Econometrica, Vol. 91, No. 6 (November, 2023), 2295–2331

MIXED STRATEGIES IN THE INDEFINITELY REPEATED PRISONER'S DILEMMA

JULIAN ROMERO
Department of Economics, University of Arizona

YAROSLAV ROSOKHA
Department of Economics, Purdue University

Identifying the strategies that are played is critical to understanding behavior in repeated games. This process is difficult because only choices (not strategies) are observable. Recently, a debate has emerged regarding whether subjects play mixed strategies in the indefinitely repeated prisoner's dilemma. We use an experimental approach to elicit mixed strategies from human subjects, thereby providing direct empirical evidence. We find that a majority of subjects use mixed strategies. However, the data also suggest subjects' strategies are becoming less mixed over time, and move toward three focal pure strategies: Tit For Tat, Grim Trigger, and Always Defect. We use the elicited strategies to provide an empirically-relevant foundation for analyzing commonly used mixture model estimation procedures.

KEYWORDS: Indefinitely repeated prisoner's dilemma, mixed strategies, experimental design, strategy elicitation, finite mixture models.

1. INTRODUCTION

THE REPEATED PRISONER'S DILEMMA has been used to study a variety of topics in economics.[1] Theoretical work has largely focused on folk theorems, which show that as long as players are sufficiently patient, any payoff profile above a minimum threshold can be obtained as an equilibrium payoff. Experimental work has provided an important complement to the theory by testing which of the plethora of equilibria will emerge in a given setting. Analysis of directly observable outcomes in experiments (e.g., cooperation and defection) is relatively straightforward, whereas analysis of the underlying (nonobservable) strategies that generate these outcomes is more difficult and often requires assumptions about the set of strategies that subjects use. Depending on the assumptions about the initial set, some studies have found strong evidence that subjects use mixed strategies, whereas others have found strong evidence that subjects use pure strategies. This paper aims to provide direct evidence on whether subjects use mixed strategies in the indefinitely repeated prisoner's dilemma. In addition, our approach can be used to provide an empirical foundation for strategy types when inferring strategies from actions using finite mixture models.

Julian Romero: julian@jnromero.com
Yaroslav Rosokha: yrosokha@purdue.edu

We thank four anonymous referees for their invaluable suggestions during the review process. This paper also benefited from discussions with and comments from Jasmina Arifovic, Luke Boosey, Gabriele Camera, Tim Cason, David Cooper, Guillaume Frechette, Christos Ioannou, Volodymyr Lugovskyy, Ryan Oprea, Sevgi Yuksel, as well participants at the 2018 CEF conference in Milan, the 2018 World ESA meetings in Berlin, 2018 Southern Economic Association meetings in Washington D.C., 2019 North American ESA meetings in Los Angeles and seminar participants at Purdue University, University of Arizona, and Washington University in St. Louis.

[1]Examples include quantity-setting oligopolies (Mailath and Samuelson (2006)), R&D races (Cockburn and Henderson (1994)), trade wars (Maggi (1999)), international relations (Powell (1993)), and labor negotiations (Kahn (1993)).

In recent years, great strides have been made to better understand the strategies that subjects play in the repeated prisoner's dilemma (Dal Bó and Fréchette (2018)). The literature has two general approaches for understanding strategies. The first involves direct observation of strategies using strategy elicitation (Selten (1967)), and the second involves indirect inference of strategies based on actions. Within the first stream of research, implementations of lab experiments vary based on the flexibility subjects have in specifying strategies. The implementation with the least flexibility is one where subjects choose from a prespecified list of strategies (e.g., Cason and Mui (2019)). A more flexible implementation involves a prespecified set of contingencies (e.g., memory-1 histories), but the subject chooses an action after each contingency (e.g., Dal Bó and Fréchette (2019)). Lastly, the most flexible implementation allows the subject to specify both contingencies and which actions to play after each contingency (e.g., Romero and Rosokha (2018)).[2] The less flexible approaches have three primary advantages: (i) the experiment is easier to explain to subjects, (ii) the strategy-choice procedure is more comparable to selecting actions in a direct-response experiment, and (iii) the analysis of strategies is less complex. The less flexible approaches also have several disadvantages. Specifically, the experimenter (i) chooses strategies ahead of time, (ii) provides subjects with an insight as to which behaviors are possible, and (iii) imposes more structure on interactions. As the implementation becomes more flexible, the advantages and disadvantages flip.

The second general approach involves indirect inference of strategies based on actions. The most common method of indirect inference is to use a finite mixture model to estimate the proportion of the population that uses a particular strategy.[3] In particular, finite mixture models have been frequently used to estimate strategies in the indefinitely repeated prisoner's dilemma (Dal Bó and Fréchette (2011), Camera, Casari, and Bigoni (2012), Fudenberg, Rand, and Dreber (2012), Breitmoser (2015), Aoyagi, Bhaskar, and Fréchette (2019)).[4] The advantage of using the indirect-inference approach over the strategy-elicitation approach is that the researcher does not interfere with or restrict subjects' behavior during the experiment. The disadvantage of the approach is that at the estimation stage, the researcher has to assume which strategies (or strategy types) subjects use. Crucially, different assumptions about the initial set may lead to different results and conclusions. For example, using data from one treatment from Dal Bó and Fréchette (2011), Fudenberg, Rand, and Dreber (2012) estimate that 55% of the subjects use the pure-strategy Tit-For-Tat (henceforth TFT), whereas Breitmoser (2015) estimate 90% of subjects use a single type of mixed strategies termed Semi-Grim (henceforth SG).

This paper makes contributions to both the strategy elicitation and the indirect-inference approaches. In regards to the strategy-elicitation literature, our experimental design is the first (to our knowledge) that allows for the elicitation of history-contingent mixed strategies in repeated games.[5] Previous studies that have elicited strategies in the

[2]An even more flexible implementation, which is not portable to the lab setting, involves subjects creating computer programs that make choices for them. Examples of this approach include Axelrod (1980a,b) and Selten, Mitzkewitz, and Uhlich (1997).

[3]An alternative approach by Engle-Warnick and Slonim (2006) determines the best-fitting set of strategies based on a fitness function and an increasing cost for larger sets of strategies.

[4]Finite mixture models have been used in games other than the prisoner's dilemma. For example, some of the earliest work that uses mixture model estimation studies level-k strategies in a p-beauty contest (Stahl and Wilson (1994, 1995), Haruvy, Stahl, and Wilson (2001)).

[5]A small set of studies experimentally investigate nonhistory-contingent mixed repeated-game strategies (Bloomfield (1994), Ochs (1995), Shachat (2002), Chmura and Güth (2011), Noussair and Willinger (2011), Cason, Friedman, and Hopkins (2013)).

indefinitely repeated prisoner's dilemma (Dal Bó and Fréchette (2019), Romero and Rosokha (2018), Cason and Mui (2019), Gill and Rosokha (2020)) have asked subjects to specify a pure strategy.[6] In our experiment, a subject specifies a mixed strategy by specifying a probability (up to 2 decimal places) with which to play one of the two actions (C or D) contingent on the actions in the previous periods. For example, a subject could specify to play C with a probability of 70% if DC was played in the previous period. In the main treatment, our design allows subjects to construct mixed repeated-game strategies in which they have different probabilities of playing C or D in the first period and after each of the four possible memory-1 histories.[7] Therefore, 101^5 possible strategies are available to the subjects.

Regarding the literature that uses indirect inference to study strategies in the indefinitely repeated prisoner's dilemma, the results of our experiments provide information about the types of strategies that should be considered. In particular, the existing literature has either restricted strategies to pure strategies (e.g., Dal Bó and Fréchette (2011)) or has specified a parameterized set of strategy types that encompass both pure and mixed strategies (e.g., Breitmoser (2015)). The literature that restricts the set of strategies to pure strategies has found that the majority of behavior is captured by just three strategies—TFT, Grim Trigger (henceforth GRIM), and Always Defect (henceforth ALLD). For example, in a recent review, Dal Bó and Fréchette (2018) report that these three pure strategies account for the majority of behavior in 15 out of 17 treatments across five papers that estimate strategies. At the same time, a study that used a parameterized set of strategy types (Breitmoser (2015)) finds that a specific mixed-strategy type, SG, explains most behavior.[8] Because we directly elicit mixed repeated-game strategies, we can empirically validate whether subjects use mixed strategies in general, and SG in particular. In addition, we can evaluate the extent to which the two indirect-inference approaches can correctly identify strategies in our experiment.

We have three main results. Our first main result concerns mixed strategies. Specifically, we find that, across all treatments of our experiment, we find that a majority of elicited strategies are mixed.[9] Despite the large percentage of mixed strategies, we do not find evidence of SG. Instead, the most common mixed strategies played in our experiment come from a class of strategies that we refer to as the mixed-tit-for-tat (henceforth MTFT) class of strategies. Although these strategies are similar to SG in that they play C after CC, D after DD, and randomize after CD and DC, they differ from SG in that they have different probabilities of cooperation after CD and DC (in the direction of TFT).

Our second main result concerns the three pure strategies—TFT, GRIM, and ALLD—commonly studied in the literature. Specifically, in the main treatment of the experiment, we find approximately 30% of subjects construct one of these three strategies exactly, despite the fact that 101^5 possible strategies are available in our experiment. We also find

[6]The exception is one treatment from Dal Bó and Fréchette (2019), in which subjects can choose a strategy that randomizes with the same (subject specified) probability in every period regardless of the history.

[7]In the robustness treatments that we discuss in Section 4, we relax the memory-1 restriction.

[8]Backhaus and Breitmoser (2018) do a further analysis of data from a wider range of repeated prisoner's dilemma experiments (32 treatments) using data-mining techniques. They still find strong support that a majority of subjects' behavior can be explained by SG.

[9]There are several reasons why a subject may want to use a mixed strategy. First, they may have a preference for randomization (e.g., Agranov and Ortoleva (2017)). Second, they may be playing a belief-free mixed-strategy equilibrium (Ely, Hörner, and Olszewski (2005)). Third, they may be playing a mixed strategy to learn about the strategies of others. For example, a subject playing TFT cannot distinguish between opponents who are playing ALLC, TFT, and GRIM. Adding noise (either via mixed strategies or via longer rules as in Romero and Rosokha (2018)) is a way to learn about others.

that another 25% of subjects construct strategies that are close to these three pure strategies. Finally, when examining learning dynamics in our experiment, we see a clear trend of subjects moving away from mixed strategies and toward these three pure strategies.

Our third main result concerns the use of finite mixture models for estimating the proportion of strategies in the indefinitely repeated prisoner's dilemma. In particular, we find that when running a finite mixture model estimation with a set of pure strategies on the data from our experiment, the approach groups the mixed-versions TFT, GRIM, and ALLD into the proportion of pure TFT, GRIM, and ALLD. At the same time, when running the estimation using the set of strategy types from Breitmoser (2015), the approach identifies SG as being one of the most popular strategies, despite the fact that no subjects actually constructed the SG strategy in the experiment. In addition, we show that the flexibility of mixed strategies can make them difficult to model and can lead to unreliable interpretations of the mixture model estimates, especially when the first period is not taken into account.

In addition to the main treatment, we run robustness treatments in which we allow subjects to specify strategies longer than memory-1. The data from these treatments suggest that the conclusions from the main treatment hold. In particular, there is randomness in a majority of strategies, but the three pure strategies—TFT, GRIM, and ALLD—account for much of the data. We also find that relative to the main treatments, subjects substitute away from memory-1 mixed strategies and toward pure strategies that condition on longer histories.

The rest of the paper is organized as follows: Section 2 presents details of our experimental design. Section 3 presents the results, including strategies observed in our experiment (Section 3.2), the finite mixture model estimation (Section 3.3), and the evolution of strategies (Section 3.4). Section 4 discusses experimental treatments designed to test the robustness of our main treatment. Lastly, Section 5 provides a concluding discussion of the results and outlines directions for future research.

2. EXPERIMENTAL DESIGN

In this section, we describe the interface and the experimental design. Specifically, we modify the experimental interface of Romero and Rosokha (2018) to allow the elicitation of mixed strategies. Figure 1 presents a screenshot of the experimental interface used for the main treatment.[10] Next, we highlight the important aspects of the experimental interface and design.

2.1. *Rules and Strategies*

The main component of our experimental interface is the ability to construct strategies using a set of "if [input]-then [output]" rules. The input of a rule is an action profile in at most one previous period, and the output of a rule is the probability of playing a particular action. Subjects are able to modify strategies by adding and subtracting rules from their rule set. The rule set will then make a choice for a subject in a given period based on the history. The choice is determined by the rule that has the same input as the last period of the history. If a rule set does not contain a rule that has the last period of the history as an input, then the default rule will be used to make the choice. Subjects are required to

[10]The interface was developed using a Simple Toolbox for Experimental Economics Programs (STEEP). Further information about the interface can be found at http://jnromero.com/research/mixedStrategyChoice.

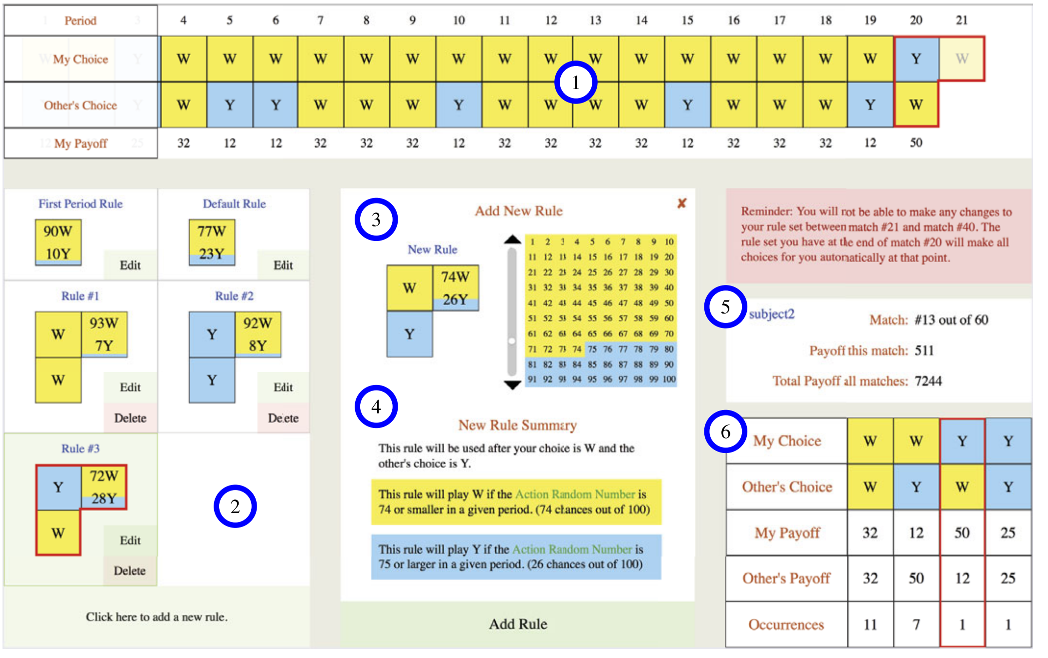


FIGURE 1.—Screenshot of the Experimental Interface. *Note*: The neutral action names W and Y correspond to the usual action names C and D from the prisoner's dilemma. The screenshot shows: (1) History, (2) Rule Set, (3) Rule Constructor, (4) New Rule Summary, (5) General Information, (6) Payoff Table.

specify both a default rule and a first-period rule before their rule set makes any choices for them, which ensures that the rule set is able to specify an action after every history.

There are two main differences between the current experiment and Romero and Rosokha (2018). First and foremost, in the current experiment, we allow subjects to specify rules with probabilistic outputs. This modification allows us to study mixed strategies in the indefinitely repeated prisoner's dilemma, which is the main goal of this paper. Second, we restrict subjects to memory-1 rules (no more than one period as an input). This modification makes the strategy-elicitation process less complex while still allowing subjects to construct strategies of interest from Dal Bó and Fréchette (2018) (e.g., GRIM, TFT, DTFT, and ALLD), SG from Breitmoser (2015), and belief-free equilibrium strategies from Ely, Hörner, and Olszewski (2005). One way to view our choice, is that we give memory-1 mixed strategies, such as SG, the best chance to succeed. In addition to our main treatment, we ran two more treatments in which we remove the restriction to memory-1. The results are discussed in Section 4.

Our experimental interface allows subjects to construct rules with probabilistic outputs by using the rule constructor (#3 in Figure 1). Specifically, subjects can use a slider to specify a cutoff between 0 and 100.[11] The cutoff determines the probability with which C is

[11]When a subject decides to create a rule, the rule constructor appears with "?"s in each box. When the subject clicks on one of the boxes corresponding to the rule input, the box changes from "?" to either "W" or "Y" randomly. If the box already has a "W" or a "Y," then it changes to the other action when clicked. To set the output of the rule, the subject needs to click the slider next to the rule. The slider has no default value. The marker on the slider is not visible until the subject clicks on it. Once the subject has clicked the slider, the proportion of squares corresponding to the probability of playing "W" are colored yellow, the proportion

selected. The way we implement randomization and explain it to subjects is by drawing an "action random number" each period. The action random number is an integer between 0 and 100 (inclusive). If the integer is less than or equal to the cutoff, then C is played. If the integer is greater than the cutoff, then D is played. Subjects are reminded that each of them receives their own independent draw of the action random number in each period.

Figure 1 (#2) presents an example of a rule set that can be constructed with the interface.[12] We denote this rule set as $\{FP \to 90; \to 77; CC \to 93; DD \to 92; DC \to 72\}$. Given this strategy, the subject will cooperate with a 90% probability in the first period, and will cooperate with probability 93%, 77%, 72%, 92% if CC, CD, DC, or DD was played in the previous period, respectively. Note that because no rule with input CD was created, the default rule will be used to make the choice after that action profile.

