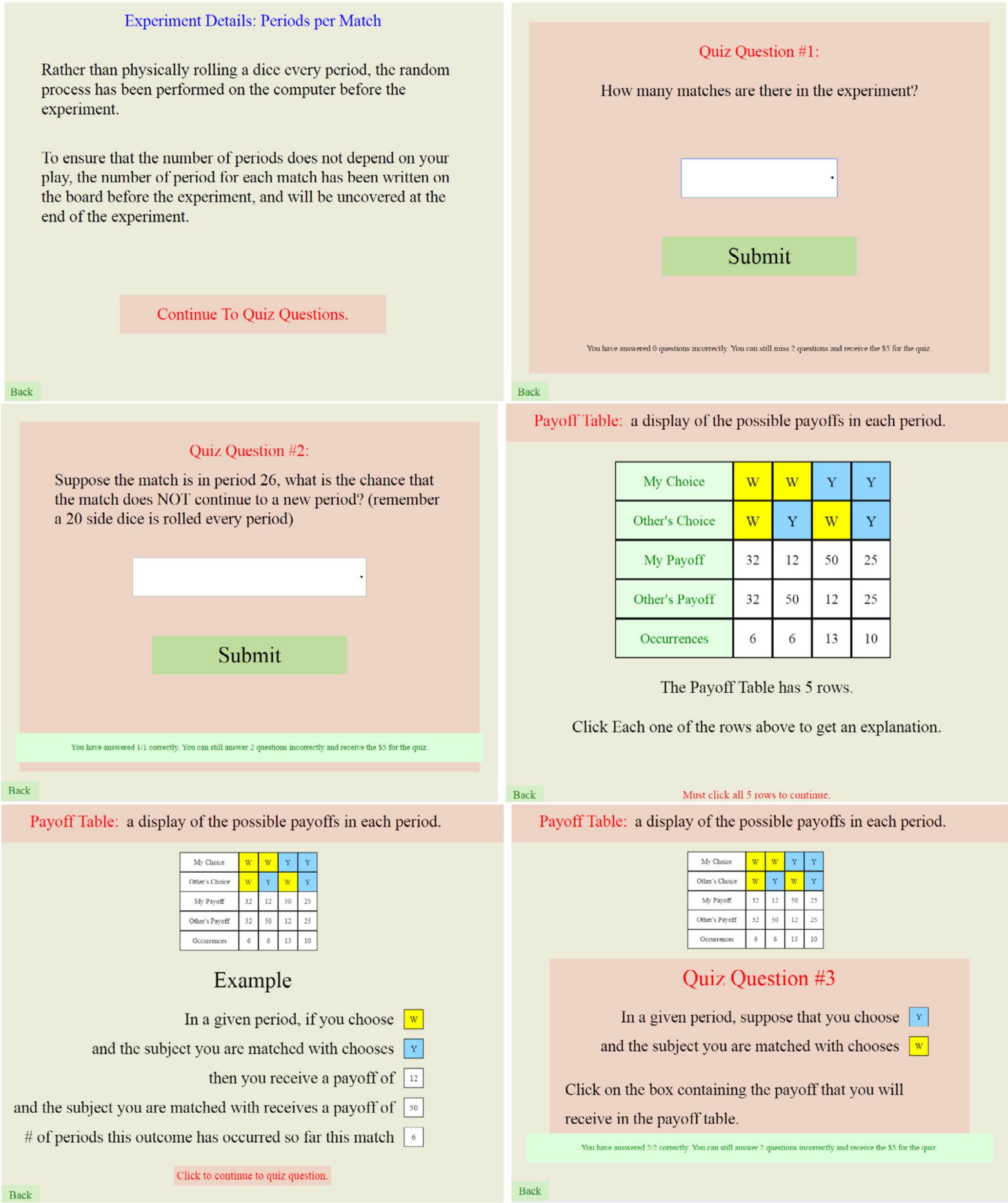
The contribution of this paper is twofold. First, we develop the interface that allows us to run experiments on long repeated games. The interface implements the strategy method, allowing us to gain a unique insight into the strategies that participants develop. Second, we conduct experiments with our interface. In particular, we study the indefinitely repeated prisoner's dilemma with a continuation probability of $\delta = .95$. We find cooperation rates consistent with those found in prior studies, and we find that our interface does not affect how subjects behave.

Our experiments yield several results. First, we find that cooperation rates decrease as the supergame progresses. Second, 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*, that plays an important role in determining the levels of cooperation within supergames. We then analyze the fully specified strategies and find that, while roughly 40% of the strategies are exact matches of those commonly used in the literature, about 75% of the strategies are close to TFT, GRIM, ALLD, and D.TFT when we perform a cluster analysis. In addition, we perform the standard maximum likelihood estimation based on only the observed actions, and we find that results are similar to the cluster analysis results.

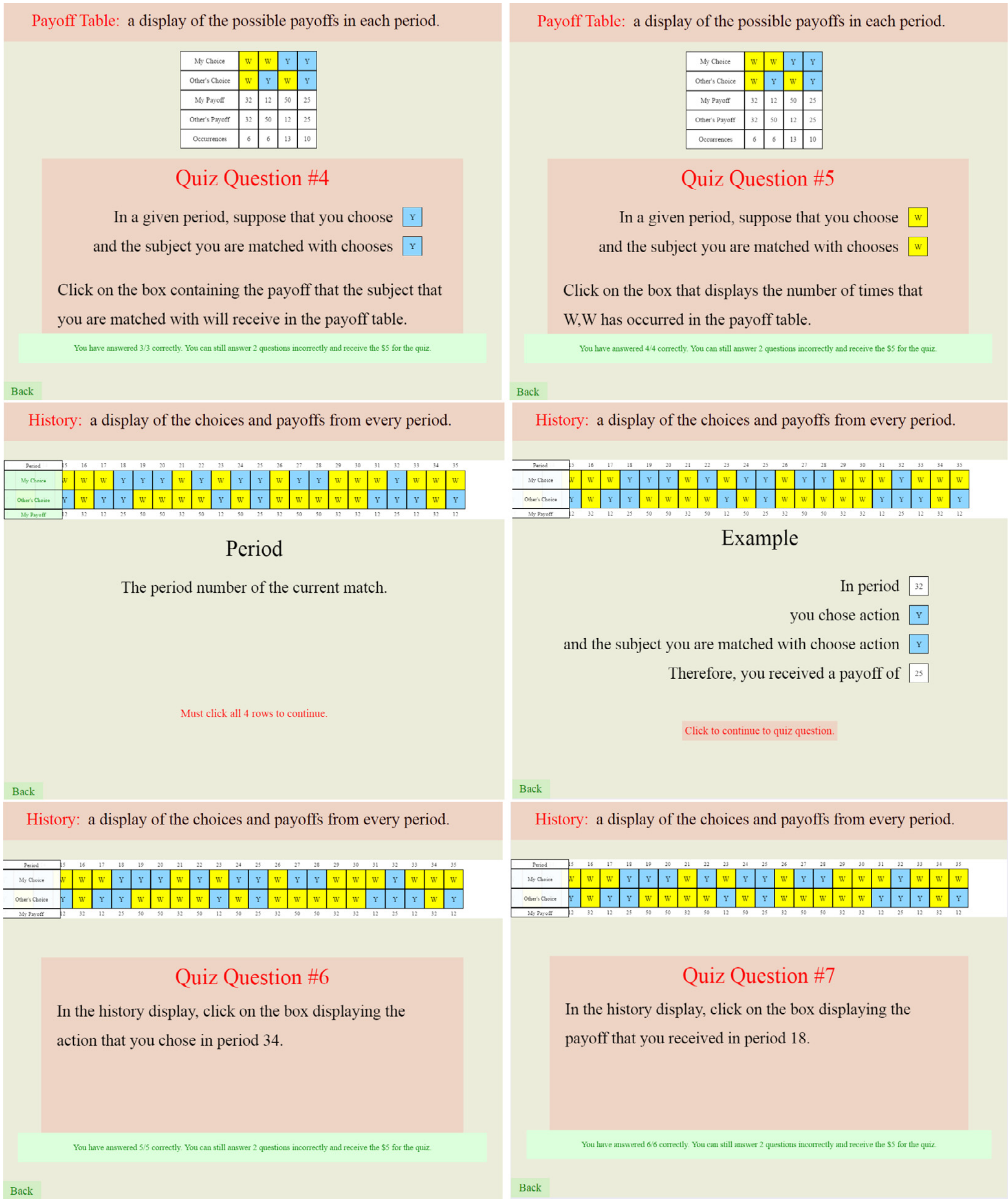
Combined with the main results, the simulations presented in Section 4 suggest that the standard maximum likelihood procedure does an excellent job of uncovering strategies in fairly complex environments. However, they also suggest that the strategy estimates may miss some important aspects of play, 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 a large impact on the level of cooperation.

