Facing the Grim Truth: Repeated Prisoner's Dilemma Against Robot Opponents*

John Duffy† Ed Hopkins‡ Tatiana Kornienko§

August 3, 2021

Abstract

We report on an experiment where subjects play an indefinitely repeated prisoner's dilemma game against robot opponents known to play the Grim trigger strategy. The game's continuation probability is varied affecting whether cooperation can be rationalized. We find that overall first round cooperation is higher than predicted and increases in the continuation probability. However, some subjects defect after cooperating, perhaps anticipating the end of a supergame and others cooperate after defecting, which is harder to rationalize. The latter type of error decreases in a measure of cognition, suggesting that cognition is important for the successful implementation of strategies.

Keywords: Cognition, Cooperation, Prisoner's Dilemma, Repeated Games, Experimental Economics.

JEL Codes: C72, C73, C92.

*We thank audiences at UC Irvine and the Behavioral Game Theory Conference at the University of East Anglia. Funding for this project was provided by the UC Irvine School of Social Science. The experimental protocol was approved by the UC Irvine Institutional Review Board, HS# 2011-8378. We thank Patrick Julius for programming the experiment. The reported experiment is a part of a project that was pre-registered on the AEA RCT registry, https://doi.org/10.1257/rct.6318-1.0.

†University of California, Irvine, duffy@uci.edu

‡University of Edinburgh, E.Hopkins@ed.ac.uk

§University of Edinburgh, Tatiana.Kornienko@ed.ac.uk

1 Introduction

Cooperation in repeated interactions is an important aspect of social behavior. It has also been the subject of much recent experimental research, surveyed in Dal Bó and Fréchette (2018). Yet it is still unclear what exactly people do in strategic prisoner's dilemma settings. One common observation is that subjects sometimes cooperate even in a one-shot prisoner's dilemma. Equally, subjects sometimes cooperate too little when a dilemma is indefinitely repeated. For example, when the continuation probability is sufficiently high that strategies supporting cooperation such as the Grim trigger strategy could be both an equilibrium and an empirical best response, a significant proportion of subjects choose always to defect, thereby leaving money on the table. However, it is difficult to interpret either behavior as mistaken, given many possible confounds, including diverse beliefs about the strategies employed by others, heterogeneous risk attitudes, social preferences and cognitive limitations.

This study attempts to simplify the analysis by conducting a novel experiment that excludes several confounding factors by design. Subjects play a series of indefinitely repeated prisoner's dilemma (IRPD) games with different continuation probabilities against a robot opponent known to play the Grim trigger strategy. This design reduces or eliminates multiple equilibria, strategic uncertainty and social preferences as factors influencing cooperation, and focuses attention on the cognitive task of trading off present gain against future reward. The optimal policy is simple in theory: a subject should cooperate if and only if the continuation probability, δ, is above a critical level, here 0.5.

We find that first round cooperation is strongly increasing in the continuation probability, ranging from 9.5% when $\delta = 0.1$ to 76% when $\delta = 0.7$. This responsiveness to δ is much greater than that estimated based on Dal Bó and Fréchette (2018) in standard subject versus subject experiments, which suggests that our design is successful in reducing strategic uncertainty. However, on average, subjects cooperate too much in the first round (48% of decisions rather than 33%), and too little overall (50% rather than 56%). Beyond the first round, there are substantial deviations from the optimal strategy. First, 52% of subjects cooperate at least once after already having defected in a supergame, behavior that is difficult to rationalize. Further, 24% of subjects make this type of mistake repeatedly, in at least 3 out of 17 relevant supergames. Second, subjects commonly defect after having started out the play of a supergame by cooperating. Specifically, we find cooperation significantly decreasing with the round number when theory suggests it should be constant. Although the supergames have an unknown, random end, subjects may be "sniping": defecting in the round they guess will be the last round of the supergame.[1] We are able to identify these behaviors only because of our novel, single-person design but our findings offer an alternative interpretation of results from other repeated game experiments.[2]

[1]As is standard in indefinitely repeated games, we employ a constant termination probability of $1 - \delta$, where δ is known to subjects. Subjects may nonetheless believe that the termination probability rises over time. Alternatively, Mengel et al. (2021) find that subjects respond to past *realized* supergame lengths.

[2]Romero and Rosokha (2018) and Cooper and Kagel (2021) also report decreasing cooperation rates in indefinitely repeated prisoner's dilemma experiments. However, there such decreased cooperation may be

We further find that test scores from a cognitive reflection test predict earnings and are negatively associated with the error of cooperating after having defected, supporting the idea that cognitive failures are a cause of deviations from optimal behavior. However, we also find that individuals with higher cognitive test scores are more likely to snipe. More generally, our results suggest that cognitive factors are important in explaining the excess cooperation observed when δ is low and the insufficient cooperation observed when δ is high.

Two prior studies, Roth and Murnighan (1978) and Murnighan and Roth (1983), used a similar idea of having groups of subjects play against a fixed strategy, as well as being the first to run experiments on supergames with an uncertain end. A fundamental difference between those two studies and the present study is that subjects in those earlier studies were *not informed* of the strategy they faced or that their opponent was in fact the experimenter. Thus, subjects in those prior studies, who participated in sessions along with other subjects, faced some strategic uncertainty.[3] By contrast, in this experiment we instruct subjects that they are playing against programmed opponents who *play the Grim trigger strategy*. Second, these two prior studies did not allow subjects to play multiple supergames with the same continuation probability. Dal Bó and Fréchette (2018) argue that such repetition is an important feature, in that more recent experiments have found significant learning effects with experience. Learning may be less important in our setting where there is no strategic uncertainty, but nonetheless we think it is important to give subjects an opportunity to learn by doing. In the most similar paper, Duffy and Xie (2016) consider play against robot players known to play the Grim trigger strategy but in an n-player Prisoner's Dilemma game under random matching, where they vary n and the stage game payoffs but *not* δ.