2.2. *Experimental Protocol*

The main treatment of the experiment consisted of six sessions run at the Vernon Smith Experimental Economics Laboratory at Purdue University in April 2018. Details of each session are provided in Table A-2 in the Appendix. Each session consisted of instructions, an incentivized quiz to ensure that subjects understood the instructions, and 60 supergames. All payoffs were displayed in Experimental Currency Units (ECUs) and were converted to USD at the end of the experiment at 2000 ECUs equals 1 USD. Next, we describe specific parts of the experimental design in more detail.

2.2.1. *Game Parameters*

We picked the parameters for the experiment to match those of Romero and Rosokha (2018) and one treatment of Dal Bó and Fréchette (2019). Specifically, in the main treatment we used the stage game payoffs that are displayed in Figure 1 (#6) and the continuation probability $\delta = 0.95$. These parameters allow a direct comparison to Romero and Rosokha (2018) and one treatment of Dal Bó and Fréchette (2019). Four sequences of 60 supergame lengths were pre-drawn using a computer according to continuation probability $\delta = 0.95$ (see Tables A-1 and A-2 in the Appendix).

2.2.2. *Instructions and Quiz*

Instructions used in this experiment consisted of a sequence of interactive screens that explained all aspects of the experiment and details of the experimental interface. The instructions contained 20 quiz questions. The quiz was incentivized as follows. Subjects earned \$5.00 if they answered at least 15 out of 20 questions correctly, and \$0.00 otherwise. Among the 158 subjects who participated in the experiment, 124 passed the quiz.[13]

of squares corresponding to the probability of playing "Y" are colored blue, and the corresponding numbers are summarized in the rule output square. In addition, a written summary of the rule is displayed below the constructed rule. Detailed screenshots of this process are presented in Figure A-1 in the Appendix.

[12] Figure A-2 in the Appendix provides a few more examples of strategies that can be constructed in our memory-1 treatment, including TFT, GRIM, SG, and MTFT. Figure A-2 in the Appendix provides some examples of strategies that can be constructed in our unrestricted treatment, including TF2T and Lenient GRIM.

[13] Passing the quiz could be correlated with IQ, which has been shown to affect cooperation (Segal and Hershberger (1999), Jones (2008), Proto, Rustichini, and Sofianos (2019, forthcoming)). However, in this experiment, we did not collect IQ variables, so we cannot assess the extent to which this is an issue. Among the demographic variables that we did collect, the statistic that stood out was whether the participant attended high school outside of the US. In particular, 47% of those that did not pass the quiz attended high school outside of the US, indicating that one of the reasons for doing poorly on the quiz may have been understanding of the language.

The subjects who passed the quiz were randomly matched into groups with each other. To keep groups relatively similar in size, we decided to have a minimum-possible group size of 8 and a maximum-possible group size of 14. Therefore, if only 12 subjects passed the quiz, they were all matched in the same group, but if 16 subjects passed the quiz, then they were divided into two groups of 8. In our experiment, each session ended up with two or three groups, the smallest of which contained 8 subjects and the largest of which contained 12 subjects.

2.2.3. *Experimental Stages*

Similar to Dal Bó and Fréchette (2019) and Romero and Rosokha (2018), we implemented three types of supergames: direct response, nonbinding, and locked response. Next, we briefly describe each stage and its purpose.

2.2.3.1. *Direct-Response Stage (Supergames 1–10).* In the direct-response stage, subjects play the game by choosing C or D each period. The direct-response stage ensures that subjects learn about the strategic tension in the game, without having to specify strategies. One difference between Romero and Rosokha (2018) and the current experiment is that we required subjects to confirm their opponent's action after each period. More specifically, they received the following message: "To continue click the key corresponding to the choice of the participant that you are matched with from the previous period on the keyboard (either W or Y)." We added this confirmation to ensure that subjects had a chance to process the choice of their opponent before making their choice in the next period. We did so to avoid situations like the following: Suppose two subjects play C for many periods, and then one subject plays D for one period, and the other subject quickly continues to play C without necessarily processing that his or her opponent played D in the previous period. This design allows the subject to progress quickly if his or her opponent plays as expected, but is more likely to pause if his or her opponent plays contrary to what was expected.

2.2.3.2. *Nonbinding Stage (Supergames 11–20).* During the nonbinding stage, subjects were provided up to 10 minutes to construct the initial set of rules and up to 2 minutes before each additional supergame. These time limits were never close to binding (see Figure A-3 in the Appendix for details). Importantly, we placed no time limit on the duration of each period during a supergame. As the supergame progressed in the nonbinding stage, subjects were informed of the action that their rule set would play each period given their draw of the action random number. Subjects were given the option to manually deviate from the prescription of their rule set in every period of the nonbinding stage.[14] When subjects deviated from the prescription of their rule set, they received a warning that reminded them that in the locked-response stage, their rule set would automatically make their choices for them.

2.2.3.3. *Locked-Response Stages (Supergames 21–40 and Supergames 41–60).* In the locked-response stage, subjects' rule sets made choices for them automatically. Subjects were not able to change their rule sets during the locked-response stage. The current experiment consisted of two locked-response stages (as opposed to only one in Romero and Rosokha (2018)), and subjects were given up to 10 minutes to edit their rule sets between

[14]The option to manually deviate from the prescription of the rule set was rarely used (see Figure OA3 in the Online Appendix in the Supplementary Material (Romero and Rosokha (2023) for details).

the locked-response stages. The locked-response stage served as an incentive to construct (and understand) strategies during the nonbinding stage. We decided to include the second locked-response stage to ensure that subjects had sufficient time and experience to evaluate mixed strategies. In addition, comparing strategies between the first and second locked-response stages allows us to assess the evolution of strategies.

3. EXPERIMENTAL RESULTS

The results section is organized into four subsections. First, we present cooperation rates and compare them with prior studies that used similar parameters. Specifically, in Section 3.1, we provide a direct comparison of cooperation rates between our experiments and Romero and Rosokha (2018) (which elicits only pure strategies) and find no significant difference between the two studies. Second, the main goal of this experiment is to examine whether subjects play mixed strategies in the indefinitely repeated prisoner's dilemma. To achieve this goal, in Section 3.2, we analyze the elicited strategies. In particular, we use a clustering approach to group similar strategies and find eight clusters, four based on pure strategies—TFT, GRIM, ALLD, and DTFT—and four mixed strategies, which roughly match mixed versions of the pure strategies. In particular, we find the most frequent mixed strategy falls into an MTFT class of strategies. Third, we consider several mixture model specifications on the action data generated by the constructed strategies in our experiment. Specifically, in Section 3.3 we compare estimates when using strategy sets from the literature with estimates when using strategy sets motivated by our elicitation. Finally, we analyze the evolution of strategies. In particular, in Section 3.4, we investigate changes that subjects made between the two locked-response stages and find evidence that the strategies are becoming less random. In addition, we consider a simple long-run learning model to extrapolate trends in strategy changes and find that the limiting strategies are TFT, GRIM, and ALLD.

3.1. *Cooperation*

Table I presents the average cooperation rate observed for mixed-strategy elicitation stages in the current experiment (labeled *Current*) and pure-strategy elicitation stages in Romero and Rosokha (2018) (labeled *RR2018*). The cooperation rates are divided into groups of supergames based on the experimental design. In particular, supergames 1–10 were direct response, supergames 11–20 were nonbinding, and supergames 21–40 were locked response in both experiments.[15] For each of these groups of supergames, we present the cooperation rates for the first period, first four periods, last four periods, and all periods, as well as the cooperation rates after each memory-1 history. The cooperation

[15]Note that Romero and Rosokha (2018) had two treatments. Treatment 1 had direct response for supergames 1–10, nonbinding for supergames 11–20, locked response for supergames 21–50, and direct response for supergames 51–60. Treatment 2 had direct response for supergames 1–20, nonbinding for supergames 21–30, and locked response for supergames 31–60. The current experiment had direct response for supergames 1–10, nonbinding for supergames 11–20, locked response for supergames 21–40, and then another locked response for supergames 41–60. Thus, we can directly compare supergames 1–10 of the current experiment with both treatments of Romero and Rosokha (2018) (82 subjects), and directly compare supergames 11–20, and supergames 21–40 of the current experiment with Treatment 1 of Romero and Rosokha (2018) (44 subjects). Finally, we have no valid comparison for supergames 41–60, because both treatments in Romero and Rosokha (2018) had a different progression of stages and opportunities for strategy revision up to supergame 41.

TABLE I

AVERAGE COOPERATION RATE.

Experiment	Current		RR2018	Current		RR2018	Current		RR2018	Current
Groups	14		6	14		3	14		3	14
Subjects	124		82	124		44	124		44	124
Supergames	1–10		1–10	11–20		11–20	21–40		21–40	41–60
Type	DR		DR	NB		NB	LR		LR	LR
First Periods	0.56 (0.04)	$\overset{0.2}{\sim}$	0.65 (0.05)	0.66 (0.05)	$\overset{0.37}{\sim}$	0.75 (0.06)	0.65 (0.05)	$\overset{0.26}{\sim}$	0.77 (0.08)	0.68 (0.05)
First 4 Periods	0.53 (0.04)	$\overset{0.15}{\sim}$	0.62 (0.04)	0.61 (0.04)	$\overset{0.5}{\sim}$	0.67 (0.05)	0.54 (0.04)	$\overset{0.29}{\sim}$	0.66 (0.1)	0.53 (0.05)
Last 4 Periods	0.47 (0.04)	$\overset{0.21}{\sim}$	0.55 (0.05)	0.55 (0.04)	$\overset{0.75}{\sim}$	0.58 (0.06)	0.44 (0.04)	$\overset{0.33}{\sim}$	0.54 (0.07)	0.42 (0.05)
All Periods	0.49 (0.04)	$\overset{0.17}{\sim}$	0.57 (0.04)	0.57 (0.04)	$\overset{0.7}{\sim}$	0.61 (0.05)	0.47 (0.05)	$\overset{0.31}{\sim}$	0.58 (0.07)	0.46 (0.05)
CC	0.86 (0.03)	$\overset{0.81}{\sim}$	0.85 (0.04)	0.84 (0.03)	$\overset{0.37}{\sim}$	0.88 (0.04)	0.78 (0.03)	0.06	0.9 (0.04)	0.79 (0.03)
CD	0.35 (0.03)	$\overset{0.16}{\sim}$	0.4 (0.04)	0.27 (0.03)	$\overset{0.64}{\sim}$	0.24 (0.05)	0.17 (0.02)	$\overset{0.24}{\sim}$	0.13 (0.03)	0.12 (0.02)
DC	0.45 (0.03)	$\overset{0.63}{\sim}$	0.43 (0.05)	0.54 (0.04)	$\overset{0.88}{\sim}$	0.53 (0.06)	0.46 (0.04)	$\overset{0.2}{\sim}$	0.55 (0.07)	0.45 (0.04)
DD	0.13 (0.02)	$\overset{0.56}{\sim}$	0.15 (0.03)	0.16 (0.02)	$\overset{0.66}{\sim}$	0.14 (0.05)	0.14 (0.02)	$\overset{0.57}{\sim}$	0.11 (0.07)	0.11 (0.02)

Note: The average cooperation rate is computed by first averaging cooperation in each of the corresponding periods in each supergame, and then averaging across the corresponding supergames. Bootstrapped standard errors (in parentheses) are calculated by drawing 1000 random samples: first randomly draw the appropriate number of groups with replacement; then, for each group, randomly draw the appropriate number of subjects with replacement. If the supergame is less than four periods, then the cooperation rate for the first four periods and the last four periods is set to the cooperation rate for all periods. Supergames 1–10 correspond to the direct-response stage (labeled DR); Supergames 11–20 correspond to the nonbinding stage (labeled NB); supergames 21–40 and 41–60 correspond to the locked-response stage (labeled LR).

rate is computed by first averaging cooperation in each of the corresponding periods in each supergame, and then averaging across the corresponding supergames. In addition, we provide p-values from the nonparametric permutation test (Good (2013)).[16]

Table I shows that cooperation rates from the current experiment are close to the cooperation rates from the prior study, both when looking at the cooperation rates across different periods and when looking at cooperation rates after each memory-1 history. In addition, in the Online Appendix, we compare the current experiment with previous studies that use the same parameters (yet different elicitation procedures and designs). The results provide further evidence that the cooperation levels seen in the current experiment are in line with previous studies.

[16]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 across groups. The distribution of the difference in the average cooperation rate between the two treatments is obtained through random permutation of treatment labels among groups (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 difference between the two corresponding columns. Unless otherwise specified, all p-values reported in this paper are obtained using this test.

When comparing cooperation across the stages, we find the greatest change in cooperation after the CD history. The table shows that the cooperation rates drop from the direct-response stage (0.35 in supergames 1–10) to the nonbinding stage (0.27 in supergames 11–20), to the locked-response stages (0.17 and 0.12 in supergames 21–40 and 41–60, respectively). Three related reasons for the observed change in cooperation after the CD history are incentives, deliberation, and learning.[17] The first reason is the different nature of incentives in the direct-response stage compared to the strategy-elicitation stage. In particular, as is common in experiments on indefinitely repeated games, we compensated subjects based on the sum of earnings across all periods. Therefore, in the direct-response stage, each choice only affected the payoff in one period, whereas in the strategy-elicitation stage, the specified strategy affected payoffs in many periods. The second reason is deliberation, which has been shown to affect cooperation (Rand (2016)). Strategy elicitation requires subjects to think about their choice after multiple contingencies. In our experiment, we provided subjects with up to 10 minutes to make their strategies. The third reason is learning. If subjects are learning to play different strategies across supergames, then we should not necessarily expect the probabilities of cooperation to remain constant across a large number of supergames. Furthermore, an introspective learning process may progress differently than an experience-based learning process, especially when some contingencies are not experienced often.

To summarize, we find that cooperation rates observed within our experiment are in line with previous studies. Furthermore, with a closer examination, we find that cooperation rates change the most after the CD history. In Section 3.4, we further examine learning and show that this finding is consistent with a reduction in mixed strategies across our experiment.

3.2. *Strategies*

Figure 2 presents all of the strategies observed in the second locked-response stage of the current experiment.[18] A strategy vector consists of the five cooperation percentages corresponding to the histories (∅, *CC*, *CD*, *DC*, *DD*). The figure shows that 83 (out of 124) subjects specify a *mixed strategy* (a strategy that randomizes after at least one of the five histories) and 2 (out of 124) subjects specify a *pseudo-mixed strategy* (a strategy that randomizes only after a history that cannot be reached). The figure highlights all of these strategies with a black bar on the left-hand side of the strategy vector.

Strategies are organized using a clustering approach that groups together similar strategies without specifying any categories beforehand. In particular, we use the affinity propagation clustering algorithm (Frey and Dueck (2007)) with Euclidean distances between strategy vectors as the similarity criteria. Cluster analysis yields eight clusters, with each cluster characterized by an exemplar—the most representative member among all the

[17]Romero and Rosokha (2018) note a similar change in cooperation rate after the CD history between the direct-response and the strategy-elicitation stages. Two reasons that can be ruled out as primary causes for this change are as follows. First, the elicitation in this experiment allows subjects to play mixed strategies. In particular, if subjects do indeed play mixed strategies, implementing a mixed strategy in the locked-response stage would have been impossible in Romero and Rosokha (2018). Second, in the current experiment, we require subjects to confirm the choice of their opponent in the direct-response stage. Such confirmation was absent in Romero and Rosokha (2018), which could have led subjects to make choices too quickly in the direct-response stage. For example, after a long sequence of mutual cooperation, a subject may accidentally choose C after the other participant selects D. Although the cooperation after the CD history in the direct-response stage has decreased (0.40 vs 0.35), this difference is not significant (p-value 0.16).

[18]Figure A-4 in the Appendix presents strategies observed in the first locked-response stage.