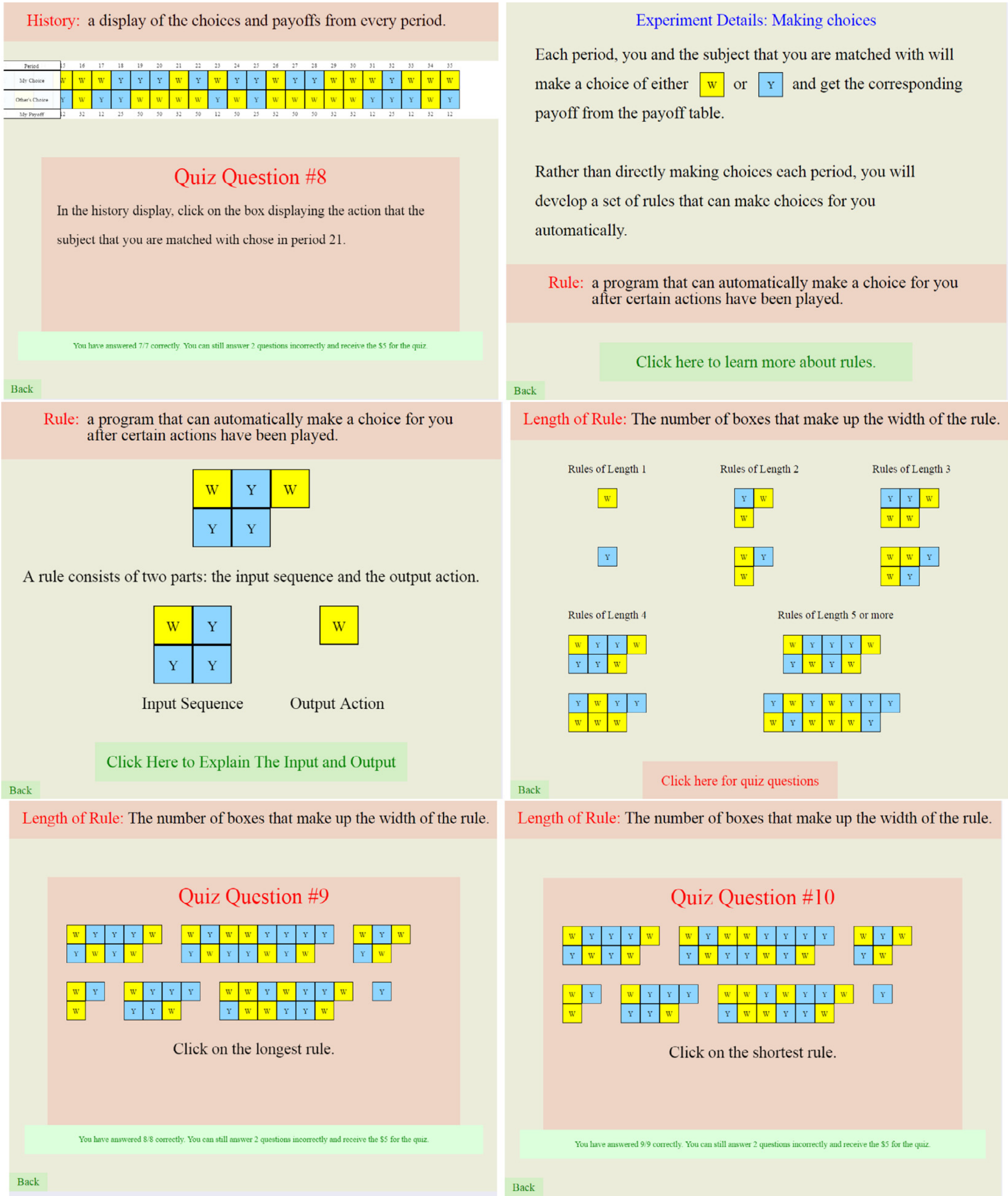
There are many interesting avenues for future research. First, it would be interesting to gain a better understanding of what causes these long rules to be played. For example, are they used only in long repeated games, like the one studied here, or would they also be present in shorter games (such as $\delta = 0.9$)? In addition, it would be useful to try to better understand how subjects learn. We find that cooperation rates increase after the locked-response stage in treatment 1, which suggests that subjects may still be learning in later supergames. Alternative designs that build on our work by expanding the use of the locked-response stage could help investigate this problem. 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 does not change? The proposed experimental interface will be a useful tool in these further investigations.

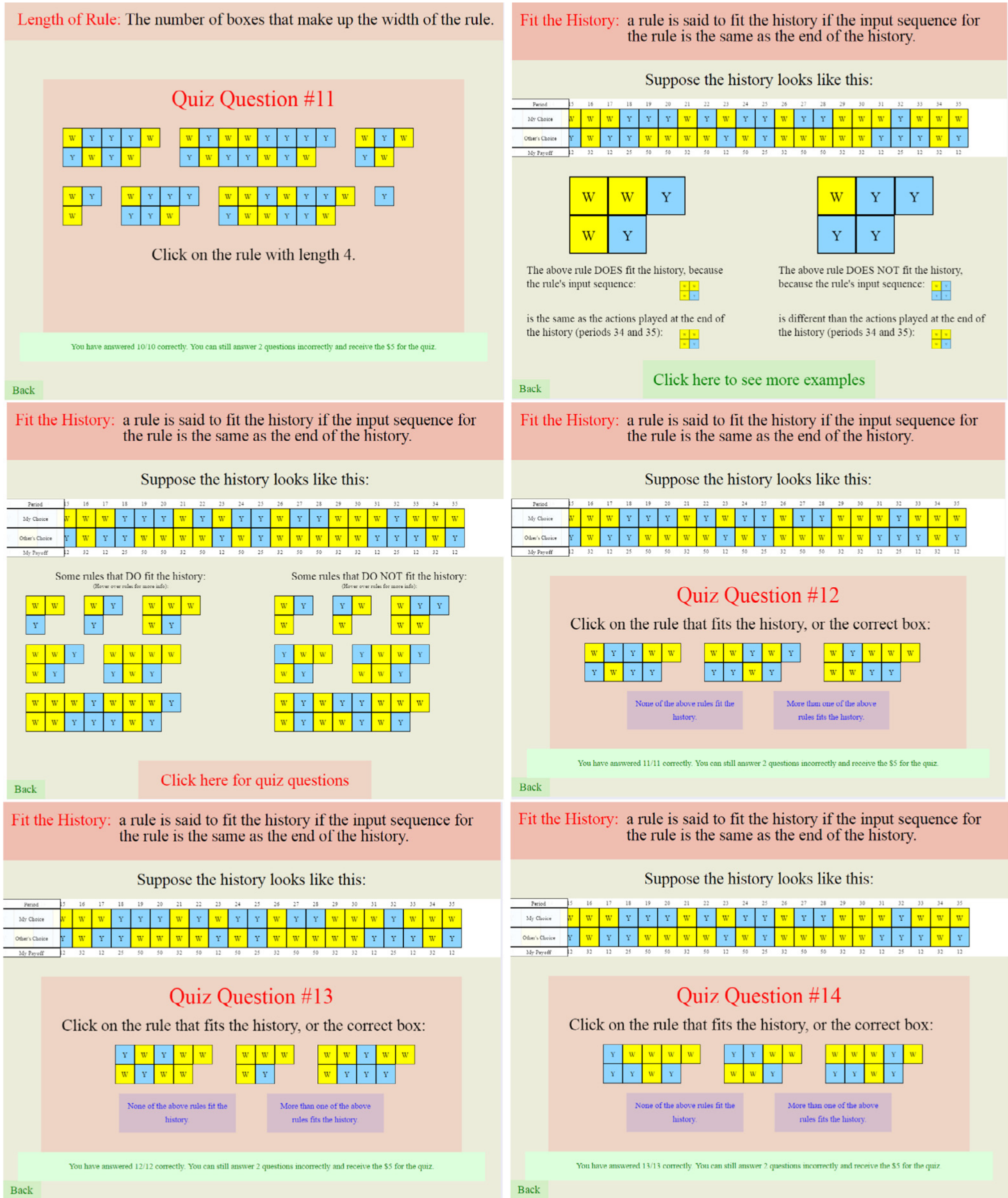
Appendix A. Experimental instructions and quiz