Of course, there are many experiments on the repeated prisoner's dilemma, where subjects play other subjects. For example, Proto et al. (2019) also find that individual differences between subjects affect play, with higher cognitive ability players being more cooperative, making fewer mistakes and earning higher payoffs. The main difference is that, in a two human subject pairing there is not a unique optimum policy as there is here, and so errors have to be inferred. For example, Proto et al. (2019) assume that playing defect directly after both players chose to cooperate is an error in implementation. However, here it seems that such behavior may represent an attempt to guess the final round. Further, as noted, with our design we can also identify excessive cooperation.

Our methodology is similar to that of Charness and Levin (2009) who show experimentally that the winner's curse phenomenon is still a factor in a single person bidding problem.

caused by beliefs that cooperation by opponents may be about to end.

[3]In Roth and Murnighan (1978), p. 194 subjects "were told that they played a programmed opponent, but were not told what strategy he would be using." The programmed opponent was in fact an experimenter playing the Tit for Tat (or "matching") strategy. In Murnighan and Roth (1983) p. 289, subjects "were told that they would be playing a different individual in each of the three sessions but that the person's identity would not be revealed. Actually all of the subjects played against the experimenter who implemented either matching [Tit for Tat] or [the] unforgiving strategy [Grim trigger]." Roth and Murnighan (1978) p. 194 explain that such design choices were made to "control for differences in subjects' behavior due to differences in their opponents."

That is, in both cases there are individual cognitive failures that are responsible for misbehavior in larger groups. Here, the individual failure is the inability to play a constant strategy in a stationary environment, which leads to suboptimal behavior even in the absence of strategic uncertainty. The difference here (besides the different game investigated) is our use of a within-subject design where subjects face situations both where cooperation is optimal and where it is not. In that sense, we are adapting the methodology of Duffy et al. (2021) and Charness et al. (2021), which also have an experimental design where subjects face opposed environments, to study repeated interactions.

2 Theory and Hypotheses

In our experiment, subjects play the indefinitely repeated prisoner's dilemma with continuation probability δ against a computer playing a fixed strategy. The specific payoffs subjects faced in the stage game are given in (1),

$$\begin{array}{ccc} & \text{X} & \text{Y} \\ \text{X} & \mathbf{75},\ \textit{75} & \mathbf{15}, \textit{120} \\ \text{Y} & \mathbf{120},\ \textit{15} & \mathbf{30},\ \textit{30} \end{array} \tag{1}$$

where X (Y) denote the cooperate (defect) actions. The main theoretical prediction tested in our experiment comes from the Folk Theorem for repeated games which (Mailath and Samuelson, 2006, p. 69) states that if players are sufficiently patient, then any pure-action profile whose payoff strictly dominates the pure-action minimax is a subgame perfect equilibrium of the repeated game in which this action profile is played in every period. This result carries over to the situation of indefinitely repeated games by replacing "players are sufficiently patient" with "the continuation probability is sufficiently high". However, here for one player, the computer, the strategy is fixed to be the Grim strategy. This converts the problem from a game with multiple equilibria to a single person decision problem with a unique optimum policy. This is to cooperate (defect) in every round of a supergame if the continuation probability exceeds (is below) a critical level δ^*, which for our parameterization (1) is 0.5.

To see this, note first that since the computer is programmed to play the Grim trigger strategy, it first cooperates and continues to cooperate so long as all previous play by its opponent has been cooperate, but after any defection, it switches to defect for all subsequent periods. Thus, any player should understand, given that the continuation probability is fixed at δ, that the return to playing cooperate (X) forever is

$$75 + 75\delta + 75\delta^2 + ... = \frac{75}{1-\delta}. \tag{2}$$

In contrast, the expected return to defecting in period one is,

$$120 + 30\delta + 30\delta^2 + ... = 120 + \frac{30\delta}{1-\delta}. \tag{3}$$

Simple calculations reveal that (2) is greater than (3) if $\delta > 0.5$. Thus, the critical continuation probability is $\delta^* = 0.5$.

Note that, because the continuation probability is constant over time, the problem is stationary and so, if it is optimal to cooperate in period one, it is also optimal to cooperate at all future periods. Thus, it cannot be optimal to switch within a supergame from cooperate to defect Further, given the fixed Grim strategy of the computer, if a player ever defects, it is always optimal to continue defecting and not to switch back to cooperating. This brings us to a simple hypothesis.

Hypothesis 1. *Rational Play: subjects should play Cooperate, X (Defect, Y) in every round of every supergame when* $\delta > (<)\ \delta^* = 0.5$.

Three important factors present in the standard two player repeated prisoner's dilemma are removed in our experimental design. First, our design reduces the problem of *multiple equilibria*. When the continuation probability is sufficiently high for cooperation to be supported, there are typically an infinite number of equilibria which presents subjects with difficult coordination problems. Given that the opponent in our design is playing the Grim strategy, the set of equilibria is reduced to just two, always cooperate or always defect. Second, our design minimizes *strategic uncertainty*. This type of uncertainty is always present in the standard design because subjects do not know which strategy their opponent is following. Indeed, a simplification used by Dal Bó and Fréchette (2018) is to suppose that strategy choices are limited to Grim trigger and the strategy of always defecting. They show that there exists $\delta^{RD} > \delta^*$ such that only if $\delta > \delta^{RD}$ is it risk dominant to choose the Grim strategy and hence start out cooperating. Or, in other words, although it is an equilibrium to cooperate as long as $\delta > \delta^*$, strategic uncertainty can make it difficult to cooperate unless $\delta > \delta^{RD}$, a higher hurdle.

Third, researchers have found evidence for *social preferences* being important in many experimental settings, see, e.g., Camerer (2003), Chaudhuri (2008). In the repeated prisoner's dilemma, Bernheim and Stark (1988) and Duffy and Muñoz-García (2012) show how social preferences, in the form of positive concerns for the other player, reduce δ^*. Thus, in conventional experiments, subjects with social preferences could cooperate even when $\delta < \delta^*$. Further, there is a second order effect. Subjects who are entirely self-interested, but who believe that other subjects have social preferences and are thus more likely to cooperate, will themselves be more cooperative than in the absence of such beliefs (see, for example, Andreoni and Samuelson (2006)). That is, strategic uncertainty and social preferences can potentially interact with one another. However, in our design, since the opponent that subjects face always is known to play the Grim strategy there is a unique optimum response and no strategic uncertainty. Further, subjects are unlikely to feel altruism toward their computer opponent, or believe that it has altruistic feelings for them. Thus, multiple equilibria, strategic uncertainty and social preferences as well as any interactions between them are minimized, if not eliminated by our design.