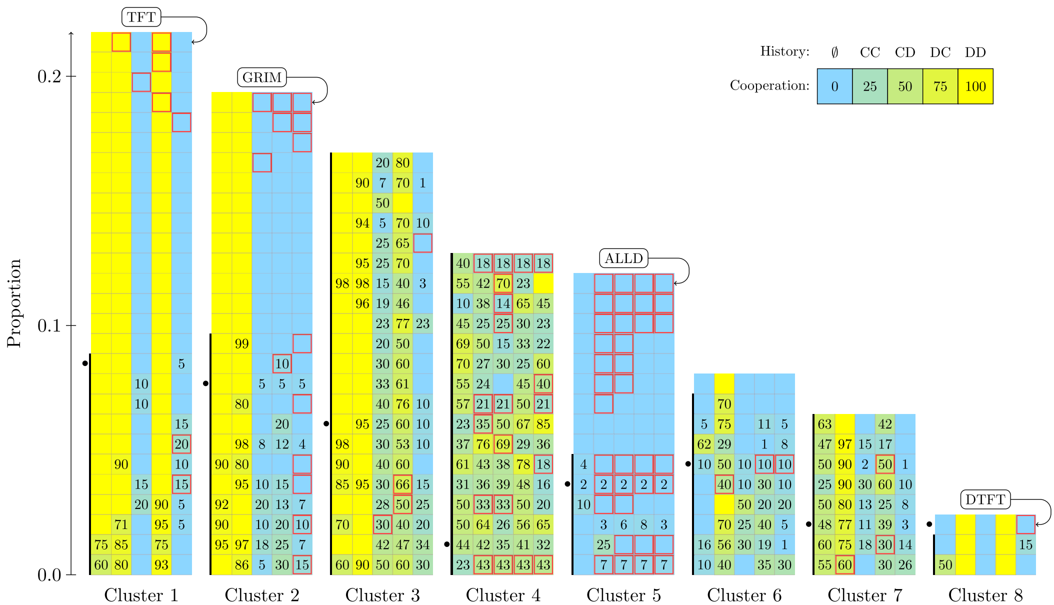


FIGURE 2.—Elicited Strategies. *Note*: Elicited strategies from the second locked-response stage are displayed. Strategies are presented as vectors of cooperation percentages after each of the five histories $(\emptyset, CC, CD, DC, DD)$. For ease of reading, 100 and 0 are not shown. A black bar on the left-hand side of the vector denotes a strategy with $M1RM > 0$. Solid squares denote a situation in which the default rule determines the percentage of cooperation. Strategies are grouped using the affinity propagation clustering algorithm. A bullet on the left-hand side denotes that the strategy is the cluster exemplar. Strategies in a given cluster are sorted according to $M1RM$. The pure strategies TFT, GRIM, ALLD, and DTFT are labeled.

strategies in that cluster. In Figure 2, we mark each exemplar with a black bullet on the left-hand side of the strategy vector. To examine the randomness of strategies across clusters, we define a memory-1 strategy randomness measure. Let *randomness measure* (*RM*) of cooperation percentage x be two times the minimum distance from x to 0 or 100. For example, a rule with an output of 80 would have an *RM* of 40 $(= 2 * (100 - 80))$, and a rule with an output of 10 would have an *RM* of 20 $(= 2 * (10 - 0))$. Furthermore, any rule with an output of 0 or 100 would have an *RM* of 0, a rule with an output of 50 would have an *RM* of 100, and a rule with an output drawn from a uniform distribution would have an expected *RM* of 50. Then the memory-1 strategy randomness measure, denoted by $M1RM$, is the average *RM* across the five rules. We find that clusters 1, 2, 5, and 8 have exemplars with $M1RM$ of at most 6.0. Strategies in these clusters are pure strategies or close to pure strategies. Combined, these four clusters contain 56% of subjects. Clusters 3, 4, 6, and 7 have exemplars with $M1RM$ of at least 32.0. Strategies in these clusters are predominately mixed strategies.

Among the pure-strategy clusters, Cluster 1 has 27 subjects, has an exemplar of $\{FP \to 100; CC \to 100; CD \to 0; DC \to 100; DD \to 5\}$ with an $M1RM$ of 2.0, and contains TFT. Cluster 2 has 24 subjects, has an exemplar of $\{FP \to 100; CC \to 100; CD \to 5; DC \to 5; DD \to 5\}$ with an $M1RM$ of 6.0, and contains GRIM. Cluster 5 has 15 subjects, has an exemplar of $\{FP \to 2; CC \to 2; CD \to 2; DC \to 2; DD \to 2\}$ with an $M1RM$ of 4.0, and contains ALLD. Finally, Cluster 8 has three subjects, has an exemplar of $\{FP \to 0; CC \to 100; CD \to 0; DC \to 100; DD \to 0\}$ with an $M1RM$ of 0.0, and contains DTFT. These four clusters represent pure strategies that have been consistently found in literature (Dal

Bó and Fréchette (2018)). To provide further evidence that strategies in these clusters are close to pure strategies, we compute (via simulations) the percentage of choices that differ from the choices made by the corresponding strategy without noise. We find that, on average, 97.2% of periods and 82.6% of supergames perfectly match the corresponding strategy.[19]

Among the mixed-strategy clusters, Cluster 3 has 21 subjects and has an exemplar of $\{FP \to 100; CC \to 95; CD \to 25; DC \to 60; DD \to 10\}$ and has an $M1RM$ of 32.0. Cluster 4 has 16 subjects and has an exemplar of $\{FP \to 44; CC \to 42; CD \to 35; DC \to 41; DD \to 32\}$ with an $M1RM$ of 77.2. Cluster 6 has 9 subjects and has an exemplar of $\{FP \to 10; CC \to 50; CD \to 10; DC \to 10; DD \to 10\}$ with an $M1RM$ of 36.0. Cluster 7 has 8 subjects and has an exemplar of $\{FP \to 48; CC \to 77; CD \to 11; DC \to 39; DD \to 3\}$ with an $M1RM$ 49.6. Although all but one strategy in these clusters are mixed, the four clusters represent a variety of randomizing behaviors. Strategies in Cluster 3 cooperate with a high probability after CC, with a low probability after DD, and with a higher probability after DC than CD. We refer to such strategies as the MTFT class of strategies. Strategies in Cluster 4 randomize with similar probability after every history. All but one strategy in Clusters 6 and 7 have the highest probability of cooperation after CC, and relatively low probabilities of cooperation after the CD, DC, and DD histories.

Notably absent from the mixed-strategy clusters is the previously studied SG strategy. Out of 124 subjects, none specified a strategy with $\{CC \to 100; CD \to \alpha; DC \to \alpha; DD \to 0\}$ for some value of $\alpha \in \{1, 2, \dots, 99\}$. As a further robustness check, we relax the criteria for SG by requiring the probability of cooperation after CC to be within x percentage points from 100, after CD and DC to be within x percentage points of each other, and after DD to be within x percentage points of 0. We find that even when $x = 25$, only 10 (out of 124) subjects could be classified as SG, while at the same time not being classified as GRIM by a similar criterion (see Figure A-5 in the Appendix).

To summarize, we find that 67% of elicited strategies are mixed, and 31% of elicited strategies are pure (the remaining 2% are pseudo-mixed strategies). Almost all of the pure strategies (37 out of 39) are the three commonly studied strategies TFT, GRIM, and ALLD. Furthermore, approximately one-third of the elicited mixed strategies are characterized as being close to these three pure strategies according to the clustering algorithm.[20] Among the remaining mixed strategies, a large proportion fall into the MTFT class of strategies. Strategies in this class cooperate after CC, defect after DD, and co-

[19]To run these simulations, we focus on the 4 clusters that correspond to the 4 commonly studied strategies TFT, GRIM, ALLD, and DTFT (clusters 1, 2, 5, and 8 in Figure 2). For each strategy in a corresponding cluster, we match that strategy against each of the 124 strategies and simulate the play for 100 supergames with continuation probability $\delta = 0.95$. All strategies used in the simulations are from the second locked-response stage. We find that in clusters TFT, GRIM, ALLD, and DTFT, constructed strategies match 97, 97, 99, and 97% of periods and 82, 81, 90, and 69% of supergames when compared to the corresponding pure strategy. Corresponding medians are 99, 99, 100, and 98% of periods and 91, 90, 100, and 55% of supergames.

[20]Another approach to test how close the elicited strategies are to pure strategies is to carry out a test similar to that performed in Dal Bó and Fréchette (2018). They find that across 32 treatments (with different discount factors and payoffs), a majority of observed sequences of play are consistent with five pure strategies: ALLC, ALLD, TFT, GRIM, and DTFT. That is, in a given supergame, if we replace a subject with a computer playing one of the five strategies, it would have played the same sequence of actions. We do this exercise on our data and find that across all 60 supergames and all subjects in the main treatment, observed play in 63.9% of supergames is identical to observed play from one of the five strategies. Furthermore, if we allow for one error, then observed play in an additional 11.7% of supergames matches one of these five strategies. If we restrict the data to the second locked-response stage, we find 68.8% identical match with an additional 10.4% when allowing for one error.

operate with a higher percentage after DC than CD. Finally, we find little evidence that subjects use the previously studied SG strategy.

3.3. *Mixture Model Estimation*

The main goal of this section is to better understand strategy-estimation techniques that have been used in the literature. When estimating strategies using mixture models, researchers are required to specify a set of input strategies to be estimated. Recent approaches have used strategy sets that have been founded on a combination of commonly studied theoretical strategies (Dal Bó and Fréchette (2011)), responses from unincentivized post-experimental survey questions (Fudenberg, Rand, and Dreber (2012)), and empirical regularities in cooperation after memory-1 histories (Breitmoser (2015)). The strategies elicited in our experiment provide an empirical comparison for the results of these strategy estimation techniques. In particular, using data from the second locked-response stage, we run the estimations based on actions and compare them with the actual strategies from the experiment.

Panel (a) of Table II presents estimation results using the strategy frequency estimation method of Dal Bó and Fréchette (2011) with the 20 pure strategies proposed in Fuden-

TABLE II

MIXTURE MODEL ESTIMATES.

(a) Strategy	(a) Proportion	(a) Parameter	(b) Strategy	(b) Proportion	(b) Parameter
TFT (1,1,0,1,0)	0.414 (0.056)	—	TFT (1,1,0,1,0)	0.196 (0.052)	—
ALLD (0,0,0,0,0)	0.22 (0.046)	—	ALLD (0,0,0,0,0)	0.145 (0.046)	—
GRIM (1,1,0,0,0)	0.214 (0.059)	—	GRIM (1,1,0,0,0)	0.179 (0.05)	—
DTFT (0,1,0,1,0)	0.058 (0.026)	—	DTFT (0,1,0,1,0)	0.024 (0.019)	—
CToD (1,0,0,0,0)	0.033 (0.021)	—	MTFT1 $(1,1,\alpha,1-\alpha,0),\alpha<.5$	0.198 (0.052)	0.292 (0.041)
2TFT —	0.027 (0.03)	—	CONST $(\alpha,\alpha,\alpha,\alpha,\alpha)$	0.08 (0.044)	0.453 (0.087)
GRIM3 —	0.016 (0.018)	—	MGRIM1 $(0,1-\alpha,\alpha,\alpha,\alpha),\alpha<.5$	0.049 (0.025)	0.094 (0.076)
ALLC (1,1,1,1,1)	0.01 (0.013)	—	MGRIM2 $(.5,1-\alpha,\alpha,\alpha,\alpha),\alpha<.5$	0.13 (0.045)	0.228 (0.098)
LL	-9044.027		LL	-6965.058	
BIC	9140.433		BIC	7022.901	

Note: For this estimation, we use the choice data from the second locked-response stage. The three columns in each panel display the strategy, the proportion of the population estimated to use that strategy, and the estimated value of the parameter ("–" if the strategy has no parameter). Bootstrapped standard errors are generated using 1000 samples by first randomly drawing groups and then randomly drawing an appropriate number of subjects for each group. Strategies are represented by the five-dimensional vector, which represents the probabilities of cooperation after the first period, and after memory-1 histories CC, CD, DC, and DD. The strategy estimation in panel (a) uses the 20 strategies from Fudenberg, Rand, and Dreber (2012). The procedure to construct the likelihood function is described in Appendix B. Eight strategies with the highest estimated frequencies are displayed. The strategy estimation in panel (b) uses the eight strategies identified from the cluster analysis in Section 3.2. To account for the mixed components of the strategies in panel (b), we adapt the procedure from Breitmoser (2015) to allow for the first period. Log-likelihood (LL) and Bayesian Information Criterion (BIC) values are displayed in the final two rows.

berg, Rand, and Dreber (2012) on our data. For comparison, panel (b) of Table II presents mixture model estimates using a strategy set motivated by our results in Section 3.2. In particular, we consider eight strategies corresponding to the eight clusters presented in Figure 2. The set contains four pure strategies (TFT, GRIM, ALLD, DTFT) and four mixed strategies (MTFT, CONST, MGRIM1, MGRIM2).[21] We find that strategies with the highest frequencies identified by the strategy frequency estimation method are the three pure strategies identified by our experiment. The strategy with the largest difference is TFT, which appears to be capturing the proportion of subjects that specify a strategy that falls in Cluster 3 that we classify as MTFT. This observation is consistent with the finding from Dal Bó and Fréchette (2019) that when a strategy is missing, its proportion goes to a related strategy. To summarize, though the panels have different proportions and likelihood values, the qualitative conclusions are similar in that the majority of subjects use TFT, GRIM, and ALLD.[22]

Panel (a) of Table III presents estimation results using a finite mixture model with the set of six parameterized strategies from Breitmoser (2015).[23] One important difference from the estimation in Table II is that the strategies proposed in Breitmoser (2015) do not take the first period into account, and are therefore only vectors of length four, corresponding to probabilities of cooperation after memory-one histories CC, CD, DC, and DD. Panel (b) of Table III presents estimation results from a set of six strategies based on the elicited strategies from the main treatment of our experiment. Because the strategies in this estimation do not take into account the first period, the strategies TFT and DTFT as well as MGRIM1 and MGRIM2 are considered the same. This leaves us with a set of six strategies—three pure strategies and three parameterized strategies.[24] We find that the strategy with the highest frequency when using the strategy set from Breitmoser (2015) is SG at 35%. This finding is surprising, given that we find little evidence of SG among the elicited strategies.

To try to understand which strategies are driving the high estimate of SG, we classify subjects based on the likelihood of each strategy generating the observed play from that subject. Specifically, for the estimated strategies from Table III(a) and each of the 124 subjects, we determine which of the strategies is most likely to generate the observed se-

[21]When determining the strategy corresponding to each cluster, we restricted to strategies that had a single parameter. MTFT corresponds to Cluster 3. Strategies in this cluster cooperate with high probability in the first period and after CC, cooperate with low probability after DD, and mix with relatively high probability after DC and relatively low probability after CD. We parameterize this strategy as $(1, 1, 1-\alpha, \alpha, 0)$, $\alpha \in (0.5, 1)$. CONST corresponds to Cluster 4. Strategies in this cluster cooperate with similar probabilities after each history. We parameterize this strategy as $(\alpha, \alpha, \alpha, \alpha, \alpha)$, $\alpha \in (0, 1)$. MGRIM1 corresponds to Cluster 6. Strategies in this cluster defect in the first period cooperate with a relatively high probability after CC and cooperate with a relatively low probability after all other histories. We parameterize this strategy as $(0, 1-\alpha, \alpha, \alpha, \alpha)$, $\alpha \in (0, 0.5)$. MGRIM2 corresponds to Cluster 7. This strategy is similar to MGRIM1 in that it cooperates with a relatively high probability after CC, and cooperates with a relatively low probability after DD, DC, and CD. However, strategies in this cluster cooperate with higher probability in the first period. We parameterize this strategy as $(0.5, 1-\alpha, \alpha, \alpha, \alpha)$, $\alpha \in (0, 0.5)$.

[22]The fit of the strategy sets presented in Table II could be improved using heterogeneous error terms for each strategy (e.g., Bland (2020)).

[23]An important difference between the mixture model estimation displayed in Table III and the procedure used in Breitmoser (2015) is that we consider the six parameterized strategies and do not carry out the strategy elimination based on the integrated classification likelihood-Bayes information criterion (ICL-BIC).

[24]Figure A-6 in the Appendix presents results of the clustering approach on vectors of cooperation percentages after (CC, CD, DC, DD). The six main clusters, each containing at least 12% of subjects, map to the six strategies used for the estimation. The clustering approach also identifies an additional seventh cluster, but it contains less than 2% of strategies.

TABLE III

MIXTURE MODEL ESTIMATES FOR MEMORY-1 MARKOV STRATEGIES.

(a)

Strategy	Proportion	Parameter
CONST (α,α,α,α)	0.145 (0.042)	0.371 (0.058)
Gen GRIM (1,α,α,α)	0.126 (0.1)	0.243 (0.129)
Gen COOP (1,α,0,0),$\alpha > .5$	0.131 (0.087)	0.5 (0.112)
Gen TFT (1,0,α,0),$\alpha > .5$	0.249 (0.062)	0.964 (0.046)
Gen WSLS (1,0,0,α),$\alpha > .5$	0.0 (0.004)	1.0 (0.183)
SG (1,α,α,0)	0.35 (0.088)	0.173 (0.136)
LL	-6407.535	
BIC	6465.378	

(b)

Strategy	Proportion	Parameter
TFT (1,0,1,0)	0.221 (0.058)	—
GRIM (1,0,0,0)	0.27 (0.067)	—
ALLD (0,0,0,0)	0.069 (0.035)	—
MTFT (1,α,1 $-$ α,0),$\alpha < .5$	0.214 (0.061)	0.309 (0.043)
CONST (α,α,α,α)	0.118 (0.039)	0.397 (0.104)
MGRIM (1 $-$ α,α,α,α),$\alpha < .5$	0.108 (0.046)	0.188 (0.124)
LL	-6234.634	
BIC	6278.017	

Note: For this estimation, we use the choice data from the second locked-response stage. The three columns in each panel display the strategy, the proportion of the population estimated to use that strategy, and the estimated value of the parameter ("–" if the strategy has no parameter). Bootstrapped standard errors are generated using 1000 samples by first randomly drawing groups and then randomly drawing an appropriate number of subjects for each group. Strategies are represented by the four-dimensional vector, which represents the probabilities of cooperation after memory-1 histories CC, CD, DC, and DD. The strategy estimation in panel (a) uses the six strategies from Breitmoser (2015). The strategy estimation in panel (b) uses the six strategies identified from the cluster analysis in Section 3.2. Log-likelihood (LL) and Bayesian Information Criterion (BIC) values are displayed in the final two rows.

quence of play for that subject in supergames 41–60. We then compare this likelihood-based classification to the clustering results presented in Figure 2 (see Figure A-7 in the Appendix). We find that strategies from five different clusters are classified as SG. The cluster with the most strategies classified as SG is the GRIM cluster, yet there are also three other clusters with at least four strategies classified as SG. Strategies classified as SG include those close to GRIM (100,100,0,0,0), TFT (100,100,20,80,0), and ALLD (0,3,6,8,3). This exercise highlights why SG has a high estimate without much empirical evidence. First, SG does not take the first period into account. Therefore, SG can match strategies that cooperate in the first period as well as those that defect in the first period. Second, SG prescribes accurate fixed behavior after the most common histories CC and DD, while providing flexibility after uncommon histories CD and DC.[25] Hence, SG can capture behavior from a variety of strategies that expect to have long periods of CC and DD such as TFT, GRIM, and ALLD, and has the ability to adjust the flexible parameter to fit behavior after rare occurrences of CD and DC.