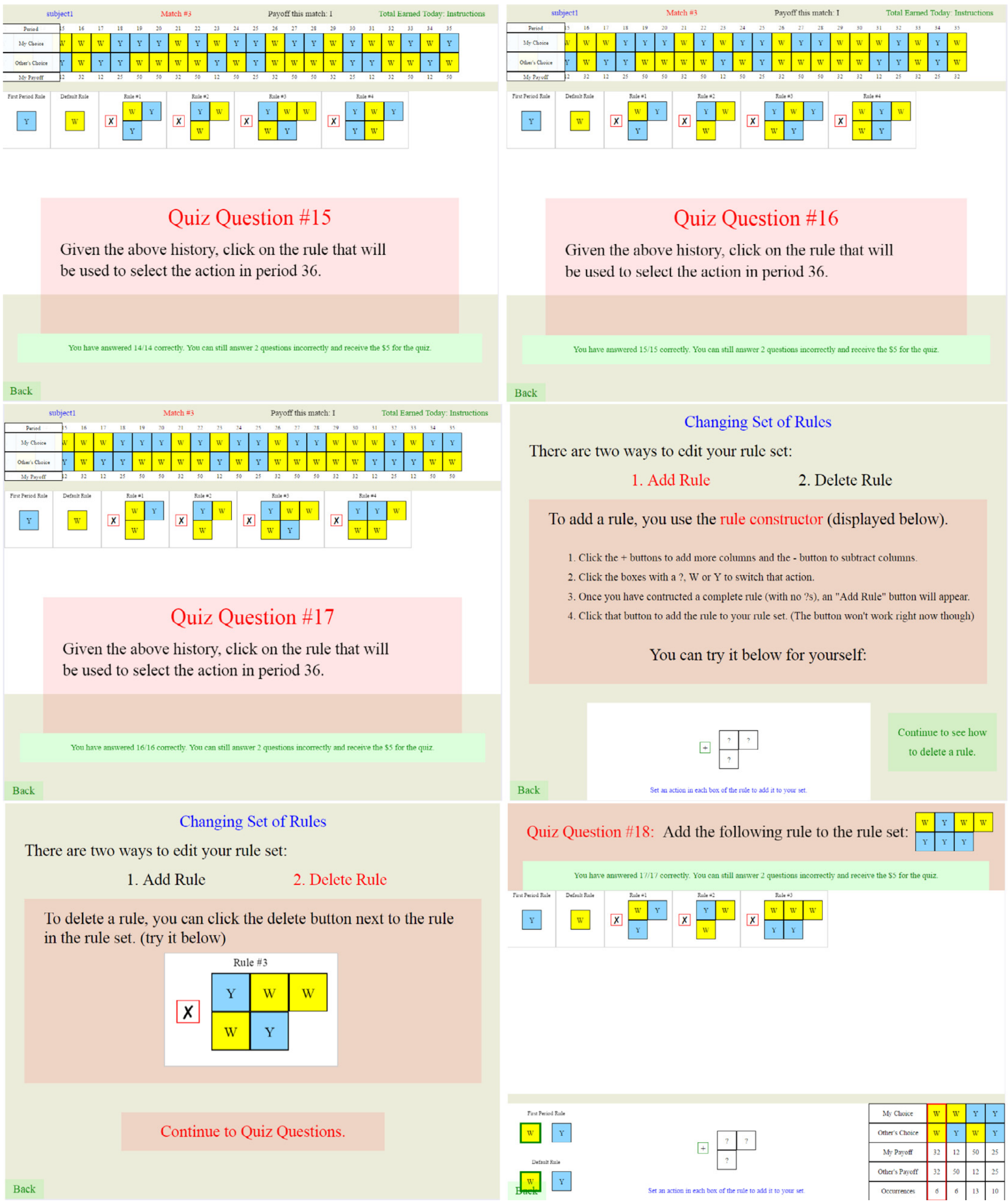
Experiment Details: Periods per Match
Rather than physically rolling a dice every period, the random process has been performed on the computer before the experiment.
To ensure that the number of periods does not depend on your play, the number of period for each match has been written on the board before the experiment, and will be uncovered at the end of the experiment.
Continue To Quiz Questions.
Back
Quiz Question #1:
How many matches are there in the experiment?
Submit
You have answered 0 questions incorrectly. You can still miss 2 questions and receive the $5 for the quiz.
Back
Quiz Question #2:
Suppose the match is in period 26, what is the chance that the match does NOT continue to a new period? (remember a 20 side dice is rolled every period)
Submit
You have answered 1/1 correctly. You can still answer 2 questions incorrectly and receive the $5 for the quiz.
Back
Payoff Table: a display of the possible payoffs in each period.
My Choice W W Y Y
Other's Choice W Y W Y
My Payoff 32 12 50 25
Other's Payoff 32 50 12 25
Occurrences 6 6 13 10
The Payoff Table has 5 rows.
Click Each one of the rows above to get an explanation.
Back
Must click all 5 rows to continue.
Payoff Table: a display of the possible payoffs in each period.
Example
In a given period, if you choose W
and the subject you are matched with chooses Y
then you receive a payoff of 12
and the subject you are matched with receives a payoff of 50
# of periods this outcome has occurred so far this match 6
Click to continue to quiz question.
Back
Payoff Table: a display of the possible payoffs in each period.
Quiz Question #3
In a given period, suppose that you choose Y
and the subject you are matched with chooses W
Click on the box containing the payoff that you will receive in the payoff table.
You have answered 2/2 correctly. You can still answer 2 questions incorrectly and receive the $5 for the quiz.
Back

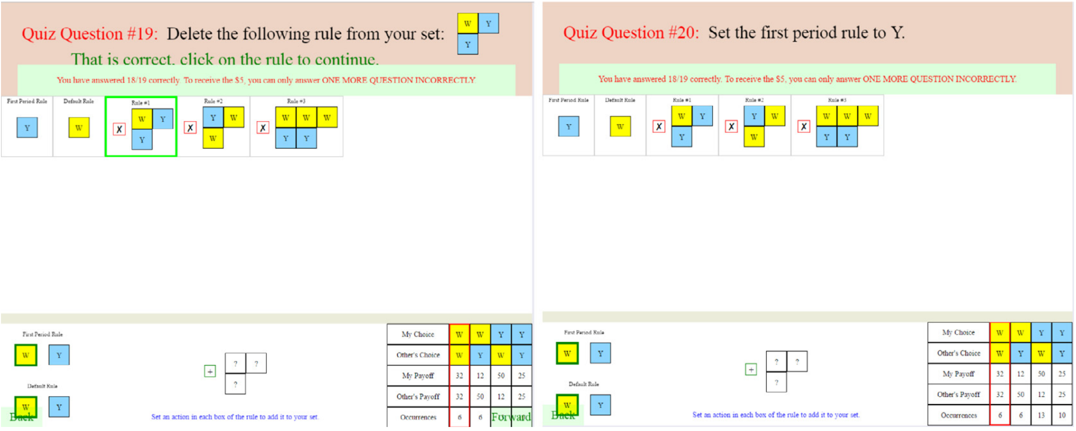

Payoff Table: a display of the possible payoffs in each period.
My Choice
Other's Choice
My Payoff
Other's Payoff
Occurrences
Quiz Question #4
In a given period, suppose that you choose Y
and the subject you are matched with chooses Y
Click on the box containing the payoff that the subject that you are matched with will receive in the payoff table.
You have answered 3/3 correctly. You can still answer 2 questions incorrectly and receive the $5 for the quiz.
Back
Payoff Table: a display of the possible payoffs in each period.
Quiz Question #5
In a given period, suppose that you choose W
and the subject you are matched with chooses W
Click on the box that displays the number of times that W,W has occurred in the payoff table.
You have answered 4/4 correctly. You can still answer 2 questions incorrectly and receive the $5 for the quiz.
Back
History: a display of the choices and payoffs from every period.
Period
The period number of the current match.
Must click all 4 rows to continue.
Back
History: a display of the choices and payoffs from every period.
Example
In period 32
you chose action Y
and the subject you are matched with choose action Y
Therefore, you received a payoff of 25
Click to continue to quiz question.
Back
History: a display of the choices and payoffs from every period.
Quiz Question #6
In the history display, click on the box displaying the action that you chose in period 34.
You have answered 5/5 correctly. You can still answer 2 questions incorrectly and receive the $5 for the quiz.
Back
History: a display of the choices and payoffs from every period.
Quiz Question #7
In the history display, click on the box displaying the payoff that you received in period 18.
You have answered 6/6 correctly. You can still answer 2 questions incorrectly and receive the $5 for the quiz.
Back