3 Experimental Design

The main experimental task consisted of the play of 24 indefinitely repeated prisoner's dilemma games or "supergames" against a computer program known by subjects to play the Grim trigger strategy. The payoff matrix for the prisoner's dilemma stage game was held constant across all treatment conditions and is shown in (1). Subjects were instructed that the rows referred to their action and the columns referred to the computer opponent's actions and that the first number in each cell (in bold) was their payoff in points and the second number in each cell (in italics) was the computerized opponent's payoff in points.[4]

The 24 indefinitely repeated games were chosen with the following considerations. First, we wanted subjects to have some experience with the same continuation probability, and we also wanted to vary the continuation probability so as to assess the subject's attentiveness to the nature of the supergame they were playing. We chose to have them face 6 different continuation probabilities 4 times each, which yields the 24 supergame total.

The set of 6 continuation probabilities $\delta \in \{0.1, 0.25, 0.33, 0.4, 0.67, 0.7\}$ were selected using several criteria. First, with this set, the expected theoretical payoff is the same for subjects who are biased towards always cooperating and for those biased towards always defecting. Second, the expected payoff from always following the theoretically optimal strategy relative to either of the fully biased strategies is substantial and results in a clear difference. Finally, since the threshold probability for sustaining cooperation in the stage game (1), $\delta^* = 0.5$, we did not want the simple heuristic of cooperating in 50% of the supergames to correspond to the optimal policy. Instead, optimal play would involve cooperating in round 1 of just 8 of the 24 supergames (those with $\delta = 0.67$ or 0.7) and always defecting in the other 16 supergames.

The experiment was computerized and conducted entirely online. It was programmed using oTree (Chen et al. (2016)). Example screenshots are provided in Appendix E. Subjects were always informed of the probability that the supergame (sequence) would continue with another round. They were also reminded of the strategy (X or Y) that their computer opponent would play in each round (following the Grim trigger strategy and based on the history of play in all prior rounds of the current supergame) on *same* decision screen where they made their own action choice (X or Y) for that same round. Thus, any strategic uncertainty should have been eliminated.

Subjects were 100 undergraduate students, 52% female, recruited using Sona system from the Experimental Social Science subject pool at the University of California, Irvine. The mean age was 21.5 years with a range of 18-34. All subjects were university students from a diverse set of majors, with 36 subjects reported majoring in engineering, 25 in social sciences, 21 in life sciences, 9 in physical sciences, 7 in education, 5 in arts and humanities, and 3 in business studies (double majors double counted).

[4]We provided the computer program's payoff so that the game setup would be comparable to two player, human-to-human games, where both players' payoffs are common knowledge.

Subjects were presented with written instructions regarding the 24 IRPD games (referred to neutrally as "sequences") they would play and they had to successfully complete a comprehension quiz that tested their understanding of payoff outcomes, their understanding of the Grim trigger strategy that the computer program would follow in various scenarios and their understanding of how the continuation probability affected the duration of the game.

For half of the subjects (50/100) or in 4 out of 8 sessions, the randomly chosen realizations for the 24 supergames (4 supergames for each continuation probability δ) and the corresponding number of rounds (in parentheses) were as follows: 0.67 (4), 0.33 (1), 0.4 (2), 0.25 (1), 0.7 (3), 0.33 (2), 0.7 (5), 0.4 (1), 0.67 (2), 0.1 (1), 0.25 (1), 0.1 (1), 0.25 (2), 0.1 (1), 0.4 (1), 0.67 (4), 0.33 (2), 0.25 (1), 0.7 (2), 0.4 (3), 0.67 (2), 0.1 (1), 0.7 (4), 0.33 (1), resulting in 48 decisions (see Figure 1 (left panel) and Table A1 in Appendix A).[5] For the other 50 subjects (or the remaining 4 out of 8 sessions), the order of these supergames was reversed.[6]

In addition to subjects' choices in the 24 IRPD games, we collected demographic and other data. Subjects were asked to answer 7 cognitive reflection test (CRT) questions based on Frederick (2005), Toplak et al. (2014) and Ackerman (2014) (see Appendix E for the list of questions). We use subjects' scores on this set of 7 CRT questions as a proxy measure for their cognitive abilities.

Subjects were instructed that at the end of the session, six supergames would be chosen from all 24 played, one from each of the six different values for δ. They were further instructed that their total point earnings from those six supergames would be multiplied by \$0.01 and this amount would comprise their monetary earnings from the repeated PD game.[7] Subjects were guaranteed \$7 for showing up and completing the study. Subjects' total earnings averaged \$17.90 for a 1 hour experiment.

4 Results

As the design involved 24 supergames, each subject faced 24 first round choices. Given the randomization, 7 supergames ended after a single round, and each subject made a further 24 choices in 17 supergames lasting 2-5 rounds (see Table A1). Given the parameters of our design, the theoretically optimal strategy involves choosing to cooperate in all rounds of the 8 supergames where $\delta = \{0.67, 0.7\}$, and to defect in all rounds of the other 16 supergames. Thus, perfect theoretically optimal behavior involves exactly 8 counts of cooperation across all first rounds of 24 supergames, and exactly 18 counts of cooperation in the subsequent

[5]These supergame lengths were drawn using a random number generator. Subjects were instructed of this procedure. To reduce noise across subjects, we used the same supergame lengths across all subjects.

[6]See Appendix B for a discussion of order effects.

[7]Following the 24 repeated PD games, subjects were randomly paired to participate in a two-player task where they could earn an additional 15-100 points that were also convertible into dollars at \$0.01 per point which we do not report on in this paper.