Another important point regarding mixture model estimation with mixed strategies is that the interpretation of a given strategy heavily depends on the parameter estimates for that strategy. For example, consider the ICL-BIC elimination as done in Breitmoser (2015). Following this procedure for the strategy set from Table III(a), we find that Gen

[25] In supergames 41–60, approximately 30%, 10%, 10%, and 50% of observed play followed each of the CC, CD, DC, and DD memory-1 histories, respectively. The difference in frequencies aligns with the observed composition of strategies. For example, a subject that plays GRIM can only encounter the CD history once within a supergame, but likely will encounter the CC or DD history more often.

COOP and Gen WSLS are eliminated while the other four strategies remain (see Table A-4 in the Appendix). Importantly, removing Gen COOP from the set causes Gen Grim to switch interpretations and change which behavior it is capturing. For example, using the likelihood-based classification above, if all six strategies are used, Gen Grim is estimated to be (1,0.243,0.243,0.243) and captures fairly random strategies in the MTFT cluster. After the elimination, Gen Grim is estimated to be (1,0.025,0.025,0.025) and captures pure and close-to-pure strategies in the GRIM and ALLD clusters (See Table A-4 in the Appendix). Importantly, only 2 strategies are classified as Gen Grim both before and after the elimination. Consequently, the estimates and interpretation of other strategies are affected.

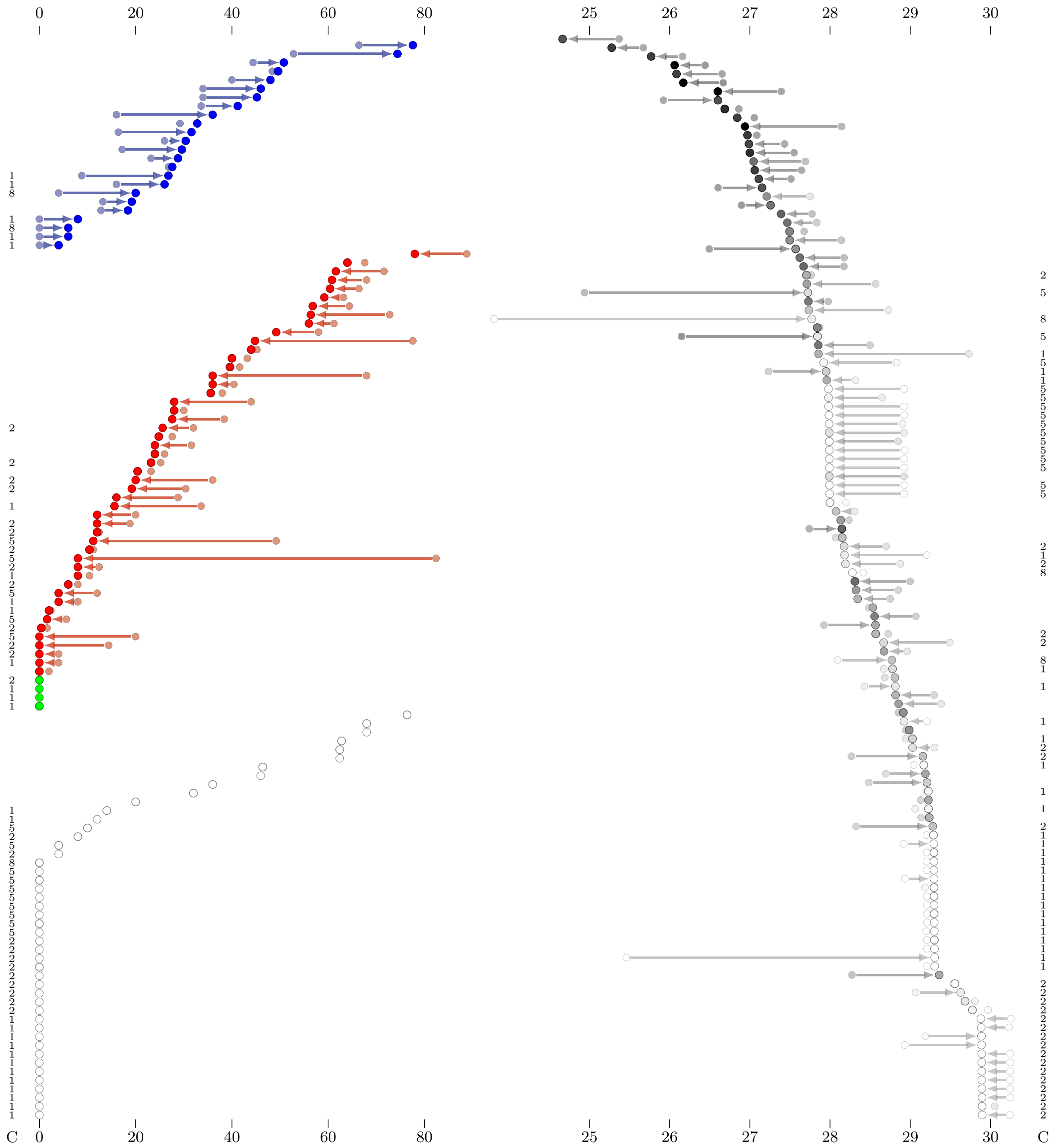
To summarize, in this section, we evaluated the performance of existing approaches using both the actions and strategies from our experiment. We find that the strategy frequency estimation approach of Dal Bo and Frechette (2011) does well in recovering qualitative trends found in our experiment—mainly, the prevalence of GRIM, ALLD, TFT, and TFT-like strategies. We also find that the specification of the details of each strategy and the interaction of strategies in the set is important. In particular, the flexibility of mixed strategies can make them difficult to model and, therefore, strategy sets consisting of mixed strategies can be unreliable in recovering trends in the data while still leading to a good fit.

3.4. *Evolution of Strategies*

Our experiment gives us a unique perspective on the evolution of strategies. We begin our analysis by considering changes that subjects make to their strategies during the experiment. In particular, we examine changes to strategy randomness between the two locked-response stages. In addition, we look at the relative performance of different types of strategies when matched against other strategies in the population. Finally, we use data on strategy changes made throughout our experiment, and simulate a learning process across the population to see where strategies are tending toward.

Panel (a) of Figure 3 presents the strategy-randomness measure ($M1RM$) for each strategy. We use arrows to denote the resulting change in $M1RM$ and sort strategies according to modification type (top: changes resulting in an increase of $M1RM$; middle: changes resulting in a decrease or no change of $M1RM$; bottom: no modification). Panel (b) of the figure presents the performance of strategies against the population of elicited strategies. Performance is calculated as the average payoff per period in a round-robin tournament in which each strategy is matched with each other strategy from the corresponding stage 1000 times. The length of each interaction is determined by the continuation probability from the experiment, $\delta = 0.95$, with the same 1000 draws for each pair of strategies. We use arrows to identify changes across the two locked-response stages.

Using the data presented in Figure 3, we identify three key observations regarding the evolution of strategies. First, subjects were less likely to make changes to pure strategies and more likely to make changes to mixed strategies. This can be seen in panel (a), where only 8 out of 38 subjects with $M1RM = 0$ in LR1 changed their strategy, whereas 67 out of 84 subjects with $M1RM > 0$ in LR1 changed their strategy (p-value < 0.01, using Fisher's exact test). Second, of the strategies that were changed, more changed to lower $M1RM$ (48) than higher $M1RM$ (24, p-value < 0.01). Finally, panel (b) of the figure shows that strategies with lower $M1RM$ perform better. Notably, the two clusters with the best per-



(a) **Strategy Randomness (M1RM)** (b) **Strategy Performance**

FIGURE 3.—Strategy Randomness and Performance. *Note*: **(a)** Strategies are sorted by modification type and $M1RM$. Modification types are based on strategy changes between LR1 and LR2. Solid circles denote subjects that changed their strategy that led to an increase, decrease, or no change in $M1RM$. Arrows denote the magnitude of the change in $M1RM$ from LR1 (semitransparent) to LR2 (opaque). Empty circles denote no change to strategy. **(b)** Strategies are sorted by performance. Performance is calculated as the average payoff per period in a round-robin tournament in which each strategy is matched with each other strategy from the corresponding stage 1,000 times. The length of each interaction is determined by the continuation probability from the experiment, $\delta = 0.95$, with the same 1,000 draws for each pair of strategies. $M1RM$ of each strategy is identified by a shade of gray (with solid black denoting a strategy with the highest $M1RM$, and solid white denoting a strategy with $M1RM$). Corresponding measures for LR1 are semitransparent. Change in the measure is identified by an arrow. "C" identifies strategies that fall into pure-strategy clusters: 1–TFT, 2–GRIM, 5–ALLD, and 8–DTFT.

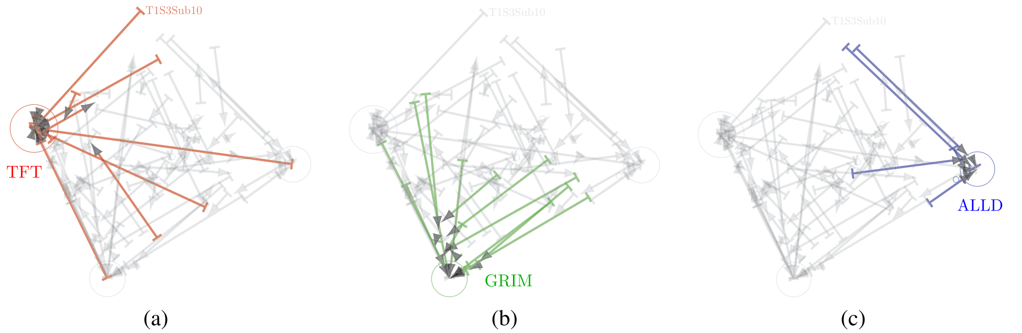


FIGURE 4.—Evolution of Strategies within Experiment. *Note*: We carry out a principal component analysis to reduce the dimensionality of the data from 5D to 2D. Circles represent subjects that did not change their strategy from the beginning to the end. The size of the circle represents the number of subjects. For example, subject T1S3Sub10 constructed strategy (0,100,100,100,100) at the beginning of supergame 11 and changed it to strategy (100,100,0,100,0) by supergame 41.

forming strategies were Clusters 2 and 1, corresponding to GRIM and TFT.[26] These three observations provide strong evidence that subjects were still learning and more subjects may have started playing well-performing pure strategies if given more time to learn.

Figure 4 presents the change in strategies observed in our experiment projected on a two-dimensional space using principal component analysis. The projected vector space has the leftmost, lowest, and rightmost points corresponding to TFT, GRIM, and ALLD, respectively. Each vector starts at a given subject's strategy observed at the beginning of supergame 11 and ends at that subject's strategy at the beginning of supergame 41. For example, subject *T1S3Sub10* constructed strategy (0,100,100,100,100) at the beginning of supergame 11 and changed it to strategy (100,100,0,100,0) by supergame 41. Panels (a), (b), and (c) highlight the vectors for all subjects whose strategies at the beginning of supergame 41 were assigned to Cluster 1 (TFT), Cluster 2 (GRIM), and Cluster 5 (ALLD), respectively.[27] This figure suggests that a number of subjects have learned to play these strategies during our experiment.

To provide evidence on where the long-run learning dynamics might go, we use Markov Chain simulations. These simulations allow us to extrapolate trends in strategy changes observed in Figure 4, by assuming a common learning process for all agents. In particular, the process is determined by the set of states, $\mathcal{S}$, the one-step transition probability matrix, $\mathcal{P}$, and the starting state, s_0. The set of states, $\mathcal{S}$, consists of all 101^5 possible memory-1 mixed strategies available in our experiment. The one-step transition probability matrix $\mathcal{P}$ consists of entries $\mathcal{P}(s, s')$ that correspond to the probability that strategy s evolves into strategy s' in the next step.[28] The starting state, s_0, is a strategy observed at the end of our experiment. To obtain the long-run distribution of strategies, we conduct Markov

[26]In the Appendix, we provide further investigation of strategy performance. Specifically, Figure A-9 presents evidence on the variability of the performance for each strategy. Notably, ALLD yields large variability in performance and a strong negative relationship with the realized supergame length, whereas TFT and GRIM generate small variability in performance and a weak positive relationship with realized supergame length.

[27]In Figure A-10 of the Appendix, we provide the plot for all of the clusters.

[28]We construct $\mathcal{P}$ based on the strategy changes observed in the experiment as follows. Let $\mathcal{I}$ be the set of 124 subjects. For each strategy s, we find the set of subjects, $c(s) \subseteq \mathcal{I}$, who played strategies most similar to s in supergame 11, and a subset of these subjects, $d(s, s') \subseteq c(s)$, who then played strategy s' in supergame

Chain simulations until convergence, by starting the process with each one of the 124 strategies from the second locked-response stage (presented in Figure 2) and repeating the process 100 times.[29] We then classify converged strategies based on the clusters from the second locked-response stage. We find that the three pure strategies—TFT, GRIM, and ALLD—are the main attractors in the learning space (see Figure A-10 in the Appendix). Specifically, 31% of converged strategies are classified as TFT, 30% are classified as GRIM, 29% are classified as behaviorally equivalent to ALLD, and the remaining 10% are classified as CONST.[30]

To summarize, we find that subjects' strategies are becoming less random across supergames during our experiment. We also find that pure strategies generally perform better than mixed strategies. In addition, we find that a number of subjects learned to play TFT, GRIM, and ALLD during our experiment. Finally, a simple learning process based on strategy changes in our experiment converges to strategies that are close to pure strategies over 90% of the time. Overall, these trends suggest that the three commonly studied strategies (TFT, GRIM, and ALLD) may be even more prevalent if the subjects are given more time to learn.

4. ROBUSTNESS TREATMENTS

We made several important choices regarding the experimental design of the main treatment presented above. First, we restricted subjects to memory-1 rules. Eliciting and analyzing rules longer than memory-1 adds significant complexity to an already complex environment for both the researcher and the experimental subjects. Because the main goal of this paper was to understand the use of mixed strategies in repeated games as motivated by the literature on memory-1 mixed strategies, we decided to try to avoid as much complexity as possible and focus solely on memory-1 strategies. Second, when creating a rule, subjects used a slider to select the output by choosing any integer between 0 and 100. Using a slider meant that the only way they could choose a nonrandom output was to select either 0 or 100, whereas the other 99 numbers all led to a random output. We made this choice for simplicity and because we wanted to give commonly studied mixed strategies, such as SG, a fair shot at emerging. Finally, we only used one probability of continuation ($\delta = 0.95$), which was the only probability of continuation that would allow us to compare results with previous experiments that had been run with a similar interface (Romero and Rosokha (2018)) as well as those run with a different interface (Dal Bó and Fréchette (2019)).

The design of our main treatment raises a few potential concerns. First, the restriction to memory-1 strategies could lead some subjects to use mixed strategies as a proxy for longer pure strategies (e.g., lenient grim or tit-for-two-tats). In this case, the results presented above may overstate the amount of mixed strategies that subjects are using.

41. Formally, $c(s) = \arg\min_{i\in\mathcal{I}} \|s - s_i^{11}\|$ and $d(s, s') = \{i | i \in c(s), s_i^{41} = s'\}$, where $s_i^{11}, s_i^{41} \in \mathcal{S}$ are strategies for subject $i \in \mathcal{I}$ in supergames 11 and 41, respectively. Finally, we define $\mathcal{P}(s, s') = \frac{|d(s,s')|}{|c(s)|}$.

[29]Sometimes the process cycles through a deterministic repetitive sequence of strategies. In this case, we say the process converges to the average of all strategies in the repetitive sequence.

[30]An alternative approach to identify long-run strategy distributions could be to use an adaptive learning model over the set of repeated-game strategies such as in Hanaki, Sethi, Erev, and Peterhansl (2005), Ioannou and Romero (2014), Romero and Rosokha (2019) or estimate individual-level learning parameters as done in Dal Bó and Fréchette (2011) and Embrey, Fréchette, and Yuksel (2018). However, given the complexity of strategies considered in this paper, extending the existing learning models is not trivial.