History: a display of the choices and payoffs from every period.
Quiz Question #8
In the history display, click on the box displaying the action that the subject that you are matched with chose in period 21.
You have answered 7/7 correctly. You can still answer 2 questions incorrectly and receive the $5 for the quiz.
Back
Experiment Details: Making choices
Each period, you and the subject that you are matched with will make a choice of either W or Y and get the corresponding payoff from the payoff table.
Rather than directly making choices each period, you will develop a set of rules that can make choices for you automatically.
Rule: a program that can automatically make a choice for you after certain actions have been played.
Click here to learn more about rules.
Back
Rule: a program that can automatically make a choice for you after certain actions have been played.
A rule consists of two parts: the input sequence and the output action.
Input Sequence
Output Action
Click Here to Explain The Input and Output
Back
Length of Rule: The number of boxes that make up the width of the rule.
Rules of Length 1
Rules of Length 2
Rules of Length 3
Rules of Length 4
Rules of Length 5 or more
Click here for quiz questions
Back
Length of Rule: The number of boxes that make up the width of the rule.
Quiz Question #9
Click on the longest rule.
You have answered 8/8 correctly. You can still answer 2 questions incorrectly and receive the $5 for the quiz.
Back
Length of Rule: The number of boxes that make up the width of the rule.
Quiz Question #10
Click on the shortest rule.
You have answered 9/9 correctly. You can still answer 2 questions incorrectly and receive the $5 for the quiz.
Back


Length of Rule: The number of boxes that make up the width of the rule.
Quiz Question #11
Click on the rule with length 4.
You have answered 10/10 correctly. You can still answer 2 questions incorrectly and receive the $5 for the quiz.
Back
Fit the History: a rule is said to fit the history if the input sequence for the rule is the same as the end of the history.
Suppose the history looks like this:
The above rule DOES fit the history, because the rule's input sequence:
is the same as the actions played at the end of the history (periods 34 and 35):
The above rule DOES NOT fit the history, because the rule's input sequence:
is different than the actions played at the end of the history (periods 34 and 35):
Click here to see more examples
Back
Fit the History: a rule is said to fit the history if the input sequence for the rule is the same as the end of the history.
Suppose the history looks like this:
Some rules that DO fit the history:
Some rules that DO NOT fit the history:
Click here for quiz questions
Back
Fit the History: a rule is said to fit the history if the input sequence for the rule is the same as the end of the history.
Suppose the history looks like this:
Quiz Question #12
Click on the rule that fits the history, or the correct box:
None of the above rules fit the history.
More than one of the above rules fits the history.
You have answered 11/11 correctly. You can still answer 2 questions incorrectly and receive the $5 for the quiz.
Back
Fit the History: a rule is said to fit the history if the input sequence for the rule is the same as the end of the history.
Suppose the history looks like this:
Quiz Question #13
Click on the rule that fits the history, or the correct box:
None of the above rules fit the history.
More than one of the above rules fits the history.
You have answered 12/12 correctly. You can still answer 2 questions incorrectly and receive the $5 for the quiz.
Back
Fit the History: a rule is said to fit the history if the input sequence for the rule is the same as the end of the history.
Suppose the history looks like this:
Quiz Question #14
Click on the rule that fits the history, or the correct box:
None of the above rules fit the history.
More than one of the above rules fits the history.
You have answered 13/13 correctly. You can still answer 2 questions incorrectly and receive the $5 for the quiz.
Back


Rules of Length 1
There are two types of rules with a length of 1:
First Period Rule:
The first period rule only specifies the action to be played in the first period of each match.
The first period rule is selected by clicking either the W box or the Y box in the panel labeled "First Period Rule" that looks like this:
Default Rule:
The default rule has a length of 1, and therefore it always fits the history.
The default rule is selected by clicking either the W box or the Y box in the panel labeled "Default Rule" that looks like this:
You will always have a first period rule and a default rule in your rule set during the experiment. This ensures that in any period, your rule set will have at least one rule that fits the history.
Back
Continue.
Which Rule Will Be Selected?
Each period your set of rules will select a single action to be played in the next period.
Since the default rule always fits the history, and you will always have a default rule in your set, there will always be at least one rule that fits the history.
If there are multiple rules that fit the history, then the longest rule that fits the history will be used to select the action in the next period.
Continue to see some examples.
Back
In the above example, there are three rules that fit the history: the Default Rule, Rule #1 and Rule #3. Since Rule #3 is the longest rule that fits the history, it will be used to select the action W in period 36.
Back
Continue.
In the above example, there are three rules that fit the history: the Default Rule, Rule #2 and Rule #6. Since Rule #6 is the longest rule that fits the history, it will be used to select the action Y in period 10.
Back
Continue.
In the above example, there is only one rule that fit the history: the Default Rule. Since the Default Rule is the longest rule that fits the history, it will be used to select the action W in period 18.
Back
Continue.
In the above example, the current period is period 1 of the match so there is no history. Since it is the first period of the match, the First Period Rule will be used to select the action Y in period 1.
Back
Click to Continue to Quiz.


subject1
Match #3
Payoff this match: 1
Total Earned Today: Instructions
First Period Rule
Default Rule
Rule #1
Rule #2
Rule #3
Rule #4
Quiz Question #15
Given the above history, click on the rule that will be used to select the action in period 36.
You have answered 14/14 correctly. You can still answer 2 questions incorrectly and receive the $5 for the quiz.
Back
Quiz Question #16
Given the above history, click on the rule that will be used to select the action in period 36.
You have answered 15/15 correctly. You can still answer 2 questions incorrectly and receive the $5 for the quiz.
Back
Quiz Question #17
Given the above history, click on the rule that will be used to select the action in period 36.
You have answered 16/16 correctly. You can still answer 2 questions incorrectly and receive the $5 for the quiz.
Back
Changing Set of Rules
There are two ways to edit your rule set:
1. Add Rule
2. Delete Rule
To add a rule, you use the rule constructor (displayed below).
1. Click the + buttons to add more columns and the - button to subtract columns.
2. Click the boxes with a ?, W or Y to switch that action.
3. Once you have contructed a complete rule (with no ?s), an "Add Rule" button will appear.
4. Click that button to add the rule to your rule set. (The button won't work right now though)
You can try it below for yourself:
Continue to see how to delete a rule.
Set an action in each box of the rule to add it to your set.
Back
Changing Set of Rules
There are two ways to edit your rule set:
1. Add Rule
2. Delete Rule
To delete a rule, you can click the delete button next to the rule in the rule set. (try it below)
Rule #3
Continue to Quiz Questions.
Back
Quiz Question #18: Add the following rule to the rule set:
You have answered 17/17 correctly. You can still answer 2 questions incorrectly and receive the $5 for the quiz.
My Choice
Other's Choice
My Payoff
Other's Payoff
Occurrences
Set an action in each box of the rule to add it to your set.



Appendix B. Additional quiz details

Table B1

Table B1
Quiz performance.

Question order	Question type	HQ correct	LQ correct	Difference
1	Matches and periods	71/82	22/34	0.22
2	Matches and periods	75/82	21/34	0.3
3	Actions and payoffs	82/82	30/34	0.12
4	Actions and payoffs	79/82	32/34	0.02
5	Actions and payoffs	82/82	34/34	0.0
6	Actions and payoffs	82/82	34/34	0.0
7	Actions and payoffs	82/82	31/34	0.09
8	Actions and payoffs	82/82	34/34	0.0
9	Rule length	81/82	32/34	0.05
10	Rule length	79/82	32/34	0.02
11	Rule length	76/82	27/34	0.13
12	Rule fit	79/82	28/34	0.14
13	Rule fit	74/82	26/34	0.14
14	Rule fit	81/82	30/34	0.11
15	Identify rule	75/82	18/34	**0.39**
16	Identify rule	75/82	18/34	**0.39**
17	Identify rule	80/82	24/33	0.25
18	Add rule	76/82	15/29	**0.41**
19	Delete rule	80/80	14/19	0.26
20	First period rule	77/77	17/17	0.0

Notes: There were twenty questions in total. Subjects were provided an explanation of the correct answer if they incorrectly answered a question.