24 decisions, amounting to exactly 26 counts of cooperation overall, out of 48 choices (see Figure 1, also Table A1). That is, by design, the theoretically optimal choices should be skewed towards defection initially, since most δs are less than 0.5 and then skewed towards cooperation later on, as it is in the longer games where cooperation is the optimal policy. Over all rounds, the theoretically optimal strategy involves always defecting in 16 supergames with $\delta < 0.5$ and always cooperating in the remaining 8 supergames with $\delta > 0.5$ (the same prediction as for first round play).

4.1 Response to the Continuation Probability δ

We find that cooperation is strongly increasing in the continuation probability (see Figure 1 (top row), also Figure D3)), with first round cooperation as low as 9.5% when $\delta = 0.1$ and as high as 76.25% when $\delta = 0.7$. This responsiveness is much greater than is observed in standard subject versus subject experiments.[8]

Finding 1. *For every round of a supergame, the rate of cooperation (defection) tends to increase (decrease) with the continuation probability* δ*.*

4.2 Choices in the First Rounds of Each Supergame

As Figure 1 (top row) shows, subjects tend to excessively cooperate in the first round of each supergame, choosing to start almost half of the 24 supergames by cooperating (see also Figure C1, top left). The mean (st.dev.) count of cooperative choices is 11.53 (5.73) which significantly exceeds the theoretical prediction of 8 (p-value=0.000). As a result, the mean (st.dev.) count of optimal choices per subject is 16.81 (3.74), which is significantly short of the theoretical prediction of 24 (p-value=0.000) (see also Figure C1).

Finding 2. *In the first rounds, subjects cooperate excessively by 44.1% on average compared to the theoretical optimum.*

4.3 Strategic Error of Cooperating after Defection (CaD)

Since the robot opponent was programmed to play the Grim trigger strategy, a defection any time in a given supergame would trigger subsequent defection by the automated opponent in all remaining rounds. Thus, choosing to cooperate after defecting earlier within the *same* supergame (CaD) is *dominated* for any δ, and is a *strategic error*. In Figure 1, such

[8] Using the probit estimates in (Dal Bó and Fréchette, 2018, p. 66, Table 4), we calculate that in subject to subject experiments that used our continuation probabilities, cooperation would be predicted to vary only from 45.5% (when $\delta = 0.1$) to 56% (when $\delta = 0.7$) among inexperienced subjects. Even after 25 supergames, cooperation in subject to subject experiments is predicted only to vary from 16.1% ($\delta = 0.1$) to 62.7% ($\delta = 0.7$).

Figure 1: Patterns of cooperation and defection using all data: 100 subjects, 2,400 supergames). Top row: Average per-subject counts of cooperation versus defection (left) and optimality versus suboptimality (right), split by δ, the first row of the horizontal axis scale and by round number, the second row of the horizontal axis scale. In the upper left panel cooperation counts, a distinction is made between undominated cooperation and dominated *cooperation after defection* (CaD). In the upper right panel suboptimal counts are divided between the error of CaD and other suboptimal choices. As the top row shows, later rounds were never reached for some δ values. (see also Table A1 in Appendix A). Bottom row (left): Average per-subject counts of first defection *within* a supergame by δ and round number. Bottom row (right): The population shares of the behavioral patterns in a supergame, by δ value. By construction, the four strategies are mutually exclusive.

suboptimal cooperation is split from un-dominated/non-erroneous cooperation (top left) and from other theoretically suboptimal choices (top right), and it amounts to 182 counts, or 7.58% of relevant observations (see also Figure D3 in Appendix C).

As Figure 2 (left) shows, only a bit less than half the subjects (48%) never made strategic errors (CaD), and 20% of subjects made at least 4 dominated choices (see also Figure D4, left panel). Some such choices could be intentional, e.g., a desire to verify the computer opponent's behavior.[9] Others could be due to a genuine "trembling hand" error of

[9]Recall, however, that the computer program's action choice of X or Y, based on the history of play and following the grim trigger strategy, was shown to subjects in advance on their decision screen prior to their making a decision.

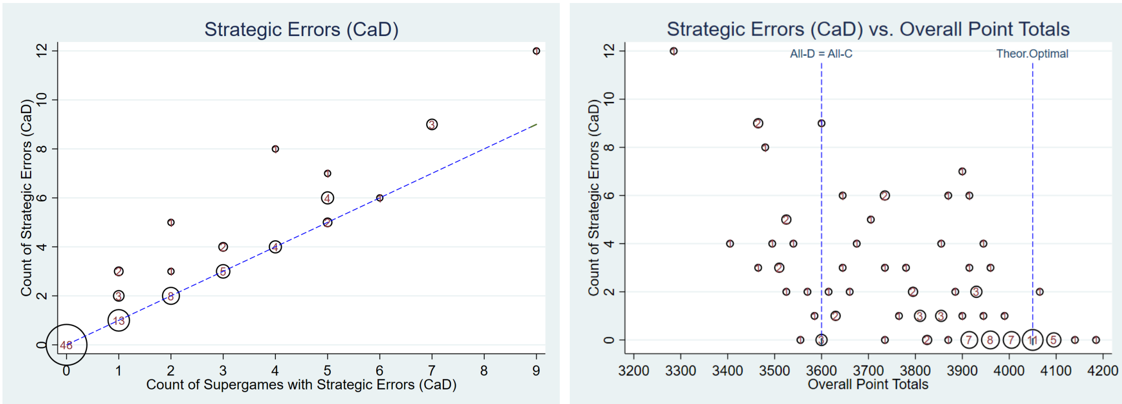

Figure 2: Strategic errors of dominated cooperation after defection (CaD), and the overall payoff. Left: overall per-subject counts of CaD instances vs count of supergames with those instances (among 17 relevant supergames). Right: the count of supergames with CaD instances vs the overall payoff. Bubble size is proportional to the share of subjects, number of subjects in the bubbles, 100 subjects total.

accidentally pressing the "defect" button without noticing it, In either case, an attentive payoff-maximizing subject would likely refrain from repeatedly making dominated choices in multiple supergames. Figure 2 (left panel) also compares the total count of strategic errors (CaD) per subject (vertical axis) versus the count of supergames where such errors were made (horizontal axis). While most strategic errors were made only once in a supergame (as revealed by the bubbles located on the diagonal in the figure), the extent of strategic errors is non-trivial, with 24% of subjects making errors in at least 3 out of 17 relevant supergames (those lasting more than 1 round) suggesting that some of the dominated CaD behavior could instead be due to inattention or a lack of strategic understanding of the game. While the prevalence of such strategic errors (CaD) is relatively small, it nevertheless complicates the interpretation of the deviations from theoretically optimal behavior.