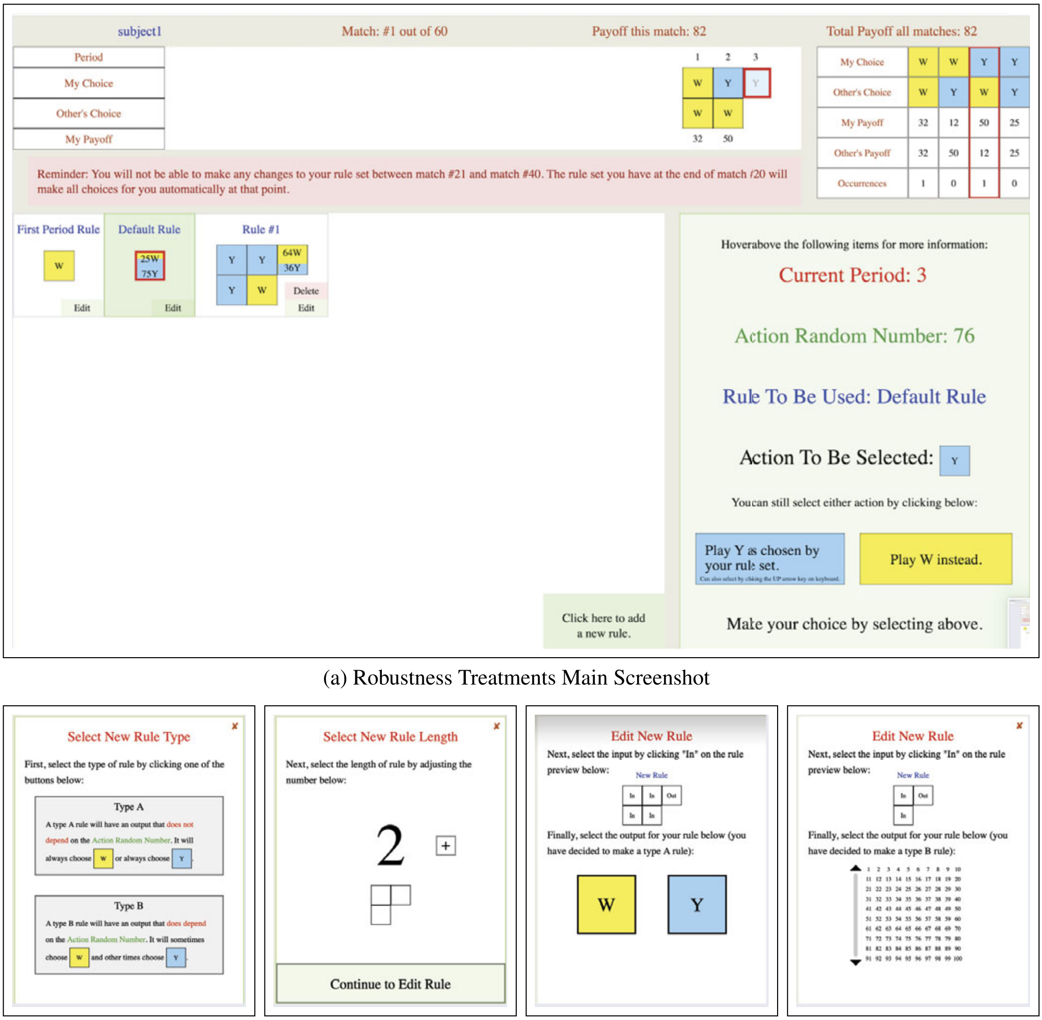
Second, the large majority of options (99 out of 101) of the output-selection slider corresponded to mixed outputs. For this reason, if a subject were to select their output randomly, the probability of them selecting a random output would be very close to 1 ($\frac{99}{101}$). This is another potential reason that the amount of mixed strategies may be overstated. Finally, we only used one continuation probability. Just using data from the main experiment, how changes to the continuation probability would affect the amount of mixed strategies is not clear. A different continuation probability would also be useful as another basis of comparison for cooperation levels and strategies played by subjects.

Given these concerns, we ran additional robustness treatments with two modifications. First, in the robustness treatments, subjects were not restricted to memory-1 rules. Instead, they could specify a rule that can condition on up to eight periods in the past. Second, they were required to go through a multistage procedure to create a new rule. The procedure (discussed below) made pure strategies more salient. These two changes allow us to address the first two concerns listed above. We address the third concern by using two different probabilities of continuation in the robustness treatments: one with continuation probability $\delta = 0.95$ (referred to as M2 + D95) and the other with continuation probability $\delta = 0.75$ (referred to as M2 + D75). Treatment M2 + D95 uses the same continuation probability as the main treatment and allows us to see what impact our design choices had on results. Treatment M2 + 75 gives us results for a new continuation probability.

A list of design changes for the robustness treatments are as follows:

1. To allow subjects to use rules longer than memory-1, we needed to redesign the interface. A screenshot of the updated experimental interface can be found in Figure 5. The major change is that the rule-set area has been expanded to allow for more rules and longer rules. In addition, it can now scroll if the user adds more rules than can fit on the screen.
2. To reduce experimenter demand of mixed strategies, the subjects now have to go through a multistage procedure to add a new rule. First, the subjects are asked if they want the output of the rule to be pure (referred to as a "Type A" rule in the experiment) or if they want the output of the rule to be mixed (referred to as a "Type B" rule in the experiment). Second, the subjects choose the length of the rule (any number between 1 and 8, which represents how many of the previous periods this rule depends on). Finally, the subjects choose the inputs and outputs. When selecting the output, if they choose a Type A rule they are given two buttons (W or Y), and if they choose a Type B rule they are given the slider, where they can choose an integer between 0 and 100. Note they can still choose a pure output. Screenshots of the rule constructor for this multistage procedure can be seen in Figure 5 panels (b) through (e).
3. Given the expanded set of possible rules, and the new procedure for adding rules, we also had to update the experimental quiz. Changes to the quiz include new screens explaining rule length, which rule will be chosen (longest rule that fits the history), and some demos for how to add rules of different types. The quiz still had the same number of questions and the same criteria for passing the quiz, and subjects were given the same amount of time.[31]
4. For the M2 + 75 treatments, subjects faced the same number of supergames but fewer periods (60*4 in expectation) than in the main experiments and M2 + 95 treatments (60*20 in expectation).

[31]A complete set of screenshots of the updated quiz from the robustness treatments can be found at http://jnromero.com/research/mixedStrategyChoice.



(a) Robustness Treatments Main Screenshot

(b) Rule Type Selection

(c) Rule Length Selection

(d) Deterministic Output Selection

(e) Mixed Output Selection

FIGURE 5.—Screenshot of the Experimental Interface for the Robustness Treatments. *Note*: The screenshots shows (a) the main interface for the robustness treatments, (b) the rule type selector (deterministic Type A or mixed Type B), (c) the rule length selector, (d) the deterministic output selector, and (e) the mixed output selector.

The two robustness treatments consisted of 12 sessions run at the Vernon Smith Experimental Economics Laboratory at Purdue University in March 2020 and March 2021. Details of each session are provided in Table A-3 in the Appendix. As in the main treatment, each session consisted of instructions, an incentivized quiz to ensure that subjects understood the instructions, and 60 supergames. The M2 + D95 treatment used the same exchange rate as the main experiment (2000 ECUs equals one US dollar). Since the M2 + D75 had one-fifth as many periods as the M2 + D95, in expectation, we set the exchange rate in the M2 + D75 treatment to 400 ECUs equals one US dollar.

It is important to note that the analysis of the rule sets in the robustness treatments is difficult due to the complexity of the strategy space.[32] Difficulties include the fact the two strategies can be very similar, for example, TFT and a strategy that is like TFT after every history except that after eight periods of alternations between CD and DC, it plays D. In addition, multiple rule combinations can lead to the same strategy (TFT can be represented with {default rule D, $FP \to C$, $CC \to C$, $DC \to C$} or {default rule C, $FP \to C$, $CD \to D$, $DD \to D$}). Therefore, rather than focusing on the strategies, we focus on simulated play of the strategies using the method from Romero and Rosokha (2018). More precisely, we fix a set of 500 predetermined action sequences of random lengths based on the continuation probability of $\delta = 0.95$. Then we have every strategy play against these sequences and analyze and compare strategies based on the *simulated play*. Focusing on the simulated play of strategies alleviates the above difficulties because similar strategies will have similar simulated play regardless of the rule-set composition.[33]

To provide a comparison between the memory-1 strategies in the main treatment and the memory-8 strategies in the robustness treatments, we construct an *empirical memory-1 strategy* by taking the simulated play and calculating the frequencies of cooperation in the first period and after each of the four memory-1 histories. For example, consider a subject that has a rule set {default rule D, $FP \to D$, $DCDC \to C$} presented in Figure 6(a). This rule set plays like ALLD after every history except that after two consecutive periods of cooperation by the other player, it switches to cooperation for one period. The empirical memory-1 frequencies for this strategy would be 0% cooperation in the first period and after CC, CD, and DD; and roughly 31% cooperation after DC. We organize these empirical memory-1 strategies using the same clustering approach as in Section 3.2. The results are presented in Figure 7. We find that consistent with the main treatment, the two largest clusters in the M2 + 95 treatment contain strategies that are similar to TFT and GRIM.[34] Also, not surprisingly, we find that in the M2 + 75 treatment, the two largest clusters contain strategies that are similar to ALLD and GRIM.

In addition to empirical memory-1 frequencies, we compute three measures that capture the intent and prevalence of randomizing behavior. First, we define a mixed strategy as a strategy that randomizes during at least one period of the simulated play. For example, the strategy presented in Figure 6(a) is a pure strategy, while the strategy presented in Figure 6(b) is a mixed strategy. Second, we calculate the M1RM for the empirical memory-1 strategy. For example, a pure strategy presented in Figure 6(a) would appear as randomizing 31% of the time after DC if viewed from the perspective of memory-1 cooperation frequency. Third, we define a strategy's action randomness measure (referred to as *ARM*) as the average randomness measure of rule outputs used to make choices

[32]Each strategy can be represented with a vector with one entry for the first-period rule, and an entry for each possible rule up to memory-8. Therefore, the set of possible strategies in this experiment is $101^{\sum_{k=0}^{8} 2^{2k}} = 10^{\log(101)87{,}381} \approx 10^{175{,}140}$. Notice that while memory-8 rules will always be used after period 8, shorter rules may still be used before period 8.

[33]The predetermined sequences are generated using a Markov process that covers a variety of behaviors (see Online Appendix for more details). We use the same set of predetermined action sequences to compare strategy sets across treatments because otherwise identical strategies may appear to be different. In addition, to keep the outcome of randomization as similar as possible, we used the same sequence of action random numbers for all strategies.

[34]Note that the clusters (and cluster exemplars) generated by the affinity propagation algorithm are endogenous. For comparison, in Figure OA4 of the Online Appendix, we provide the empirical memory-1 strategies organized using the exemplars from Figure 2. One observation is that the composition of cluster 4 in Figure OA4 is substantially different than in Figure 2. In particular, we find fewer uniformly randomizing strategies in the robustness treatments.

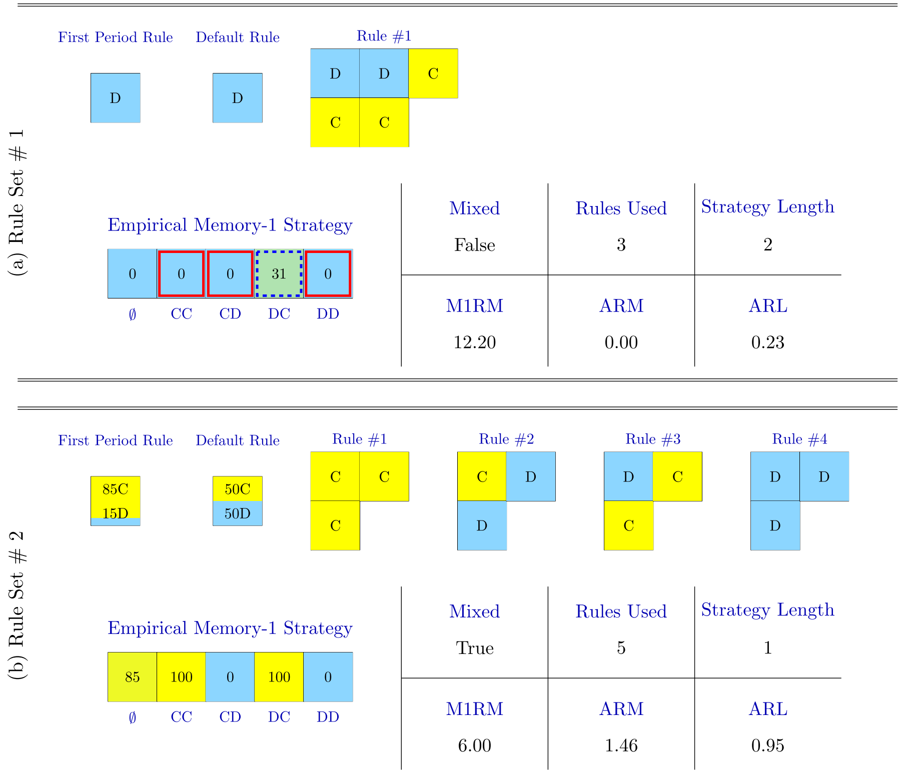


FIGURE 6.—Example Rule Sets and Strategy Statistics. *Note*: Rule sets are displayed in the top row of each panel. All strategy statistics are calculated using simulated play (see Online Appendix for details). Empirical memory-1 strategies display cooperation frequency after each of the five histories ($\emptyset$, CC, CD, DC, DD). Solid squares denote a situation in which the default rule determines the percentage of cooperation after that history. Dashed squares denote a situation in which a memory-2 or greater rule was used to make a choice at least once after the specified history. Both rule sets were constructed by subjects in the second locked-response stage of the M2+D95 treatment. Rule Set #1 is the exemplar of the ALLD cluster, and rule set #2 is the exemplar of the TFT cluster.

across all periods of simulated play.[35] This measure tells how random the choices from a subject's rule set are due to rules with mixed outputs. For example, the ARM of the pure strategy in Figure 6(a) is 0, because all choices are made by rules with deterministic output. Whereas the ARM of the mixed strategy in Figure 6(b) is close to 1.50 because the randomness measure of actions in the first period is 30, and in the simulations, the first period occurs approximately 1/20 of the time because the interaction length is determined using continuation probability $\delta = 0.95$.

We present a summary of the three measures in Table IV. In particular, the table displays results obtained from rule sets of the main treatment (labeled M1D95), the first robustness treatment (M2 + D95), and data from Romero and Rosokha (2018) (labeled RR2018), all of which implemented continuation probability $\delta = 0.95$. In the fourth column, we include results from the second robustness treatment (M2 + D75), which con-

[35] $ARM = \frac{1}{T}\sum_{t=1}^{T} RM_t$, where RM_t is the randomness measure of the output of the rule used to make a choice in period t, and T is the number of interactions in the simulated play. Recall that the randomness measure is defined as $RM = 2 * \min\{100 - p, p\}$, where p is the cooperation percentage.

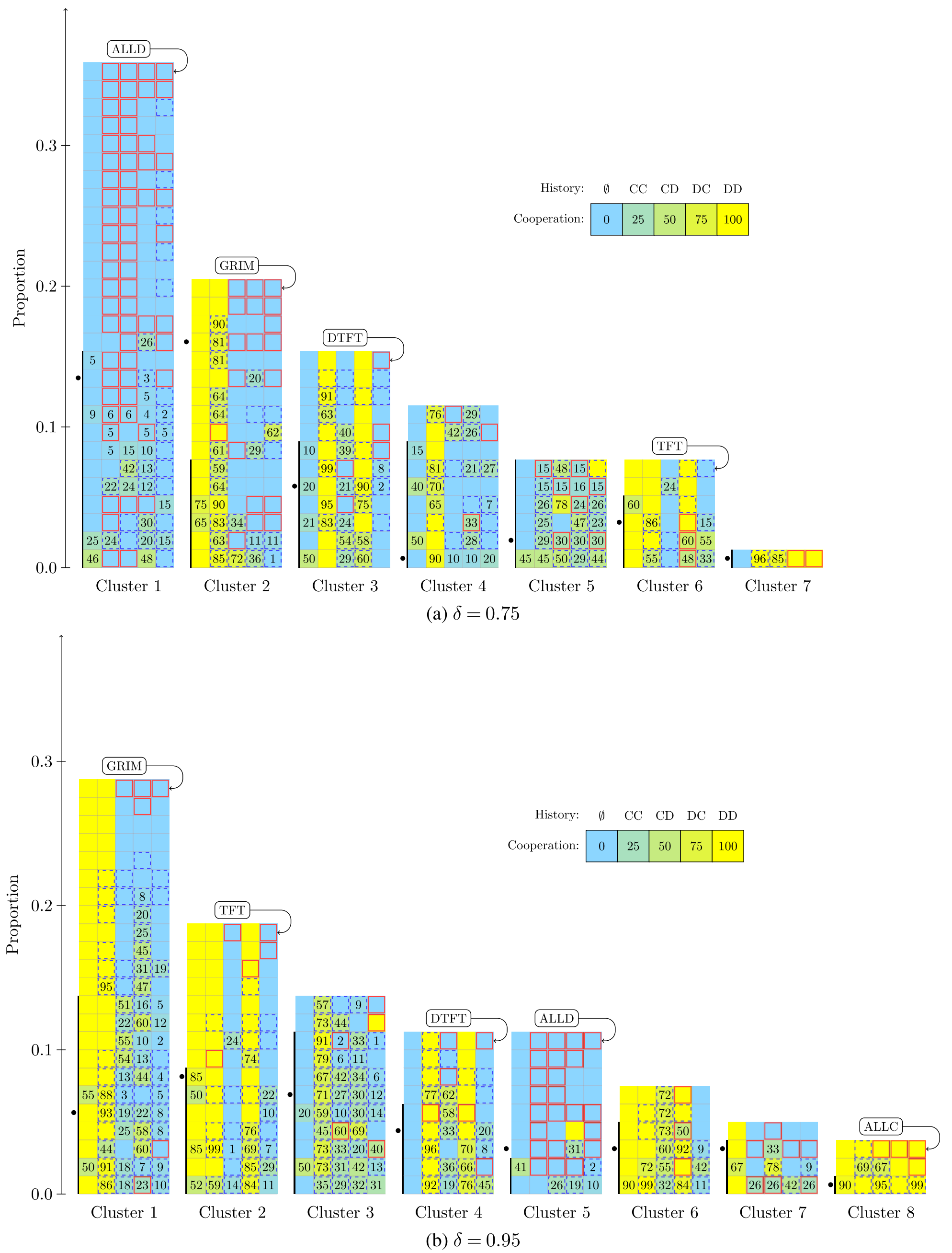


FIGURE 7.—Empirical Memory-1 Strategies in Robustness Treatments. *Note*: Empirical memory-1 strategies are presented as vectors of cooperation frequency after each of the five histories ($\emptyset$, CC, CD, DC, DD). Displayed strategies are generated using rule sets from the second locked-response stage. Strategies are grouped using the affinity propagation clustering algorithm. Strategies in a given cluster are sorted according to ARM. A black bar on the left side of a strategy indicates $ARM > 0$. A bullet on the left side indicates the cluster exemplar. Solid squares denote a situation in which the default rule determines the percentage of cooperation after that history. Dashed squares denote a situation in which a memory-2 or greater rule was used to make a choice at least once after the specified history. The pure strategies TFT, GRIM, ALLD, DTFT, and ALLC are labeled.

sists of 60 supergames with continuation probability $\delta = 0.75$. All analyses are carried out using rule sets from supergame 41 of the corresponding treatment. Table IV shows that the proportion of subjects that construct rule sets that use random outputs has marginally decreased relative to the main treatment (0.669 vs. 0.525, p-value of 0.11; 0.669 vs. 0.551, p-value of 0.14). In addition, while the extent of randomness is not significantly different according to M1RM (21.523 vs. 17.899, p-value of 0.30; 21.523 vs. 15.325, p-value of 0.13), it significantly decreases according to ARM (20.664 vs. 10.066, p-value of 0.01; 20.664 vs. 10.237, p-value of 0.02). This finding suggests that focusing on cooperation frequencies after memory-1 histories (M1RM) may lead to the overestimation in the extent of randomizing behavior during game-play (ARM). Combined, our results suggest that while about half of the subjects are still using mixed strategies in the robustness treatments, the amount of randomness in the play has gone down.