Appendix C. Additional interface screenshots

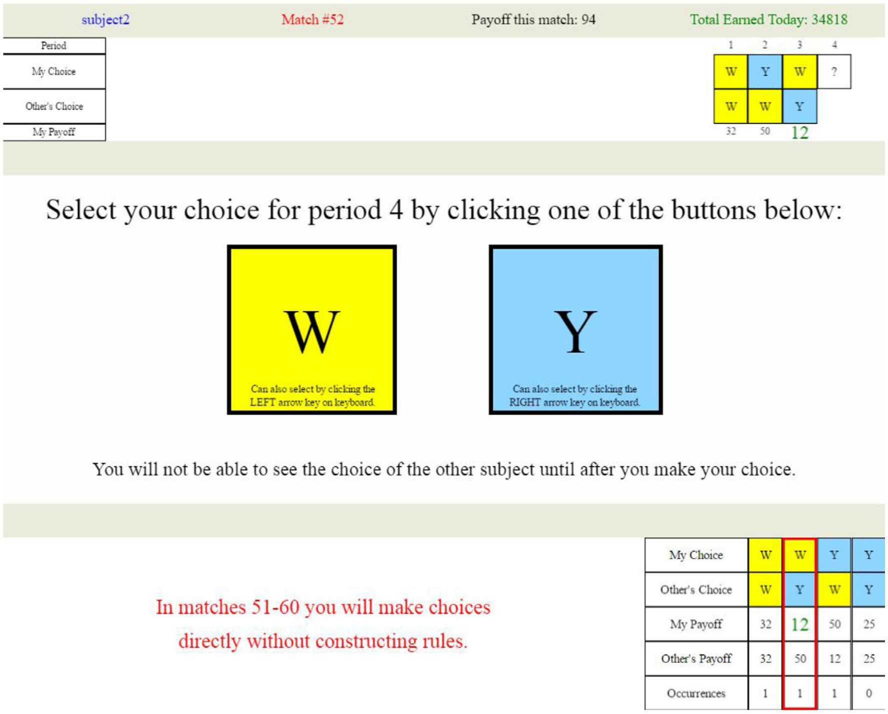


Fig. C1. Direct-response stage screenshot.

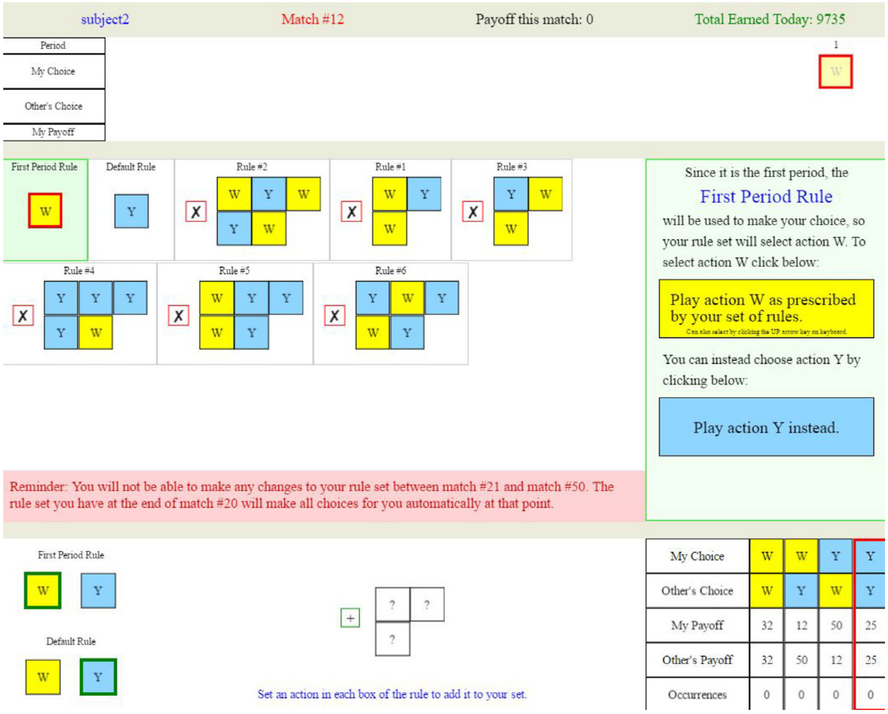


Fig. C2. Non-binding stage screenshot.

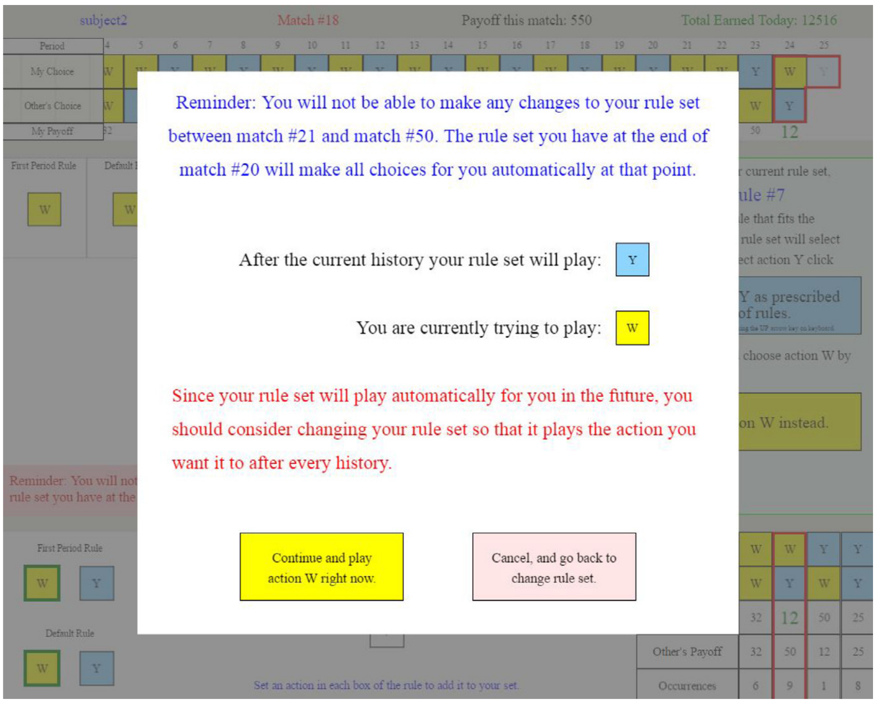


Fig. C3. Alert message screenshot. *Notes*: Alert message appears in the non-binding stage when subjects click to play an action different from the action prescribed by their rule set.

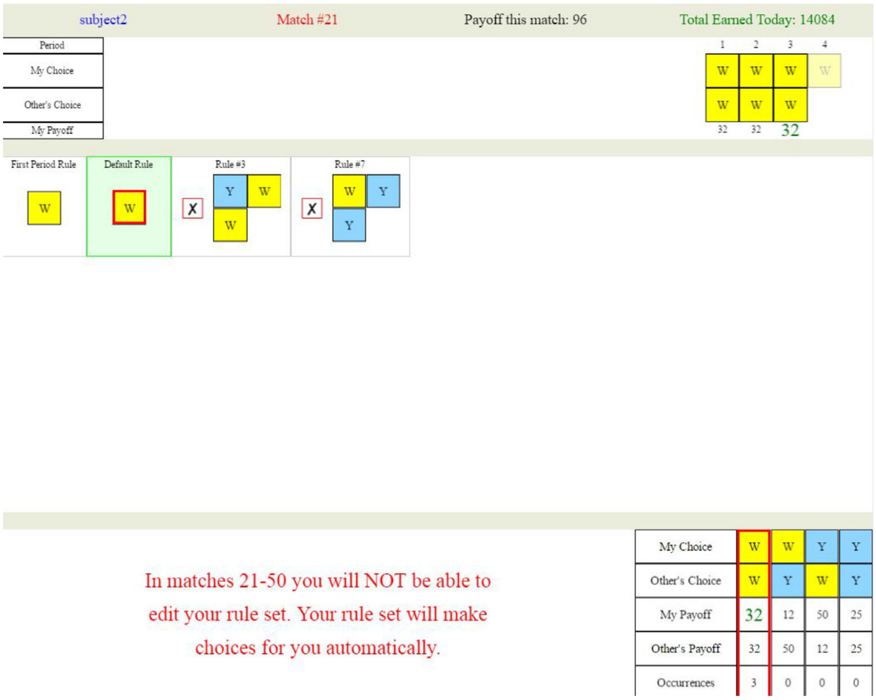


Fig. C4. Locked-responsestage screenshot.

Table D2
Supergame length realizations.

Supergame number:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
Realization #1:	23	31	20	26	10	1	14	5	84	4	99	26	15	7	31	13	3	25	18	2	13	15	41	12	13	3	1	9	69	45
Realization #2:	33	39	4	13	29	15	13	5	19	4	33	14	19	9	1	10	48	21	9	1	16	9	11	2	4	5	9	8	22	10
Realization #3:	10	55	15	1	48	49	74	13	4	33	11	28	4	13	7	1	24	31	27	12	17	3	4	8	9	31	33	6	23	22

Supergame number:	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
Realization #1:	10	19	5	17	13	37	4	18	3	24	14	5	10	57	9	55	46	10	29	15	170	21	35	13	7	8	8	14	1	13
Realization #2:	32	23	2	28	5	19	5	20	1	31	8	5	68	15	12	4	13	13	37	37	5	18	40	20	25	2	24	47	41	25
Realization #3:	5	14	97	3	8	3	65	16	20	5	35	41	35	8	7	52	23	18	29	7	1	48	20	28	16	7	2	12	1	8

Table E3
Additional treatment comparisons.