Finding 3. *A majority of subjects (52%) made at least one strategic error of choosing to cooperate after defecting earlier within the same supergame (CaD), i.e., after triggering a "grim" response. Overall, suboptimal, excessive cooperation amounts to 7.58% of relevant observations, with 24% of subjects making dominated choices in at least 3 out of 17 relevant supergames.*

4.4 Overall Choices Across All Supergames

Across all 48 choices in all 24 supergames, the overall optimal choice counts are significantly short of the theoretical prediction of 48, with a mean (st.dev.) of 31.05 (8.06) (see Figures 1, top right, and C2, bottom right). The mean (st.dev.) of the overall count of cooperative choices is only 24.09 (11.24), which is marginally lower than the theoretical prediction of 26 (one-sided $t = 1.670$, p-value=0.046) (see Figures 1, top left, and C2, top left).

While subjects start by cooperating excessively in the first rounds, the mean (st.dev.) of cooperation counts in the subsequent rounds (given by the difference in the overall and first round cooperation counts) is only 12.56 (6.29), which is significantly lower than the theoretical prediction of 18 (p-value=0.000). Note that this is despite the excessive cooperation of 1.82 counts per subject on average due to strategic errors (CaD).

Finding 4. *Compared with theoretical predictions, on average, subjects cooperate too much at the beginning of supergames with* $\delta < 0.5$ *and stop cooperating too early in supergames with* $\delta > 0.5$. *Combined with an average of 1.82 strategic errors (CaD) per subject, only 64.69% of all choices are theoretically optimal.*

Figure 3 shows the heterogeneity of behavioral patterns within each supergame. Only two subjects (represented by solid diamonds) behaved fully theoretically optimally, and four more (represented by hollow diamonds) were only two supergames away from full optimality. Similarly, there are two solid circles and one hollow circle representing subjects who always and almost always cooperated, as well as one solid triangle and four hollow triangles representing subjects who always and almost always defected. (See also Figure C2, top right).

Finding 5. *There is a notable heterogeneity in subjects' choices to cooperate or defect. Only two out of 100 subjects always followed the theoretically optimal strategy, and four more did so in 22 out of 23 supergames. Two (one) subjects are fully biased towards cooperation (defection), while one (four) are biased in 22 out of 23 supergames. The rest of the subjects appear to pursue strategies that are neither theoretically optimal nor purely biased.*

4.5 Overall Point Totals

As Figure 2 (right) shows, strategic errors (CaD) reduce the overall total of awarded points (i.e., the sum of point earnings across all 48 decisions). In fact, 16 subjects could have achieved a strictly higher total of 3600 points by choosing either to always cooperate (All-C) or always defect (All-D), as by design, the expected payoff to these two extremely biased strategies is the same. The empirical range of overall point totals is $[3285, 4185]$ points, with a mean (st.dev.) of 3835.05 (203.38). Note that the mode is at the theoretically optimal point total of 4050, and that the maximum point total is still higher. Overall, 19 subjects were able to achieve at least this optimal point value, despite only 2 subjects behaving in a way that was fully theoretically optimal. Thus, given the random realization of the supergames, some subjects were able to achieve at least as much as the theoretical payoff despite pursuing strategies that were not theoretically optimal.

Finding 6. *16% of subjects earned less than what they could have achieved by either always cooperating or always defecting. 17% of subjects were able to achieve an overall payoff at least as high as the expected theoretical payoff without following the theoretically optimal strategy.*