Table IV also includes descriptive measures of rule sets. Specifically, we define a strategy's length as the length of the longest rule used to make a choice during simulated play. For example, the length of the strategy presented in Figure 6(a) is 2, and the length of the strategy presented in Figure 6(b) is 1. We define rules used as the total number of rules used to make a choice during simulated play. For example, the strategy set presented in Figure 6(b) contains five rules because the default rule is never used. We define action rule length (*ARL*) as the average length of the rule that makes a choice during simulated play. For example, *ARL* of the strategy presented in Figure 6(b) is 0.95 because for an interaction of 20 periods, the first-period rule (length 0) is used once, and a rule with length 1 is used in every other period. We find that the length of the longest rule (row 4), the number of rules used (row 5), and the average length of a rule that make a choice during simulated play (row 6) are all increasing (p-values < 0.01 for all comparisons except ARL M1D95 vs. M2 + D75, which has a p-value of 0.80). Not surprisingly, this confirms that subjects have more rules and longer rules in their sets in the robustness treatments.

The last three rows of Table IV show how close the elicited strategies are to three commonly observed strategies: ALLD, TFT, and GRIM. The seventh row of the table shows the percentage of strategies that exactly match one of those three strategies. Next, in row eight, we present the proportion of strategies in the same cluster as one of the three pure strategies. Finally, the last row shows the proportion of ALLD, TFT, and GRIM estimated using the SFEM procedure on the actions from the experiment using the set of 20 pure strategies proposed in Fudenberg, Rand, and Dreber (2012). Notably, for the robustness treatments, the proportions of subjects classified to the ALLD, TFT, and GRIM clusters are close to the results of the SFEM. In addition, the fact that 20% of subjects are able to construct one of the three common strategies exactly, despite the complexity of the experimental interface and the fact that the set of possible strategies contains $101^{\sum_{k=0}^{8} 2^{2k}}$ strategies indicates the focal nature of these strategies. These results suggest that a majority of the strategies used in the robustness treatments are close to the three commonly studied pure strategies, even if they do not match exactly.

Several results from the robustness treatments are in line with previous papers that study strategies for the same stage-game payoffs. The first three results pertain to changes in strategies in response to changes in δ and are consistent with Dal Bó and Fréchette (2019). First, as δ decreases, the prevalence of the ALLD strategy increases significantly (0.11% vs. 0.35%). Second, the ratio of TFT to (TFT + GRIM) decreases as δ decreases (39.4 vs. 27.3). Third, the use of memory-1 strategies increases as δ decreases (ARL of 0.987 vs. 0.800). The last result pertains to potential experimenter demand effects for randomization and longer strategies induced by our interface and design. In particular, using data from supergames 11 through 20 of Dal Bó and Fréchette (2011) and Dal Bó and Fréchette (2019), we calculate the proportion of supergames that can be perfectly accounted for by the five most common memory-1 strategies —ALLC, ALLD, TFT, GRIM, and DTFT—and compare it to our M2 + D75 treatment. We find that 84.0% of the supergames in Dal Bó and Fréchette (2011), 91.8% of supergames in Dal Bó and Fréchette (2019), and 87.8% of supergames in our M2 + D75 treatment can be perfectly accounted for by these five strategies, suggesting that the experimenter demand effects from our design are minimal.

To summarize, Table IV shows that a majority of subjects in the robustness treatments are still using mixed strategies (row 1). The table also shows that subjects are constructing rule sets with longer (rows 4 and 6) and fewer random rules (row 3). This finding is not surprising given the design of the robustness treatments, which allows subjects to construct longer rules and deemphasizes mixed rules. Finally, the table shows that a majority of subjects are playing strategies close to the three focal pure strategies TFT, GRIM, and ALLD (according to clusters in row 8 and SFEM in row 9). Altogether, these results suggest that while subjects shifted to longer and less random strategies in the robustness treatments, the primary conclusions from the main treatments still hold.[36]

5. DISCUSSION

The goal of this paper is to address the recent debate regarding whether subjects use mixed strategies in the indefinitely repeated prisoner's dilemma. In particular, we conduct both an experimental and econometric investigation into the types of strategies that subjects play. Our experiment is the first (to our knowledge) to elicit mixed repeated-game strategies. The experimental design allows us to directly observe the strategies that subjects use, thereby shedding light on whether subjects use mixed repeated-game strategies. We use the elicited strategies to provide an empirically-relevant foundation for analyzing commonly used mixture model estimation procedures.

The first main takeaway is that a majority (67%) of subjects played mixed strategies in our experiment. In particular, we find a large proportion of subjects used mixed strategies that we refer to as the mixed-tit-for-tat class of strategies. These strategies cooperate with high probability after mutual cooperation, defect with high probability after mutual defection, and randomize otherwise. This class of strategies, to a large degree, is similar to SG (proposed in Breitmoser (2015)) because it has the same trends after CC and DD, and also randomizes after CD and DC. However, SG is defined as having an equal probability of cooperation after CD and DC, and our data suggest that there is a consistent difference in these probabilities in favor of more cooperation after DC than after CD.

The second main takeaway is that three commonly studied pure strategies (TFT, GRIM, and ALLD) play an important role in our experiment. First, a large proportion of subjects (32%) constructed these three strategies exactly. This result is particularly striking in the context of our experiment because these three strategies emerged even when the strategy space contained 101^5 strategies. Second, the cluster analysis shows that many of the mixed

[36]Recall that all results in this section are based on the simulated play, because it provides the cleanest comparison of constructed strategies across treatments. As a robustness check, we remove subjects who played ALLD from the data and present the results in Table OA2 in the Online Appendix. As a further robustness check, we provide the corresponding results (for all subjects) based on the actual play during the second locked-response stage of the experiment in Table OA3 in the Online Appendix. Finally, the proportion mixed, M1RM of the memory-1 part of the strategy, the strategy length, and the number of rules in the set can be calculated from the rule sets directly. We present them in Table OA4 in the Online Appendix. The main conclusions still hold across these robustness checks.

TABLE IV

STRATEGY COMPARISON ACROSS TREATMENTS.

	M1D95	M2+D95	RR2018	M2+D75
Proportion Mixed	0.669	0.525	0.000	0.551
	(0.057)	(0.088)	(0.000)	(0.085)
Memory-1 Randomness (M1RM)	21.523	17.899	9.646	15.325
	(3.002)	(2.331)	(1.776)	(3.185)
Action Randomness (ARM)	20.664	10.066	0.000	10.237
	(3.041)	(2.145)	(0.000)	(3.146)
Strategy Length	0.927	2.775	4.744	2.744
	(0.033)	(0.221)	(0.639)	(0.286)
Rules Used	4.581	7.600	8.378	5.974
	(0.104)	(0.553)	(0.576)	(0.495)
Action Rule Length (ARL)	0.782	0.987	1.112	0.800
	(0.032)	(0.074)	(0.060)	(0.084)
Proportion of ALLD, TFT, GRIM	0.323	0.212	0.293	0.231
	(0.055)	(0.085)	(0.063)	(0.065)
Close to ALLD, TFT, GRIM (In Cluster)	0.540	0.588	0.732	0.641
	(0.065)	(0.091)	(0.072)	(0.100)
SFEM proportion of ALLD, TFT, GRIM	0.840	0.580	0.600	0.680

Note: The columns of the table correspond to the main treatment (M1D95), the robustness treatments with $\delta = 0.95$ (M2 + D95), the treatments from Romero and Rosokha (2018), and the second robustness treatment with $\delta = 0.75$ (M2 + D75). All statistics are calculated using strategies from the second locked-response stage. Bootstrapped standard errors (in parentheses) are calculated by drawing 1000 random samples: first, randomly draw the appropriate number of groups with replacement; then, for each group, randomly draw the appropriate number of subjects with replacement; and finally, calculate the desired statistic.

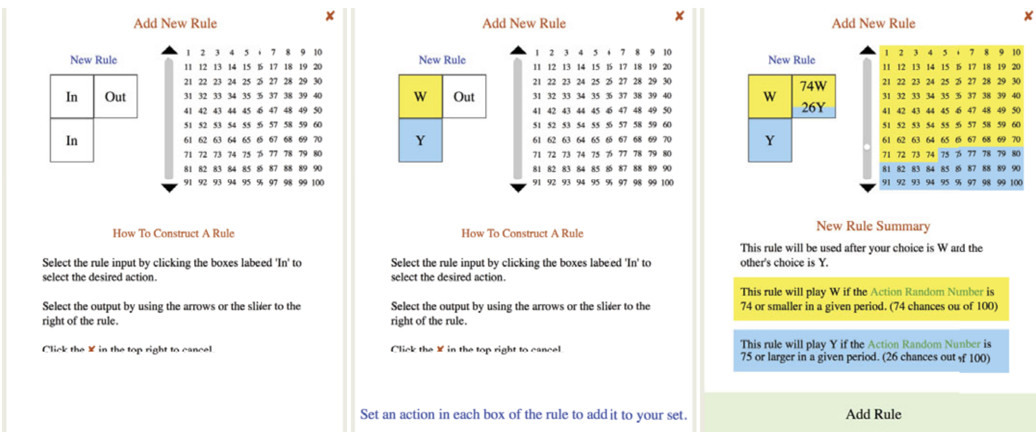
strategies are close to these three pure strategies as well. In total, 54% of all subjects' strategies fall into the three clusters corresponding to these three pure strategies. Finally, we find evidence that subjects' strategies are becoming less mixed over the course of the experiment. Taking this observation one step further, we implement a model of learning that suggests that 90% of the strategies would converge to the three commonly studied pure strategies if subjects had more time to learn.

The third main takeaway is that mixture model estimation is sensitive to the set of strategies that are being used. This is particularly relevant when considering mixed strategies because the researcher needs to specify how randomization is modeled within a strategy. We use the set of strategies observed in our experiment as a benchmark to compare with previously considered strategy sets using observed actions from our data. We find that some strategy sets lead to estimates that do not match the underlying distribution of strategies.

In addition to our main experiment, we ran two robustness treatments that make pure strategies less salient and allow for a larger set of strategies. The results of the robustness treatments suggest that some subjects appear to substitute away from memory-1 mixed strategies toward pure strategies that condition on longer histories. Despite this, the primary conclusions from our main experiment still hold: a majority of subjects use mixed strategies; many subjects exactly construct ALLD, TFT, or GRIM; and many of the mixed strategies are close to those three pure strategies.

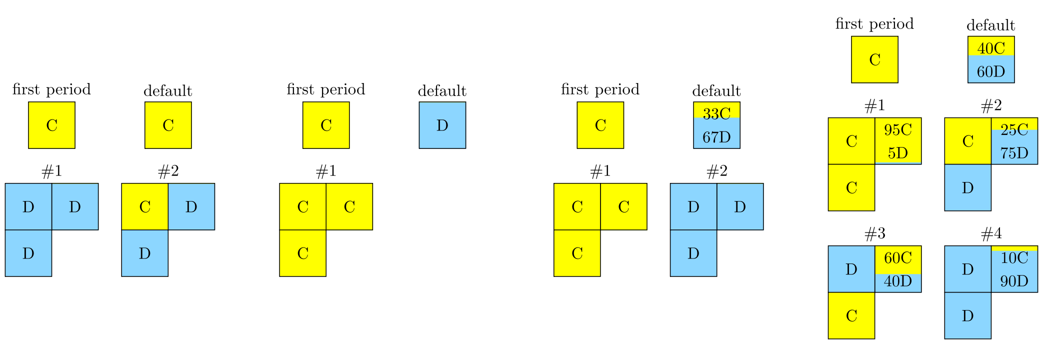
Several avenues for future research are promising. First, it would be interesting to test whether the large proportion of pure strategies and the identified type of mixed strategies are observed in treatments with different parameters. In particular, parameters that are studied in Breitmoser (2015) may show more evidence of SG because the SG equilibrium described in that paper is fairly close to GRIM for our experimental parameters. Second, it would be interesting to focus on the learning process with mixed strategies in more detail. In particular, building a reinforcement learning or a belief learning model that may explain behavior in our experiment would be of great interest. Finally, future research can use our elicitation approach for evaluating theoretical refinements in repeated-game strategies and different strategy-estimation procedures.

APPENDIX A: ADDITIONAL TABLES AND FIGURES



(a) (b) (c)

FIGURE A-1.—Rule Constructor Screen-shots. *Note*: (a) Before any selection has been made; (b) After the input has been set, but before the output has been set; (c) After both input and output have been set. Subjects could make selections regarding inputs and output in any order they choose.



(a) TFT (b) GT (c) SG (d) mixed-TFT

FIGURE A-2.—Examples of Rule Sets. *Note*: (a) TFT—Tit-for-Tat; (b) GT—Grim Trigger; (c) SG—Semi–Grim; (d) mixed-TFT—mixed tit-for-tat. Note that there are multiple ways to construct the same strategy. All rule-sets must always have a first-period rule and a default rule. If the rule set has all four rules corresponding to the four possible memory-1 histories (as in the mixed-TFT rule set displayed in panel (d)), then the default rule will never make a choice.

TABLE A-1

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 ($\delta = 0.95$):	39	34	16	2	34	1	47	21	16	4	4	17	23	9	30	27	6	14	8	33	18	3	38	39	13	5	45	25	8	22
Realization #2 ($\delta = 0.95$):	19	25	7	9	2	7	4	54	30	19	24	16	3	15	39	62	14	12	13	9	18	47	11	51	40	4	31	3	4	8
Realization #3 ($\delta = 0.95$):	9	33	12	7	1	76	27	39	6	29	77	34	1	38	25	12	1	12	26	8	3	32	29	15	1	27	31	4	9	58
Realization #4 ($\delta = 0.95$):	42	16	21	25	8	15	90	10	1	19	2	6	20	8	9	15	18	3	2	89	9	29	1	29	23	19	19	8	1	32
Realization #5 ($\delta = 0.75$):	1	2	2	5	5	1	14	3	4	1	2	1	3	1	2	1	1	5	4	1	4	1	4	6	7	1	15	5	1	9
Realization #6 ($\delta = 0.75$):	2	3	2	2	5	11	8	1	3	7	1	11	1	5	1	5	4	6	3	4	2	1	6	1	2	3	3	6	1	1
Realization #7 ($\delta = 0.75$):	5	5	7	1	1	6	6	2	6	6	4	6	2	2	1	1	2	2	1	2	3	5	1	2	4	1	2	1	2	15
Realization #8 ($\delta = 0.75$):	13	6	4	10	4	10	1	4	2	6	7	8	6	1	1	1	1	3	10	10	3	2	2	9	4	8	11	8	8	3

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 ($\delta = 0.95$):	2	22	34	18	24	22	9	6	1	22	13	5	61	7	6	19	5	25	2	1	51	94	12	19	11	32	12	8	6	3
Realization #2 ($\delta = 0.95$):	9	31	22	3	18	34	20	18	43	14	17	2	26	13	23	4	29	7	17	3	4	6	13	28	2	27	9	14	15	79
Realization #3 ($\delta = 0.95$):	49	2	6	4	22	5	50	26	7	14	3	2	34	1	7	2	2	14	56	11	29	32	1	3	7	2	25	11	14	23
Realization #4 ($\delta = 0.95$):	11	45	67	43	16	28	2	40	20	15	6	8	40	38	2	31	28	13	34	57	8	9	108	17	4	15	9	27	23	20
Realization #5 ($\delta = 0.75$):	1	4	6	10	2	3	7	6	6	2	1	1	2	10	11	2	10	10	1	1	1	2	1	2	3	2	13	17	5	3
Realization #6 ($\delta = 0.75$):	1	2	1	7	1	4	11	2	1	3	2	4	1	10	3	14	5	10	1	3	17	5	3	7	4	1	2	3	9	3
Realization #7 ($\delta = 0.75$):	1	20	2	4	1	2	2	2	4	6	1	4	13	5	3	1	1	2	4	6	7	1	3	6	5	3	2	1	6	7
Realization #8 ($\delta = 0.75$):	8	7	3	3	2	1	1	4	2	4	2	16	6	6	1	3	3	1	1	3	1	9	7	1	6	4	1	2	5	1

TABLE A-2

SESSION SUMMARY FOR MAIN TREATMENT.

Date	Session #	Group	# Subj	Realization	Avg Pay
Apr 10, 2018	1	high1	8	1	18.91
	1	high2	8	2	18.55
Apr 11, 2018	2	high1	10	1	18.04
	2	high2	8	2	18.73
Apr 12, 2018	3	high1	12	1	17.89
	3	high2	10	2	18.57
Apr 01, 2019	4	high1	8	3	18.34
	4	high2	8	4	19.39
	4	high3	8	4	20.61
Apr 02, 2019	5	high1	10	3	17.98
	5	high2	10	4	20.74
Apr 02, 2019	6	high1	8	3	18.27
	6	high2	8	4	19.70
	6	high3	8	3	18.40

TABLE A-3

SESSION SUMMARY FOR ROBUSTNESS TREATMENTS.