(a) Average cooperation rate				(b) Average earnings per period			
Treatment	1		2	Treatment	1		2
Supergames		11–20		Supergames		11–20	
Type	NB		DR	Type	NB		DR
First periods	0.75 (0.06)	$\overset{0.35}{\sim}$	0.67 (0.07)	First periods	31.17 (1.09)	$\overset{0.72}{\sim}$	30.59 (1.14)
First 4 periods	0.67 (0.04)	$\overset{0.24}{\sim}$	(0.71) (0.05)	First 4 periods	30.38 (0.34)	$\overset{0.72}{\sim}$	30.58 (0.44)
	$\overset{0.0}{\ggg}$		$\overset{0.0}{\ggg}$		$\overset{0.01}{\gg}$		$\overset{0.04}{\gg}$
Last 4 periods	0.58 (0.04)	$\overset{0.34}{\sim}$	0.63 (0.05)	Last 4 periods	29.48 (0.34)	$\overset{0.6}{\sim}$	29.74 (0.33)
All periods	0.61 (0.04)	$\overset{0.24}{\sim}$	0.66 (0.05)	All periods	29.81 (0.32)	$\overset{0.5}{\sim}$	30.1 (0.27)

Notes: **(a)** The unit of observation is a subject. The cooperation rate is determined by averaging the fraction of periods that each subject cooperated within a supergame across the ten supergames. **(b)** The unit of observation is the subjects' average per period earnings in a supergame, averaged over the ten supergames. Bootstrapped standard errors are in parentheses. If the supergame is less than four periods, then the cooperation rate for the first four and the last four is set to the cooperation rate for all periods. Tests between treatments are carried out using Wilcoxon ranked-sum test. Tests within treatments are carried out using Wilcoxon signed-rank test. *p*-values are provided for each test. $<$, $\ll$, $\lll$ denote significant difference at the 0.10, 0.05, and 0.01 levels, respectively.

Appendix D. Supergame length details

Appendix E. Additional comparisons

In this section, we further test the differences between the two treatments by running additional robustness checks. We focus on the differences for supergames 11–20 because 1) the two treatments are identical before supergame 11; 2) subjects have the same amount of experience playing the game through supergame 20; 3) the two treatments use different interface starting in supergame 11. Table E3 runs additional robustness checks using Wilcoxon rank-sum and Wilcoxon signed-rank tests to compare cooperation and payoffs between the two treatments. The results in this table provide additional evidence that there is no difference in cooperation or payoffs between the two treatments. Table E4 provides average cooperation rates after each of the four possible memory-1 histories. We find no difference for CC, DD, and DC histories, but a significant difference in average cooperation rate after CD history.

Table E4
Average cooperation rate after each memory-1 history. *Notes*: The unit of observation is a subject. The average cooperation rate after each memory-1 history is the fraction of periods that a subject cooperated, averaged over the 10 supergames. Bootstrapped standard errors are in parentheses. Tests between treatments are carried out using Wilcoxon ranked-sum test. *p*-values are provided for each test. $<$, $\ll$, $\lll$ denote significant difference at the 0.10, 0.05, and 0.01 levels, respectively.

Average cooperation rate after each memory-1 history			
Treatment	1		2
Supergames		11–20	
Type	NB		DR
After CC	0.88 (0.03)	$\overset{0.84}{\sim}$	0.85 (0.05)
After CD	0.28 (0.05)	$\overset{0.01}{\ll}$	0.47 (0.05)
After DC	0.54 (0.06)	$\overset{0.89}{\sim}$	0.57 (0.05)
After DD	0.16 (0.03)	$\overset{0.57}{\sim}$	0.18 (0.03)

Appendix F. Clustering robustness

To investigate how many sequences would be sufficient, we run the following robustness check:

1. For a given value of N, repeat the following 50 times:
   (a) Generate an input vector consisting of N sequences generated by the Markov process with expected length of 20 (with length of each sequence generated using the geometric distribution corresponding to $\delta = .95$).
   (b) Generate an output vector for each of the 82 rule sets and 20 commonly studied strategies.
   (c) Cluster the rule sets based on these output vectors.
   (d) Assign a strategy for each rule set based on the strategy that is closest to the exemplar of the cluster that contains that rule set.
2. For each of the 82 rule sets, determine the number of different strategies that it has been assigned to over the 50 iterations. Then take the average of these.

As N gets large, the differentiation becomes better, and the strategy should be in the same cluster for all 50 realizations. We investigate this as we vary N and present our results in Fig. F5. We find that even for a relatively small number of sequences ($N = 100$), subjects are classified to 1.4 clusters (on average). As we increase N that number gets closer to 1. For the analysis

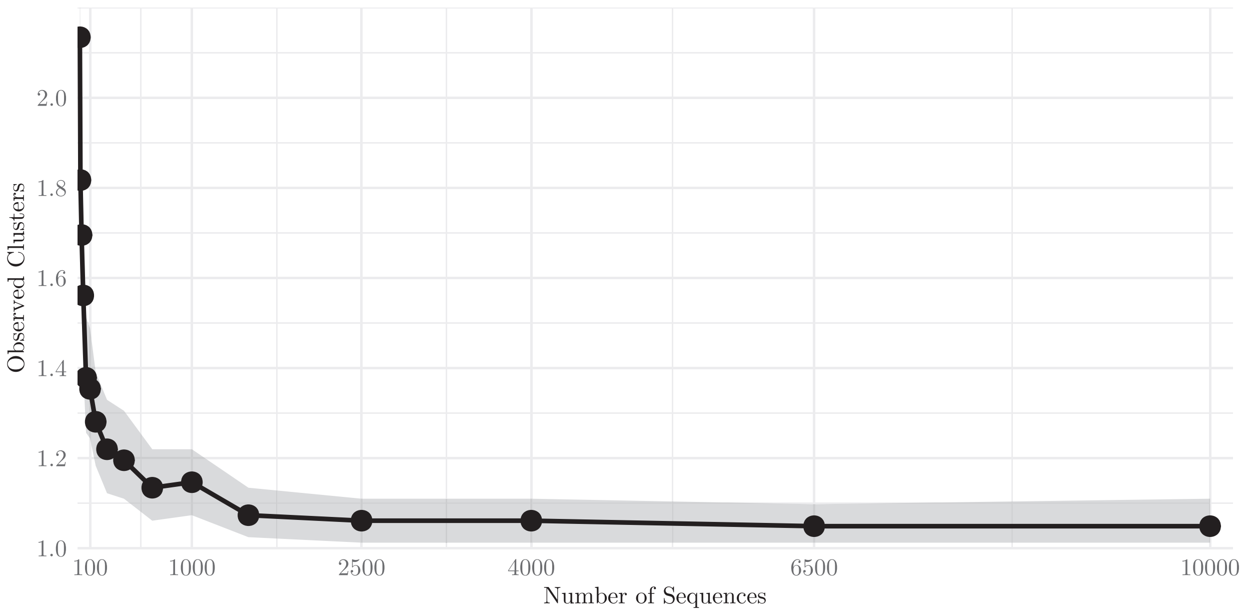


Fig. F5. Average number of clusters for different numbers of input sequences.

Table G5
Maximum likelihood estimates. *Notes*: Estimates use 20 initial periods of the given range of 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.

Treatment	Supergames	Type	TFT	GRIM	D.TFT	ALLD	2TFT	GRIM2	TF2T	WSLS	ALLC	CToD	GRIM3	D.TF2T	D.TF3T	2TF2T	WSLS2	D.GRIM3	β	Cooperation	# of Subjects
1	16-20	NB	0.30^{***} (0.08)	0.10^{*} (0.06)	0.14^{***} (0.05)	0.11^{***} (0.05)	0.10^{*} (0.06)	0.05 (0.05)	0.10^{*} (0.06)	0.03 (0.03)		0.02 (0.02)					0.03 (0.03)	0.02 (0.02)	0.94	0.63	44
	26-30	LR	0.30^{***} (0.08)	0.26^{***} (0.09)	0.09^{**} (0.04)	0.14^{***} (0.05)		0.09 (0.04)		0.02 (0.02)	0.05^{*} (0.03)	0.02 (0.02)					0.02 (0.02)		0.97	0.59	44
	31-50	LR	0.34^{***} (0.08)	0.13^{***} (0.05)	0.09^{**} (0.04)	0.14^{***} (0.05)	0.08^{*} (0.04)		0.11^{**} (0.05)	0.05^{*} (0.03)	0.05^{*} (0.03)	0.02 (0.02)							0.96	0.56	44
2	16-20	DR	0.18^{**} (0.08)	0.07 (0.06)	0.11^{**} (0.06)	0.08^{**} (0.04)	0.08^{*} (0.06)	0.23^{**} (0.09)				0.04 (0.03)	0.02 (0.04)		0.08^{*} (0.05)	0.05 (0.05)		0.03 (0.03)	0.92	0.68	38
	26-30	NB	0.32^{***} (0.09)	0.15^{*} (0.09)	0.13^{**} (0.06)	0.08^{**} (0.04)	0.08 (0.08)		0.04 (0.04)	0.08 (0.06)			0.07 (0.05)	0.06^{*} (0.04)					0.95	0.67	38
	31-50	LR	0.32^{***} (0.07)	0.16^{**} (0.07)	0.11^{**} (0.05)	0.11^{**} (0.05)	0.07^{*} (0.05)	0.03 (0.03)	0.03 (0.03)		0.05^{*} (0.04)	0.02 (0.02)	0.02 (0.02)	0.05^{*} (0.03)					0.96	0.67	38

in the paper, we chose to set $N = 4000$ because that provides a nice balance of strategy differentiation while not requiring too much computation.