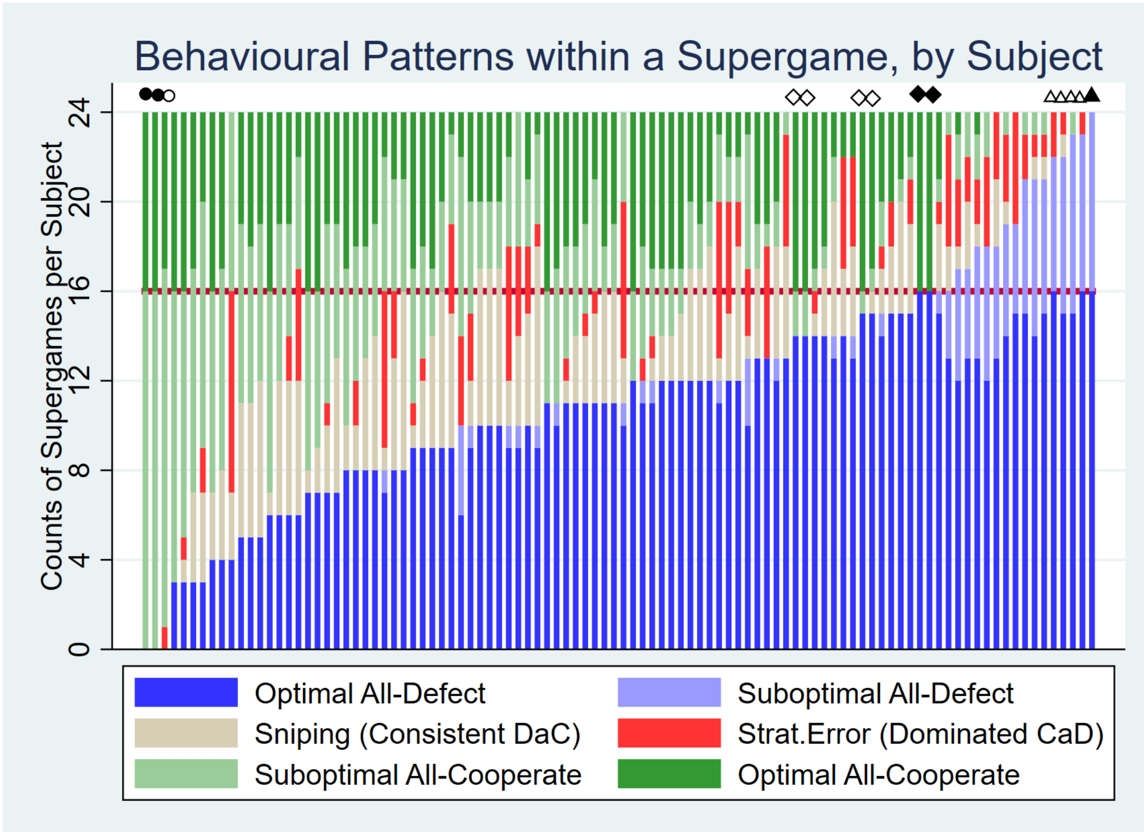

Figure 3: Subject heterogeneity in patterns of choices within supergames, out of 24 supergames, by subject, ordered by the count of supergames with All-Defect choices (100 subjects total). The theoretically optimal strategy involves always defecting in 16 supergames and always cooperating in the remaining 8 supergames (represented by a horizontal line). Diamonds, circles, and triangles represent subjects who made optimal, All-C, and All-D decisions, respectively, in at least 22 supergames, with solid markers representing those making such decisions in all 24 supergames.

4.6 "Sniping"

Note that when playing against a robot known to play the Grim trigger strategy, one could achieve up to 4680 points, a much higher point total than under the theoretically optimal strategy, *if one knew in advance exactly when each supergame would end* by always cooperating prior to the final round of a supergame, and defecting in the final round. Such a "sniping" strategy would allow one to earn the temptation payoff without triggering the "grim" punishment as the supergame ends.[10] Of course, subjects did not have such knowledge in our experiment, but they may have formed some expectations as to when a sequence (supergame) might end in an effort to employ such a strategy.

[10]While we refer to this type of behavior as sniping borrowing terminology from the auction literature, it is also an instance of "gambler's fallacy," Cowan (1969). This is the erroneous belief that the probability of some event is lowered when that event has occurred recently even though the probability of that event is known to be independent from one instance to the next.

Indeed, Figure 1 (bottom left) shows that, for some δs, some subjects defect for the first time (thus triggering subsequent defection by the automated opponent) *later* in the sequence, rather than in the first round (if ever) as predicted by the theory. [11] We hypothesize that this pattern of behavior could be due to some subjects using a "sniping" strategy, which we define as consistently defecting after the earlier play of cooperation in the same supergame, or (DaC) for short. Such a sniping strategy is *risky*, as it is most profitable if the first defection happens in the final round of the supergame.

As Figure 1 (bottom right) shows, the shares of the supergames where subjects always defected (All-D) are declining as δ increases. However, this does not translate into an increase in the prevalence of the always cooperate (All-C) strategy as delta increases. Instead, as δ (and thus the expected duration of a supergame) increases, both the prevalence of strategic errors (CaD) and "sniping" (consistent DaC) strategies increases. Note that interpreting All-C strategies is complicated by attrition, as a subjects might have *intended* to snipe, but a supergame ended earlier than expected. Similarly, All-D strategies in low δ supergames could be not only due to theoretically optimal behavior, but also due to the intended use of a sniping strategy.

Indeed, if the behavior were theoretically optimal, then in Table 1 (specifications 1-2) the coefficients on $\delta = \{0.25, 0.33, 0.4\}$ would have been insignificantly different from the baseline of $\delta = 0.1$, and would only be significantly different for $\delta = \{0.67, 0.7\}$. In addition, the round dummies would all be insignificantly different from the baseline of the first round. Instead, as this table demonstrates, subjects' tendency to choose cooperation increases with δ, and decreases with the round number, consistent with the use of the sniping strategy.

Finding 7. *Some subjects appear to use a "sniping" strategy, by trying to time their first defection with the unknown final round of a supergame. While following this strategy could lead to a higher payoff, only a few subjects earned more than the theoretically optimal payoff.*

4.7 The Effect of Cognitive Abilities

As noted earlier, we asked all of our subjects to answer 7 cognitive reflection test (CRT) questions as part of the study. Subjects' total score on this 7-item, cognitive reflection test, CRT7, was used as a proxy for cognitive ability. The mean (st.dev.) of the CRT7 score was 3.78 (2.26) with a median of 4.

As Table 2 shows, CRT7 predicts the rational aspects of subjects' behavior rather well. Specifically, it is positively correlated with their overall point payoff total (specifications 1-2) and negatively with a count of the number of supergames with strategic errors (dominated

[11] Note that the expected final round $\frac{1}{1-\delta}$ (as calculated from the perspective of round 1), as well as the average realized final round increase with δ - see Table A1. Mengel et al. (2021) report that subjects respond to the *realized* supergame length, and are more likely to cooperate when they have experienced supergames of longer duration.