Date	Session #	δ	Group	# Subj	Realization	Avg Pay
Mar 10, 2020	1	0.95	high1	8	1	26.21
	1	0.95	high2	8	2	25.93
	1	0.95	high3	10	3	26.64
Mar 10, 2020	2	0.95	high	12	4	29.63
Mar 11, 2020	3	0.75	high1	10	6	26.89
	3	0.75	high2	8	7	27.10
Mar 11, 2020	4	0.75	high1	10	8	29.52
	4	0.75	high2	8	5	27.64
Mar 04, 2021	5	0.75	high	12	6	18.30
Mar 08, 2021	6	0.75	high	10	7	25.29
Mar 08, 2021	7	0.75	high	10	8	29.88
Mar 09, 2021	8	0.95	high	10	1	26.98
Mar 09, 2021	9	0.95	high	10	2	26.75
Mar 12, 2021	10	0.95	high	10	3	26.36
Mar 12, 2021	11	0.95	high	12	4	29.66
Mar 12, 2021	12	0.75	high	10	5	26.63

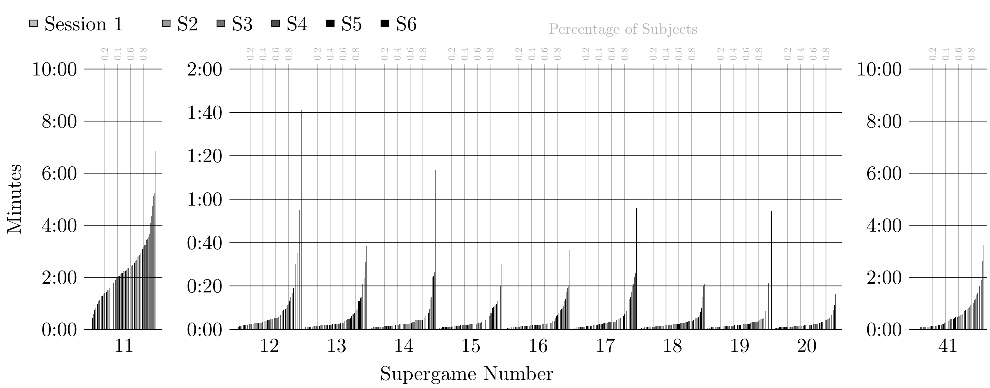


FIGURE A-3.—Time Before Start Match Button Click. *Note*: Cumulative distribution of times at which subjects clicked "start match" button. They had up to 10 minutes before the first supergame in the nonbinding stage (supergame 11) to construct the initial rule set. Once a subject clicked "start match" button, they were still able to make changes to their rule set until everyone clicked "start match" button. Subject had up to 2 minutes before each of the other supergames in the nonbinding stage. Subjects had up to 10 minutes before the second locked-response stage.

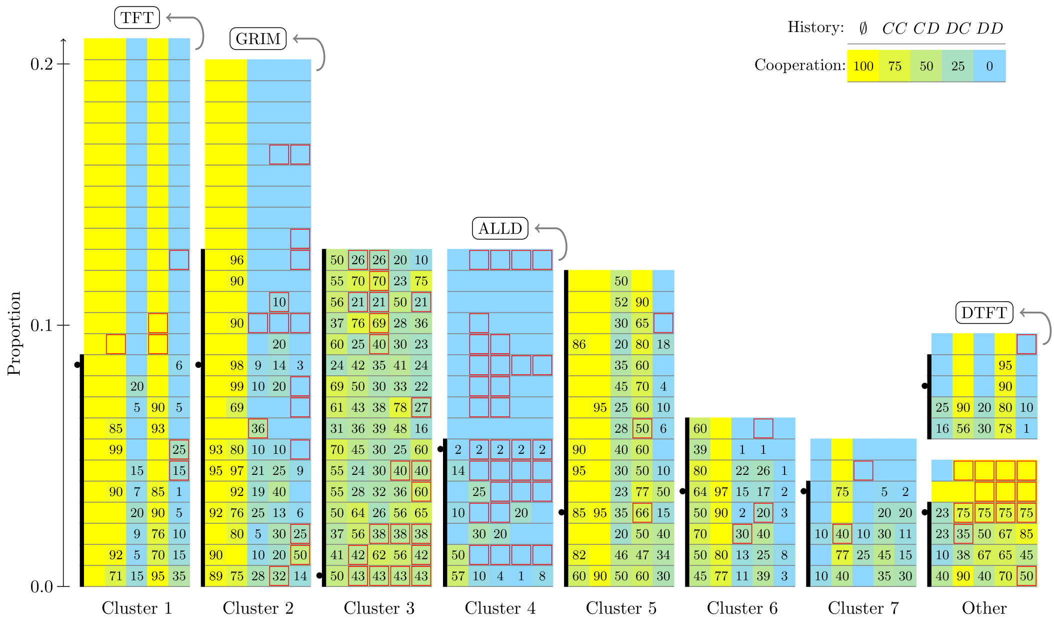


FIGURE A-4.—Elicited Strategies (First Locked-Response Stage). *Note*: Displayed strategies are elicited strategies from the first locked-response stage. Strategies are presented as vectors of cooperation percentages after each of the five histories ($\emptyset$, CC, CD, DC, DD). A black line on the left-hand side of the vector denotes a mixed strategy. Red rectangles denote a situation in which the percentage of cooperation is determined by the default rule. A black bullet on the left-hand side denotes that the strategy is the cluster exemplar. For ease of reading, 100 and 0 are not shown. Strategies are grouped using the affinity propagation clustering algorithm. Strategies in a given cluster are sorted according to $M1RM$. The pure strategies TFT, GRIM, ALLD, and DTFT are labeled.

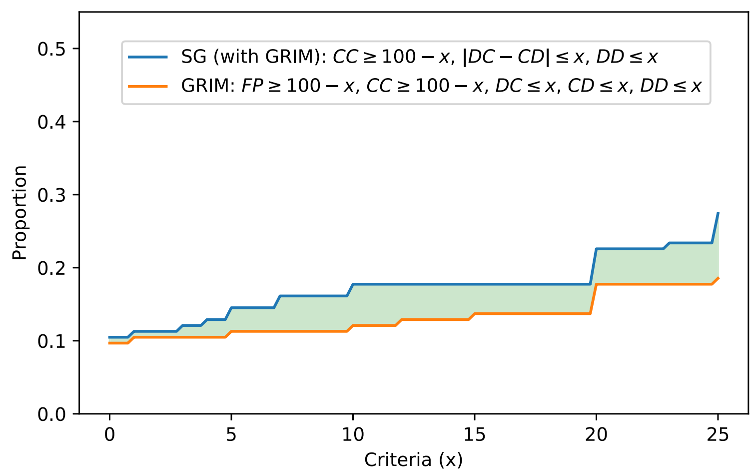


FIGURE A-5.—Evidence of Semi-Grim. *Note*: Proportion of SG and GRIM in the second locked-response stage of our experiment. For SG strategy, we require the probability of cooperation after CC to be within x percentage points from 100, after CD and DC to be within x percentage points of each other, and after DD to be within x percentage points from 0. For GRIM strategy, we require the probability of cooperation in the first period and after CC to be within x percentage points from 100, whereas after CD, DC, and DD to be within x percentage point from 0.

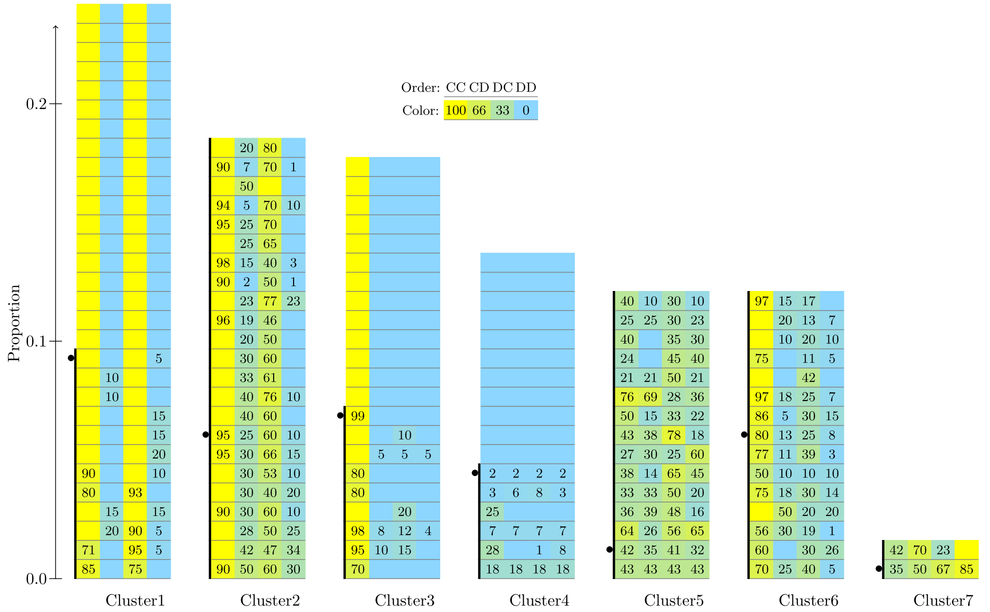


FIGURE A-6.—Cluster Analysis Without First Period. *Note*: Displayed strategies are elicited strategies from the second locked-response stage. Strategies are presented as vectors of cooperation percentages after each of the four histories (CC, CD, DC, DD). A black line on the left-hand side of the vector denotes a mixed strategy. A black bullet on the left-hand side denotes that the strategy is the cluster exemplar. For ease of reading, 100 and 0 are not shown. Strategies are grouped using the affinity propagation clustering algorithm. Strategies in a given cluster are sorted according to $M1RM$.

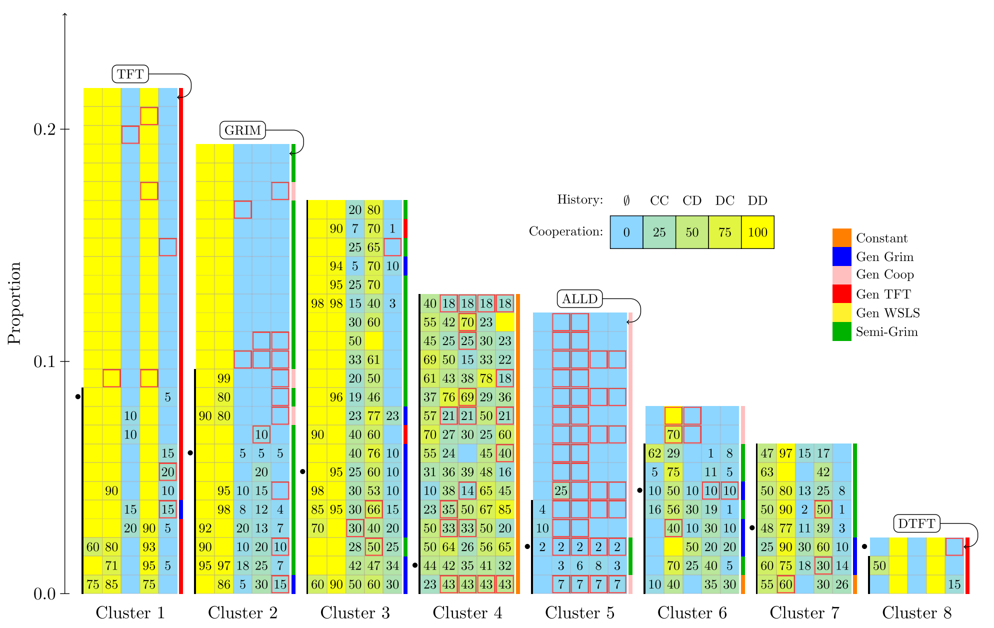


FIGURE A-7.—Likelihood Classification of Elicited Strategies. *Note*: This figure replicates Figure 2 with the following change. For each of the 124 elicited strategies, we determine which of the six estimated strategies from Table 3(a) has the highest likelihood of matching observed play. The highest likelihood strategy is denoted by a colored bar on the right side of the corresponding strategy.

TABLE A-4

MIXTURE MODEL ESTIMATES FOR MEMORY-1 MARKOV STRATEGIES (WITH ICL-BIC ELIMINATION).

	Initial Estimate		After Elimination 1		After Elimination 2	
Strategy	Proportion	Parameter	Proportion	Parameter	Proportion	Parameter
CONST	0.145	0.371	0.145	0.371	0.153	0.366
Gen GRIM	0.126	0.243	0.126	0.243	0.335	0.025
Gen COOP	0.130	0.500	0.131	0.500		
Gen TFT	0.249	0.964	0.249	0.964	0.255	0.969
Gen WSLS	0.000	0.501				
SG	0.350	0.173	0.350	0.173	0.257	0.355
LL	-6407.535		-6407.535		-6416.006	
ICL	6452.252		6447.431		6446.672	

Note: This table replicates Table III(a) except that elimination is performed based on the integrated classification likelihood-Bayes information criterion (ICL-BIC) procedure as in Breitmoser (2015).

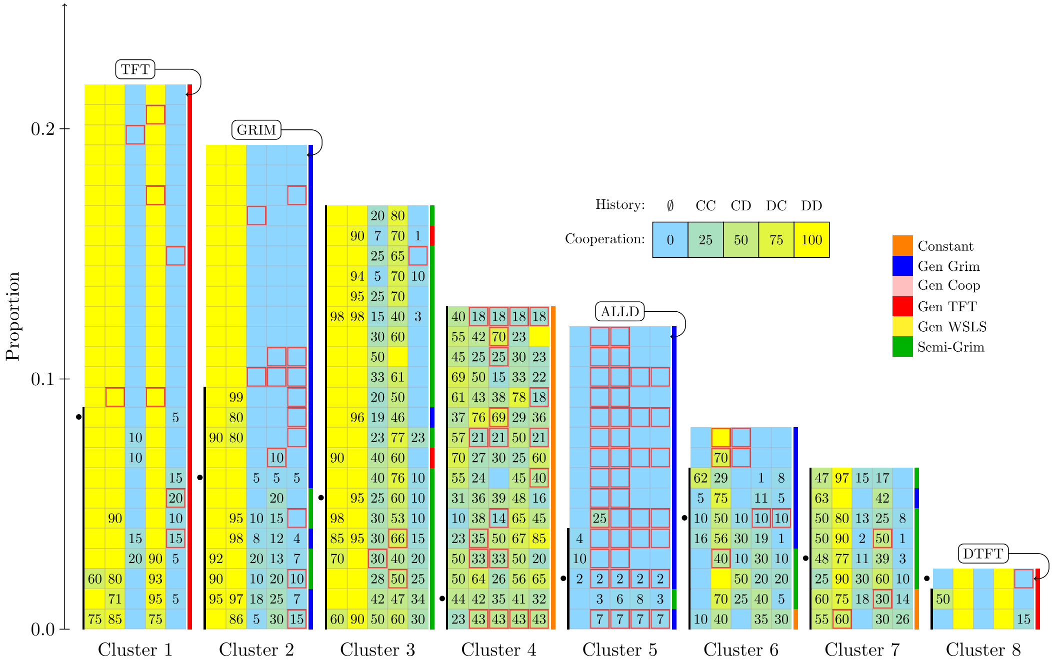


FIGURE A-8.—Likelihood Classification of Elicited Strategies (with ICL-BIC elimination). *Note*: This table replicates Figure A-7 with the following change. Strategies Gen WSLS and Gen Coop have been removed based on the integrated classification likelihood-Bayes information criterion (ICL-BIC) elimination procedure as in Breitmoser (2015).

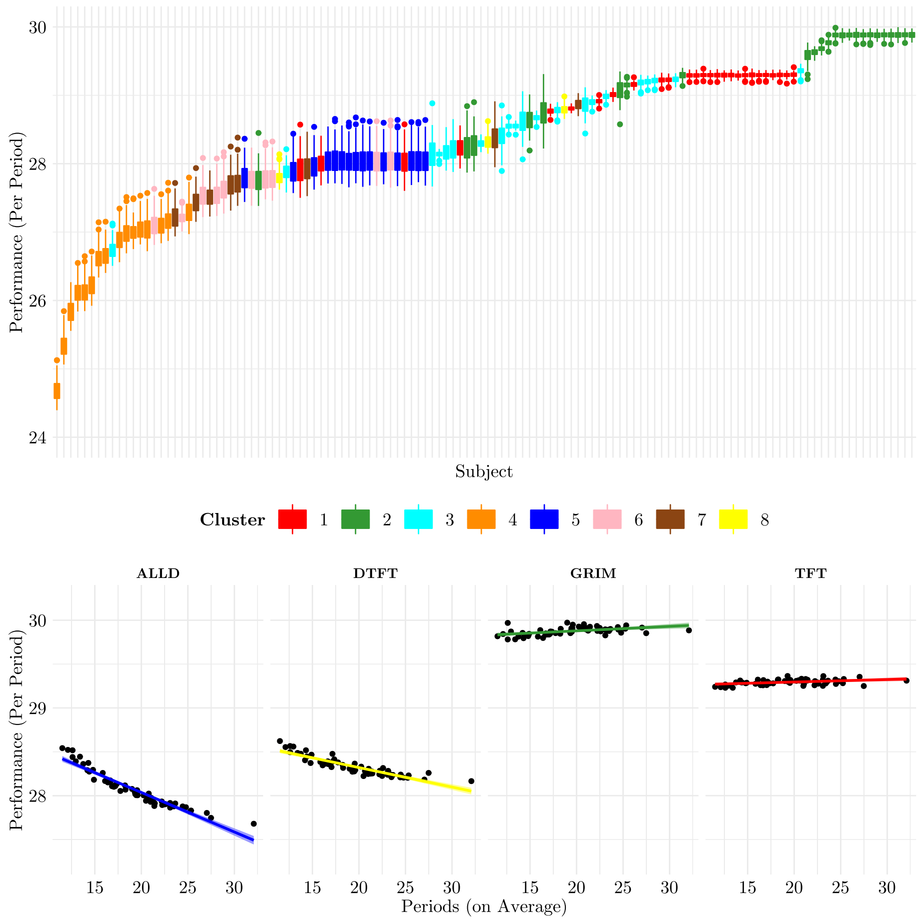
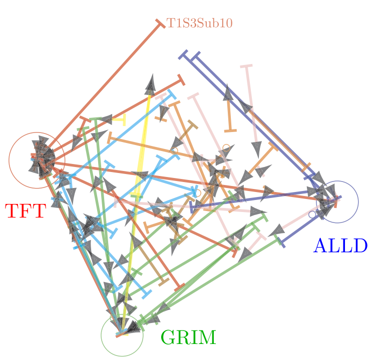


FIGURE A-9.—Strategy Performance Variability. *Note*: We ran 100 round-robin tournaments in which each strategy from the second locked-response stage was matched with each other's strategy from the population for 20 supergames of random duration (supergame lengths were the same for each pair within a tournament, but different across tournaments). **Top**: Variability in strategy performance across 100 tournaments. **Bottom**: Each point denotes the strategy performance (the average payoff per period) in one tournament and the realized average supergame length in the same tournament. Payoff is calculated as the average earnings per period. Clusters are identified by color.