Appendix G. Additional MLE estimates

We use the strategy frequency estimation method (Dal Bó and Fréchette, 2011; Fudenberg et al., 2012) to find strategies that are the most likely among the population, given the observed actions.[27] 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

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

And the likelihood function is:

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

Table G5 presents the estimation results for the full set of 20 strategies used in Fudenberg et al. (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 et al. (2012) can be constructed with a finite number of rules in our setting. For each treatment, bootstrapped standard errors are calculated by drawing 200 random samples size. Specifically, to get each random sample we take all subjects in a given treatment, and draw them with replacement until the sample has as many subjects as the treatment. Next, we estimate the strategy frequencies corresponding to each of these samples. Finally, we calculate the standard deviation of the sampling distribution (Efron and Tibshirani, 1986).

We find strong evidence of TFT, GRIM, ALLD, and D.TFT, which were also prominent in the cluster analysis in Section 3.3. Next, we investigate MLE estimates for each cluster of rule sets.

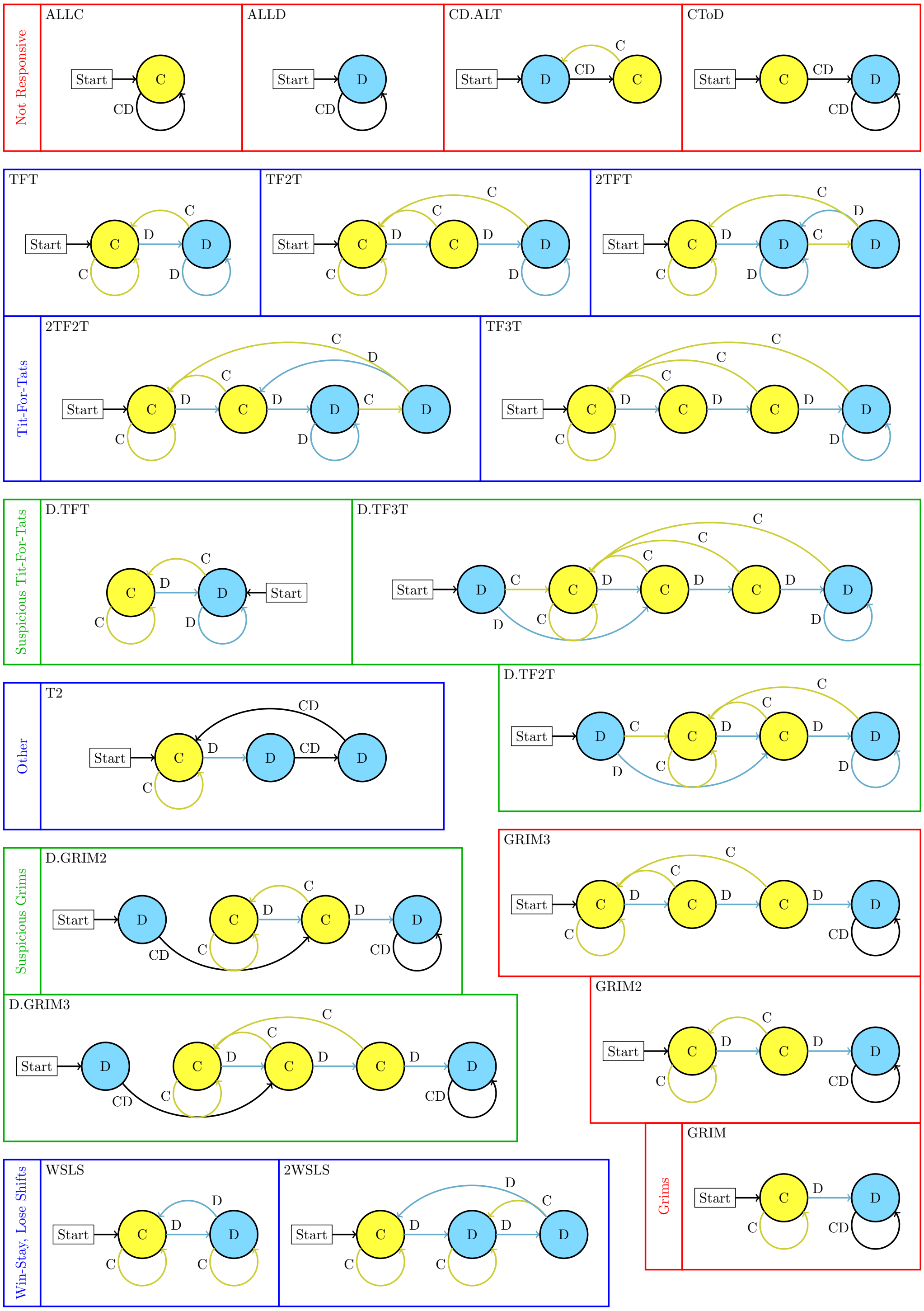
We run an MLE on eight separate subsets of the data, with each corresponding to one of the eight representative strategies obtained in Section 3.3. Table G6 presents the results. We find that the estimated strategies and realized cooperation rates across the clusters are substantially different. Specifically, estimates within each cluster match the cluster exemplar very well (one exception is GRIM2 cluster). The fact that MLE estimates for the obtained clusters have very little overlap, provide further evidence that participants' behavior among the clusters obtained in Section 3.3 is substantially different.

[27] More papers that use this approach are cited in Dal Bó and Fréchette (2017), Vespa (2011), Camera et al. (2012), and Fréchette and Yuksel (2017).

Table G6
Maximum likelihood estimates for strategy clusters. *Notes*: Only subject-supergames that were in the three cluster where included for the estimation. 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.

Cluster	TFT	GRIM	D.TFT	ALLD	2TFT	GRIM2	TF2T	ALLC	WSLS	D.TF2T	T2	CToD	GRIM3	TF3T	β	Cooperation	# of Subjects
TFT	0.83^{***} (0.08)				0.09^{*} (0.06)			0.04 (0.04)	0.04 (0.04)						0.96	0.68	23
GRIM		0.84^{***} (0.09)				0.13^{*} (0.08)						0.03 (0.04)			0.96	0.54	16
D.TFT			0.80^{***} (0.14)							0.20^{*} (0.14)					0.94	0.45	10
ALLD				0.90^{***} (0.09)								0.10 (0.09)			0.96	0.05	10
2TFT				0.14 (0.13)	0.71^{***} (0.17)						0.14 (0.13)				0.97	0.58	7
GRIM2	0.83^{***} (0.15)					0.17 (0.15)									1.00	0.69	6
TF2T	0.25 (0.22)						0.75^{**} (0.22)								0.94	0.81	4
ALLC								0.80^{***} (0.20)					0.11 (0.10)	0.09 (0.09)	0.93	0.93	5
WSLS									1.00 (0.00)						1.00	0.83	1

Appendix H. Description of strategies



Appendix I. 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 other subjects in the experiment. Specifically, we take each the 82 strategies from the locked-response stage of the experiment and match it with each of the other 81 strategies for 30 supergames (with supergame