Cooperate/Marginals	All				$CRT7 \leq 4$		$CRT7 > 4$	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
δ=0.25	0.23***	0.22***	0.22***	0.22***	0.25***	0.25***	0.18***	0.18***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)	(0.05)	(0.05)
δ=0.33	0.31***	0.31***	0.31***	0.31***	0.31***	0.31***	0.29***	0.29***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.05)	(0.05)	(0.05)	(0.05)
δ=0.4	0.47***	0.47***	0.47***	0.47***	0.45***	0.45***	0.47***	0.47***
	(0.04)	(0.04)	(0.04)	(0.04)	(0.05)	(0.05)	(0.06)	(0.06)
δ=0.67	0.66***	0.66***	0.66***	0.66***	0.62***	0.61***	0.69***	0.69***
	(0.04)	(0.04)	(0.04)	(0.04)	(0.06)	(0.06)	(0.06)	(0.06)
δ=0.7	0.69***	0.69***	0.69***	0.69***	0.62***	0.62***	0.75***	0.74***
	(0.04)	(0.04)	(0.04)	(0.04)	(0.06)	(0.06)	(0.06)	(0.06)
Round 2	-0.04	-0.03	-0.04	-0.03	-0.00	-0.00	-0.09*	-0.09*
	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.04)	(0.04)
Round 3	-0.12***	-0.12***	-0.12***	-0.12***	-0.06	-0.06	-0.21***	-0.20***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)	(0.05)	(0.05)
Round 4	-0.19***	-0.18***	-0.18***	-0.18***	-0.10	-0.10	-0.29***	-0.29***
	(0.04)	(0.04)	(0.04)	(0.04)	(0.06)	(0.06)	(0.06)	(0.06)
Round 5	-0.17**	-0.17**	-0.17**	-0.17**	-0.13	-0.13	-0.23**	-0.23**
	(0.05)	(0.05)	(0.05)	(0.05)	(0.07)	(0.07)	(0.08)	(0.08)
Supergame	-0.00**	-0.00**	-0.00**	-0.00**	-0.00**	-0.00**	-0.00	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Order Long	0.05	0.04	0.05	-0.00	0.05	-0.04	0.04	0.01
	(0.04)	(0.10)	(0.04)	(0.10)	(0.06)	(0.08)	(0.06)	(0.19)
Prior Defect	-0.20***	-0.20***	-0.20***	-0.20***	-0.17***	-0.17***	-0.24***	-0.24***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)	(0.05)	(0.05)
Female			-0.12**	-0.15*	-0.11	-0.11	-0.13*	-0.15
			(0.04)	(0.06)	(0.07)	(0.11)	(0.06)	(0.11)
Age			-0.00	-0.00	-0.01	-0.00	-0.00	0.00
			(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
CRT7			0.00	0.00	0.01	0.01	0.03	0.02
			(0.01)	(0.01)	(0.02)	(0.02)	(0.03)	(0.03)
Controls	No	Yes	No	Yes	No	Yes	No	Yes
chi2	266.09	276.82	318.74	318.47	175.91	188.92	163.70	180.21
p	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
N	4800	4800	4800	4800	2688	2688	2112	2112

Table 1: Choices to cooperate: mixed-effects probit regressions, marginals (dy/dx), robust errors in parentheses. "Supergame" stands for the number of the supergame in the sequence, "Order Long" stands for the order with the first supergame involving $\delta = 0.67$, "Ever Defected" stands for whether the subject defected in prior rounds of a given supergame. Controls stands for session controls. Chi2 and corresponding p-values are for the odds regressions (see Table D2). (Significance * 0.05 ** 0.01 *** 0.001.)

CaD) (specifications 3-4). Interestingly, neither the count of theoretically optimal choices (specifications 5-6), nor the count of apparent sniping (consistent DaC) (specifications 7-8) are correlated with the CRT7 score. By construction, these two counts are *mutually exclusive*. However, as discussed above, some choices which are consistent with following the theoretically optimal strategy, could instead be part of an intended sniping strategy. Thus, it is not a surprise that the combined count of whether subjects follow the theoretically optimal strategy (which involves either All-D for $\delta < 0.5$ or All-C for $\delta > 0.5$), or engage in sniping (which involves consistent defection after cooperation, or DaC) *is* correlated with the CRT7 score (see specifications 9-10).

Furthermore, as the specifications 3-4 in Table 1 show, the CRT7 score does not correlate with the choice to cooperate or defect in all 48 rounds of prisoners' dilemma. Note that while the coefficient on the female dummy in that specification is significantly negative, and the CRT7 score is negatively correlated with being female ($r = -0.2365, pvalue =$

	Overall Point Totals (OLS)		Dominated(CaD) (Tobit, ll=0)		Theor.Optimal (Tobit, ul=24)		Sniping(DaC) (Tobit, ll=0)		Th.Opt.+Snipe(DaC) (Tobit, ul=24)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Order Long	57.11	84.46	-0.22	0.66	0.87	2.33	0.34	0.12	1.01	2.13
	(38.27)	(78.57)	(0.68)	(1.22)	(0.79)	(1.37)	(0.59)	(1.08)	(0.78)	(1.56)
Female	-65.92	-21.23	1.40*	0.92	0.09	1.52	-0.88	-0.63	-0.52	1.13
	(41.98)	(53.92)	(0.69)	(0.80)	(0.80)	(0.96)	(0.59)	(0.78)	(0.82)	(1.15)
Age	-4.85	-5.42	0.12	0.13	0.25	0.20	-0.24	-0.23	0.07	0.04
	(9.17)	(9.72)	(0.14)	(0.12)	(0.18)	(0.17)	(0.14)	(0.14)	(0.18)	(0.18)
CRT7	26.37**	26.25**	-0.51**	-0.52**	0.32	0.29	0.11	0.14	0.46*	0.47**
	(9.32)	(9.28)	(0.16)	(0.16)	(0.19)	(0.18)	(0.14)	(0.14)	(0.19)	(0.17)
Constant	3845.38***	3806.58***	-1.03	-1.25	7.46	6.76	7.47*	6.73*	13.78**	12.38**
	(212.24)	(235.42)	(3.06)	(2.91)	(3.95)	(4.00)	(3.09)	(3.24)	(4.19)	(4.42)
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
F	5.47	2.80	4.52	2.78	1.44	2.26	1.80	0.93	2.72	4.04
p	0.00	0.00	0.00	0.00	0.23	0.02	0.14	0.51	0.03	0.00
Observations	100	100	100	100	100	100	100	100	100	100

Table 2: Individual differences in rationality. (Significance: * 0.05 ** 0.01 *** 0.001).

0.0178), the coefficient on the CRT7 score remains insignificant if we exclude the age and gender demographic variables (results available on request). However, when we split the sample according based on the median CRT7 score, we find that subjects with relatively high proxies for cognitive ability (CRT7>4), tend to respond more strongly to the continuation probability δ (particularly for the highest δ value 0.7) and exhibit a stronger tendency for following the "sniping" strategy, as their coefficients on the round dummies tend to be more strongly significant. Subjects with relatively lower proxies for cognitive ability (CRT7≤ 4) do not exhibit any systematic sensitivity to the round number.