		0.31	0.30	0.00	0.10	0.15	0.14	0.00	0.00
TFT	0.22	0.93	0.04		0.00	0.01	0.02		
GRIM	0.19	0.09	0.75			0.07	0.09		
MTFT1	0.17	0.20	0.33		0.05	0.16	0.27		
CONST	0.13	0.01	0.16		0.62		0.20		
ALLD	0.12	0.24				0.76			
MTFT2	0.08	0.17	0.27		0.12		0.45		
MGRIM	0.06	0.09	0.55		0.01	0.14	0.21		
DTFT	0.02	0.25	0.63			0.12			

FIGURE A-10.—Long-Run Evolution of Strategies.

APPENDIX B: SFEM ESTIMATION DETAILS

There are three main steps to the strategy frequency estimation method (Dal Bó and Fréchette (2011)). The first step is to specify the set of K strategies considered for estimation. The second step is to determine the likelihood that strategy $k \in K$ generates the choices made by each subject $n \in N$ over multiple supergames. The third step is to formulate the likelihood function. In what follows, we present the matrix approach to SFEM described in Rosokha and Wei (2020):

- Let X denote a $K \times N$ matrix of the number of correct matches for all combinations of subjects and strategies. That is, for each entry in the matrix $X(k, n)$ we compare subject n's actual play with how strategy k would have played in her place and find the number of periods in which subject n's play correctly matches the play of strategy k.
- Let Y denote a $K \times N$ matrix of the number of mismatches when comparing subjects' play with what the strategies would do in their place.
- Define a Hadamard-product P:

$$P = \beta^X \circ (1 - \beta)^Y, \tag{1}$$

  where β is the probability that a subject plays according to a strategy and $(1 - \beta)$ is the probability that the subject deviates from that strategy. Thus, each entry $P(k, n)$ is the likelihood that strategy k generated the observed choices by subject n.

TABLE B-1

SFEM ESTIMATES FOR OUR DATA.

Treatment	Supergames	ALLD	TFT	GRIM	DTFT	2TFT	D.COOP	TF2T	GRIM2	ALLC	GRIM3	TF3T	DTF2T	DTF3T	2TF2T	β
1.M1D95	**S06–10**	0.14 (0.03)	0.23 (0.04)	0.09 (0.04)	0.11 (0.04)	0.03 (0.01)	0.01 (0.01)	0.05 (0.04)	0.08 (0.02)	0.01 (0.01)	0.06 (0.02)	0.04 (0.02)	0.01 (0.02)	0.03 (0.02)	0.06 (0.04)	0.92 (0.01)
1.M1D95	**S16–20**	0.14 (0.03)	0.33 (0.05)	0.12 (0.03)	0.12 (0.03)	0.07 (0.05)	0.03 (0.03)	0.07 (0.03)		0.04 (0.02)	0.02 (0.02)	0.02 (0.03)		0.04 (0.01)		0.91 (0.01)
1.M1D95	**S21–40**	0.18 (0.03)	0.35 (0.05)	0.19 (0.04)	0.07 (0.02)	0.06 (0.02)	0.07 (0.02)			0.04 (0.02)	0.01 (0.01)					0.9 (0.01)
1.M1D95	**S41–60**	0.22 (0.03)	0.41 (0.04)	0.21 (0.04)	0.06 (0.01)	0.03 (0.02)	0.03 (0.02)			0.01 (0.01)	0.02 (0.01)					0.91 (0.01)
2.M2D95	**S06–10**	0.1 (0.02)	0.23 (0.05)	0.07 (0.04)	0.18 (0.03)	0.09 (0.04)		0.04 (0.04)	0.06 (0.04)		0.08 (0.04)	0.07 (0.05)	0.01 (0.01)	0.02 (0.02)	0.03 (0.03)	0.93 (0.01)
2.M2D95	**S16–20**	0.16 (0.04)	0.24 (0.07)	0.08 (0.03)	0.13 (0.03)	0.05 (0.03)	0.02 (0.01)	0.1 (0.04)	0.01 (0.03)		0.04 (0.03)	0.06 (0.04)	0.03 (0.02)			0.93 (0.01)
2.M2D95	**S21–40**	0.19 (0.03)	0.22 (0.05)	0.11 (0.04)	0.08 (0.03)	0.03 (0.01)	0.09 (0.03)	0.06 (0.03)	0.03 (0.02)	0.09 (0.02)	0.01 (0.01)		0.06 (0.02)		0.03 (0.02)	0.92 (0.01)
2.M2D95	**S41–60**	0.21 (0.05)	0.25 (0.03)	0.12 (0.02)	0.1 (0.03)	0.06 (0.03)	0.07 (0.03)	0.02 (0.01)	0.04 (0.03)	0.04 (0.02)		0.03 (0.02)	0.05 (0.03)	0.01 (0.01)	0.01 (0.01)	0.93 (0.01)
3.M2D75	**S06–10**	0.41 (0.06)	0.19 (0.07)	0.09 (0.04)	0.14 (0.04)	0.04 (0.05)			0.04 (0.04)	0.03 (0.02)		0.01 (0.01)				0.9 (0.01)
3.M2D75	**S16–20**	0.42 (0.07)	0.05 (0.03)	0.28 (0.07)	0.18 (0.05)				0.02 (0.02)	0.02 (0.02)						0.92 (0.01)
3.M2D75	**S21–40**	0.44 (0.05)	0.12 (0.03)	0.12 (0.05)	0.16 (0.03)		0.07 (0.03)					0.01 (0.01)		0.03 (0.02)	0.01 (0.01)	0.91 (0.01)
3.M2D75	**S41–60**	0.49 (0.06)	0.04 (0.03)	0.15 (0.04)	0.16 (0.04)	0.05 (0.04)	0.03 (0.02)		0.01 (0.01)				0.01 (0.01)	0.03 (0.02)	0.01 (0.0)	0.94 (0.01)

- Then, using the matrix dot product, we define the log-likelihood function $\mathcal{L}$:

$$\mathcal{L}(\beta, \phi) = \ln\bigl(\phi' \cdot P\bigr) \cdot \mathbf{1}, \tag{2}$$

where ϕ is a vector of strategy frequencies.
- The optimization objective is to find $\beta \geq 0.5$ and ϕ to maximize $\mathcal{L}$.

Table B-1 presents SFEM estimates on the data from our experiment using the set of 20 strategies from Fudenberg, Rand, and Dreber (2012). Values of 0.00 are dropped for ease of reading. Only strategies with at least 10% total proportion across various treatments/supergames are shown.

REFERENCES

AGRANOV, MARINA, AND PIETRO ORTOLEVA (2017): "Stochastic Choice and Preferences for Randomization," *Journal of Political Economy*, 125 (1), 40–68. [2297]

AOYAGI, MASAKI, VENKATARAMAN BHASKAR, AND GUILLAUME R. FRÉCHETTE (2019): "The Impact of Monitoring in Infinitely Repeated Games: Perfect, Public, and Private," *American Economic Journal: Microeconomics*, 11 (1), 1–43. [2296]

AXELROD, ROBERT (1980a): "Effective Choice in the Prisoner's Dilemma," *Journal of conflict resolution*, 24 (1), 3–25. [2296]

——— (1980b): "More Effective Choice in the Prisoner's Dilemma," *Journal of Conflict Resolution*, 24 (3), 379–403. [2296]

BACKHAUS, TERESA, AND YVES BREITMOSER (2018): "God Does not Play Dice, but Do We?" Tech. rep., CRC TRR 190 Rationality and Competition. [2297]

BLAND, JAMES R. (2020): "Heterogeneous Trembles and Model Selection in the Strategy Frequency Estimation Method," *Journal of the Economic Science Association*, 6 (2), 113–124. [2308]

BLOOMFIELD, ROBERT (1994): "Learning a Mixed Strategy Equilibrium in the Laboratory," *Journal of Economic Behavior & Organization*, 25 (3), 411–436. [2296]

BREITMOSER, YVES (2015): "Cooperation, but No Reciprocity: Individual Strategies in the Repeated Prisoner's Dilemma," *The American Economic Review*, 105 (9), 2882–2910. [2296-2299,2307-2309,2320,2322, 2327]

CAMERA, GABRIELE, MARCO CASARI, AND MARIA BIGONI (2012): "Cooperative Strategies in Anonymous Economies: An Experiment," *Games and Economic Behavior*, 75 (2), 570–586. [2296]

CASON, TIMOTHY N., AND VAI-LAM MUI (2019): "Individual versus Group Choices of Repeated Game Strategies: A Strategy Method Approach," *Games and Economic Behavior*, 114, 128–145. [2296,2297]

CASON, TIMOTHY N., DANIEL FRIEDMAN, AND ED HOPKINS (2013): "Cycles and Instability in a Rock–Paper–Scissors Population Game: A Continuous Time Experiment," *Review of Economic Studies*, 81 (1), 112–136. [2296]

CHMURA, THORSTEN, AND WERNER GÜTH (2011): "The Minority of Three-Game: An Experimental and Theoretical Analysis," *Games*, 2 (3), 333–354. [2296]

COCKBURN, IAIN, AND REBECCA HENDERSON (1994): "Racing to Invest? The Dynamics of Competition in Ethical Drug Discovery," *Journal of Economics & Management Strategy*, 3 (3), 481–519. [2295]

DAL BÓ, PEDRO, AND GUILLAUME R. FRÉCHETTE (2011): "The Evolution of Cooperation in Infinitely Repeated Games: Experimental Evidence," *The American Economic Review*, 101 (1), 411–429. [2296,2297,2307, 2313,2319,2320,2329]

——— (2018): "On the Determinants of Cooperation in Infinitely Repeated Games: A Survey," *Journal of Economic Literature*, 56 (1), 60–114. [2296,2297,2299,2305,2306]

——— (2019): "Strategy Choice in the Infinitely Repeated Prisoner's Dilemma," *American Economic Review*, 109 (11), 3929–3952. [2296,2297,2300,2301,2308,2313,2319,2320]

ELY, JEFFREY C., JOHANNES HÖRNER, AND WOJCIECH OLSZEWSKI (2005): "Belief-Free Equilibria in Repeated Games," *Econometrica*, 73 (2), 377–415. [2297,2299]

EMBREY, MATTHEW, GUILLAUME R. FRÉCHETTE, AND SEVGI YUKSEL (2018): "Cooperation in the Finitely Repeated Prisoner's Dilemma," *The Quarterly Journal of Economics*, 133 (1), 509–551. [2313]

ENGLE-WARNICK, JIM, AND ROBERT L. SLONIM (2006): "Inferring Repeated-Game Strategies From Actions: Evidence From Trust Game Experiments," *Economic Theory*, 28 (3), 603–632. [2296]

FREY, BRENDAN J., AND DELBERT DUECK (2007): "Clustering by Passing Messages Between Data Points," *Science*, 315 (5814), 972–976. [2304]

FUDENBERG, DREW, DAVID G. RAND, AND ANNA DREBER (2012): "Slow to Anger and Fast to Forgive: Cooperation in an Uncertain World," *American Economic Review*, 102 (2), 720–749. [2296,2307,2308,2319, 2330]

GILL, DAVID, AND YAROSLAV ROSOKHA (2020): "Beliefs, Learning, and Personality in the Indefinitely Repeated Prisoner's Dilemma," Available at SSRN 3652318. [2297]

HANAKI, NOBUYUKI, RAJIV SETHI, IDO EREV, AND ALEXANDER PETERHANSL (2005): "Learning Strategies," *Journal of Economic Behavior & Organization*, 56 (4), 523–542. [2313]

HARUVY, ERNAN, DALE O. STAHL, AND PAUL W. WILSON (2001): "Modeling and Testing for Heterogeneity in Observed Strategic Behavior," *Review of Economics and Statistics*, 83 (1), 146–157. [2296]

IOANNOU, CHRISTOS A., AND JULIAN ROMERO (2014): "A Generalized Approach to Belief Learning in Repeated Games," *Games and Economic Behavior*, 87, 178–203. [2313]

JONES, GARETT (2008): "Are Smarter Groups More Cooperative? Evidence From Prisoner's Dilemma Experiments, 1959–2003," *Journal of Economic Behavior & Organization*, 68 (3), 489–497. [2300]

KAHN, LAWRENCE M. (1993): "Unions and Cooperative Behavior: The Effect of Discounting," *Journal of Labor Economics*, 11 (4), 680–703. [2295]

MAGGI, GIOVANNI (1999): "The Role of Multilateral Institutions in International Trade Cooperation," *American Economic Review*, 89 (1), 190–214. [2295]

MAILATH, GEORGE J., AND LARRY SAMUELSON (2006): *Repeated Games and Reputations: Long-Run Relationships*. Oxford university press. [2295]

NOUSSAIR, CHARLES, AND MARC WILLINGER (2011): "Mixed Strategies in an Unprofitable Game: An Experiment," Tech. rep., Citeseer. [2296]

OCHS, JACK (1995): "Games With Unique, Mixed Strategy Equilibria: An Experimental Study," *Games and Economic Behavior*, 10 (1), 202–217. [2296]

POWELL, ROBERT (1993): "Guns, Butter, and Anarchy," *American Political Science Review*, 87 (1), 115–132. [2295]

PROTO, EUGENIO, ALDO RUSTICHINI, AND ANDIS SOFIANOS (2019): "Intelligence, Personality, and Gains From Cooperation in Repeated Interactions," *Journal of Political Economy*, 127 (3), 1351–1390. [2300]

——— (forthcoming): "Intelligence, Errors and Strategic Choices in the Repeated Prisoners' Dilemma," Review of Economic Studies. [2300]

RAND, DAVID G. (2016): "Cooperation, Fast and Slow: Meta-Analytic Evidence for a Theory of Social Heuristics and Self-Interested Deliberation," *Psychological science*, 27 (9), 1192–1206. [2304]

ROMERO, JULIAN, AND YAROSLAV ROSOKHA (2018): "Constructing Strategies in the Indefinitely Repeated Prisoner's Dilemma Game," *European Economic Review*, 104, 185–219. [2296-2302,2304,2313,2316,2317, 2321]

——— (2019): "The Evolution of Cooperation: The Role of Costly Strategy Adjustments," *American Economic Journal: Microeconomics*, 11 (1), 299–328. [2313]

——— (2023): "Supplement to 'Mixed Strategies in the Indefinitely Repeated Prisoner's Dilemma'," *Econometrica Supplemental Material*, 91, https://doi.org/10.3982/ECTA17482. [2301]

ROSOKHA, YAROSLAV, AND CHEN WEI (2020): "Cooperation in Queueing Systems," Available at SSRN 3526505. [2329]

SEGAL, NANCY L., AND SCOTT L. HERSHBERGER (1999): "Cooperation and Competition Between Twins: Findings From a Prisoner's Dilemma Game," *Evolution and Human Behavior*, 20 (1), 29–51. [2300]

SELTEN, REINHARD (1967): "Die Strategiemethode zur Erforschung des eingeschränkt rationalen Verhaltensim Rahmen eines Oligopolexperiments," in *Beiträge zur Experimentellen Wirtschaftsforschung*, ed. by Heinz Sauermann. Tiubingen: J. C. B. Mohr, 136–168. [2296]

SELTEN, REINHARD, MICHAEL MITZKEWITZ, AND GERALD R. UHLICH (1997): "Duopoly Strategies Programmed by Experienced Players," Econometrica: Journal of the Econometric Society, 517–555. [2296]

SHACHAT, JASON M. (2002): "Mixed Strategy Play and the Minimax Hypothesis," *Journal of Economic Theory*, 104 (1), 189–226. [2296]

STAHL, DALE O., AND PAUL W. WILSON (1994): "Experimental Evidence on Players' Models of Other Players," *Journal of economic behavior & organization*, 25 (3), 309–327. [2296]

——— (1995): "On Player's Models of Other Players: Theory and Experimental Evidence," *Games and Economic Behavior*, 10 (1), 218–254. [2296]

Co-editor Alessandro Lizzeri handled this manuscript.

Manuscript received 11 July, 2019; final version accepted 26 March, 2023; available online 7 August, 2023.