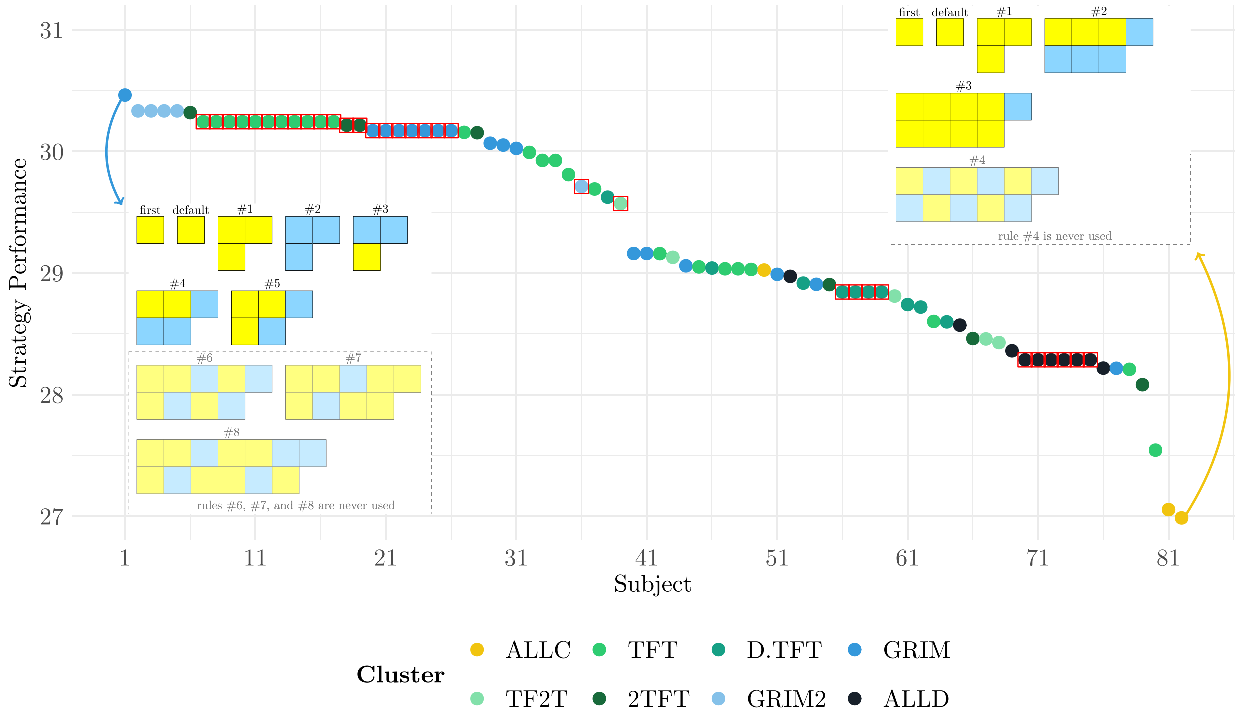


Fig. I6. Strategy performance across subjects. *Notes*: Clusters are identified by color. Red square denotes an exact match to one of the 20 common strategies. Best-performing rule-set is presented in the bottom left corner. Worst-performing rule-set is presented in the top right corner.

length determined randomly using continuation probability $\delta = .95$). We then calculate the average earnings per period and rank observed strategies from best to worst. Fig. I6 presents the results with the cluster of each strategy identified.

Fig. I6shows that, on average, GRIM and TFT strategies do best, while ALLD and ALLC strategies do poorly. The best-performing rule set in the experiment is one that combines features of GRIM and GRIM2 strategies and belongs to the GRIM cluster. This strategy differs from GRIM in that the rule $CD \to D$ is divided into two separate rules (Rule #4 and Rule #5). This allows this strategy to immediately punish a defection by the other player in every period except for the first period. Therefore, it is lenient in the first two periods of a supergame, but immediately triggers in the later periods of the supergame. Given a relatively large fraction of D.TFT, the success of this strategy is unsurprising because, unlike TFT and GRIM, it always cooperates with D.TFT.

The worst-performing rule set in the experiment belongs to the ALLC cluster, though it does not match the ALLC strategy exactly. Specifically, it differs from the ALLC strategy in two ways. First, it punishes once for every three defections by the opponent, which could be useful when matched with a strategy that defects a lot (e.g., ALLD or GRIM). Second, it has a memory-4 *CsToD* rule, which defects after four periods of mutual cooperation. This rule could cause a breakdown in cooperation when matched with a cooperative but unforgiving strategy such as GRIM, but could be useful against a more lenient strategy such as GRIM2. Notice that even the worst-performing strategy receives a higher payoff against the population than the mutual defection payoff of 25.

References

Aoyagi, M., Bhaskar, V., Frechette, G. R., 2015. The Impact of Monitoring in Infinitely Repeated Games: Perfect, Public, and Private. Working Paper.
Axelrod, R., 1980a. Effective choice in the prisoner's dilemma. J. Confl. Resolut. 24 (1), 3–25.
Axelrod, R., 1980b. More effective choice in the prisoner's dilemma. J. Confl. Resolut. 24 (3), 379–403.
Bigoni, M., Casari, M., Skrzypacz, A., Spagnolo, G., 2015. Time horizon and cooperation in continuous time. Econometrica 83 (2), 587–616.
Bodenhofer, U., Kothmeier, A., Hochreiter, S., 2011. APCluster: an R package for affinity propagation clustering. Bioinformatics 27 (17), 2463–2464.
Breitmoser, Y., 2015. Cooperation, but no reciprocity: Individual strategies in the repeated Prisoner's Dilemma. Am. Econ. Rev. 105 (9), 2882–2910.
Bruttel, L., Kamecke, U., 2012. Infinity in the lab. how do people play repeated games? Theory Dec. 72 (2), 205–219.
Camera, G., Casari, M., Bigoni, M., 2012. Cooperative strategies in anonymous economies: an experiment. Games Econ. Behav. 75 (2), 570–586.
Cason, T.N., Mui, V.-L., 2017. Individual Versus Group Choices of Repeated Game Strategies: A Strategy Method Approach. Purdue Working Paper.
Dal Bó, P., Fréchette, G.R., 2011. The evolution of cooperation in infinitely repeated games: experimental evidence. Am. Econ. Rev. 101 (1), 411–429.
Dal Bó, P., Fréchette, G.R., 2016. On the determinants of cooperation in infinitely repeated games: a survey. J. Econ. Lit. 56 (1), 60–114.
Dal Bó, P., Fréchette, G.R., 2017. Strategy Choice in the Infinitely Repeated Prisoners' Dilemma. Working Paper.
Duffy, J., 2008. Macroeconomics: a survey of laboratory research, Handbook of experimental economics 2.
Efron, B., Tibshirani, R., 1986. Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. Stat. Sci. 1 (1), 54–75.
Embrey, M., Fréchette, G.R., Yuksel, S., 2015. Cooperation in the Finitely Repeated Prisoner's Dilemma. Working Paper.

Embrey, M., Mengel, F., Peeters, R., 2016. Eliciting Strategies in Indefinitely Repeated Games of Strategic Substitutes and Complements. Working Paper.
Engle-Warnick, J., Slonim, R.L., 2004. The evolution of strategies in a repeated trust game. J. Econ. Behav. Org. 55 (4), 553–573.
Fréchette, G.R., Yuksel, S., 2017. Infinitely repeated games in the laboratory: four perspectives on discounting and random termination. Exp. Econ. 20 (2), 279–308.
Frey, B.J., Dueck, D., 2007. Clustering by passing messages between data points. Science 315 (5814), 972–976.
Fudenberg, D., Rand, D.G., Dreber, A., 2012. Slow to anger and fast to forgive: cooperation in an uncertain world. Am. Econ. Rev. 102 (2), 720–749.
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., 2015. Subject pool recruitment procedures: organizing experiments with Orsee. J. Econ. Sci. Assoc. 1 (1), 114–125.
Lugovskyy, V., Puzzello, D., Sorensen, A., Walker, J., Williams, A., 2017. An experimental study of finitely and infinitely repeated linear public goods games. Games Econ. Behav 102, 286–302.
Phipson, B., Smyth, G.K., 2010. Permutation p-values should never be zero: calculating exact P-values when permutations are randomly drawn. Stat. Appl. Genet. Mol. Biol. 9 (1).
Romero, J., Rosokha, Y., 2016. The Evolution of Cooperation: The Role of Costly Strategy Adjustments. Purdue Working Paper.
Roth, A.E., Murnighan, J.K., 1978. Equilibrium behavior and repeated play of the Prisoner's Dilemma. J. Math. Psychol. 17 (2), 189–198.
Selten, R., 1967. Die strategiemethode zur erforschung des eingeschrÃ rationalen verhaltensim rahmen eines oligopolexperiments. BeitrÃzur Experimentellen Wirtschaftsforschung, Tiubingen: J. C. B. Mohr, pp. 136–168.
Selten, R., Mitzkewitz, M., Uhlich, G.R., 1997. Duopoly strategies programmed by experienced players. Econometrica: J. Econ. Soc. 517–555.
Stahl, D.O., 2011. Cooperation in the sporadically repeated prisoners' dilemma via reputation mechanisms. J. Evol. Econ. 21 (4), 687–702.
Stahl, D.O., 2013. An experimental test of the efficacy of a simple reputation mechanism to solve social dilemmas. J. Econ. Behav. Org. 94, 116–124.
Vespa, E., 2011. Cooperation in Dynamic Games: An Experimental Investigation. SSRN. Available at SSRN 1961450