Finding 8. *Subjects with higher proxy values for cognitive ability are more likely to behave in a payoff-maximizing fashion, and engage in sniping.*

5 Conclusions

In this paper, we report on an experiment in which subjects play the repeated prisoner's dilemma against a robot player known to play the Grim trigger strategy. This design converts the original strategic situation into a single-person decision problem, for which there is a unique optimal strategy. We use a within-subject design in which subjects play many different supergames with differing continuation probabilities. We can therefore identify systematic errors made by subjects and relate them to individual characteristics, and, in particular, cognitive abilities. Our novel design has revealed several important and interesting findings. First, first-round cooperation strongly increases with increases in δ, more so than in human-to-human studies, suggesting that strategic uncertainty may play an important role in reducing cooperation rates in those studies. Second, we find that a majority (52%) of our subjects make at least one strategic error of cooperating after defection. Third, some subjects employ a sniping strategy, consistently defecting after initially choosing to cooperate (DaC) in the same supergame that can yield them higher payoffs than the theoretically optimal strategy. Finally, we show that these different behaviors are correlated with our proxy measure for cognitive ability. We hope that these findings will be useful in differentiating intentional strategies from errors in repeated games more generally.

References

Ackerman, R. (2014): "The diminishing criterion model for metacognitive regulation of time investment." *Journal of Experimental Psychology: General*, 143, 1349.

Andreoni, J. and L. Samuelson (2006): "Building rational cooperation," *Journal of Economic Theory*, 127, 117–154.

Bernheim, B. D. and O. Stark (1988): "Altruism within the family reconsidered: Do nice guys finish last?" *The American Economic Review*, 1034–1045.

Camerer, C. F. (2003): *Behavioral game theory: Experiments in strategic interaction*, Princeton University press.

Charness, G. and D. Levin (2009): "The origin of the winner's curse: a laboratory study," *American Economic Journal: Microeconomics*, 1, 207–36.

Charness, G., R. Oprea, and S. Yuksel (2021): "How do people choose between biased information sources? Evidence from a laboratory experiment," *Journal of the European Economic Association*, 19, 1656–1691.

Chaudhuri, A. (2008): *Experiments in economics: Playing fair with money*, Routledge.

Chen, D. L., M. Schonger, and C. Wickens (2016): "oTree—An open-source platform for laboratory, online, and field experiments," *Journal of Behavioral and Experimental Finance*, 9, 88–97.

Cooper, D. J. and J. H. Kagel (2021): "Infinitely repeated prisoner dilemma games: Comparing teams with individuals," *working paper*.

Cowan, J. L. (1969): "The gambler's fallacy," *Philosophy and Phenomenological Research*, 30, 238–251.

Dal Bó, P. and G. R. Fréchette (2018): "On the determinants of cooperation in infinitely repeated games: A survey," *Journal of Economic Literature*, 56, 60–114.

Duffy, J., E. Hopkins, and T. Kornienko (2021): "Lone wolf or herd animal? Information choice and learning from others," *European Economic Review*, 134, 103690.

Duffy, J. and F. Muñoz-García (2012): "Patience or fairness? Analyzing social preferences in repeated games," *Games*, 3, 56–77.

Duffy, J. and H. Xie (2016): "Group size and cooperation among strangers," *Journal of Economic Behavior & Organization*, 126, 55–74.

Frederick, S. (2005): "Cognitive reflection and decision making," *Journal of Economic Perspectives*, 19, 25–42.

Mailath, G. J. and L. Samuelson (2006): *Repeated games and reputations: long-run relationships*, Oxford University Press.

Mengel, F., S. Weidenholzer, and L. Orlandi (2021): "Match Length Realization and Cooperation in Indefinitely Repeated Games," *Available at SSRN 3777155*.

Murnighan, J. K. and A. E. Roth (1983): "Expecting continued play in prisoner's dilemma games: A test of several models," *Journal of Conflict Resolution*, 27, 279–300.

Proto, E., A. Rustichini, and A. Sofianos (2019): "Intelligence, personality, and gains from cooperation in repeated interactions," *Journal of Political Economy*, 127, 1351–1390.

Romero, J. and Y. Rosokha (2018): "Constructing strategies in the indefinitely repeated prisoner's dilemma game," *European Economic Review*, 104, 185–219.

Roth, A. E. and J. K. Murnighan (1978): "Equilibrium behavior and repeated play of the prisoner's dilemma," *Journal of Mathematical Psychology*, 17, 189–198.

Toplak, M. E., R. F. West, and K. E. Stanovich (2014): "Assessing miserly information processing: An expansion of the Cognitive Reflection Test," *Thinking & Reasoning*, 20, 147–168.

Appendices (Not Intended for Publication)

A Description of Supergames

Delta	Duration		Duration (Rounds)					Number of	
δ	Expected $\left(\frac{1}{1-\delta}\right)$	realized (Ave.)	1	2	3	4	5	Supergames	Choices
.1	1.11	1.00	4	0	0	0	0	4	4
.25	1.33	1.25	3	1	0	0	0	4	5
.33	1.49	1.50	2	2	0	0	0	4	6
.4	1.67	1.75	2	1	1	0	0	4	7
.67	3.03	3.00	0	2	0	2	0	4	12
.7	3.33	3.50	0	1	1	1	1	4	14
Total Supergames			11	7	2	3	1	24	
Total Choices			24	13	6	4	1		48

Table A1: The distribution of the supergames, split by continuation probability δ. The average theoretical and realized supergame durations are 1.99 rounds and 2 rounds, respectively.

B Order Effects

As was documented recently by Mengel et al. (2021) early exposure to relatively long sequences could affect subsequent behavior in the prisoner's dilemma. While the mean (st.dev.) first round per-subject counts of cooperation in the long and reverse orders are 10.96 (6.48) and 12.10 (4.86), respectively (out of 24), this difference is insignificant ($t = 1.00$, Kolmogorov-Smirnov one-sided p-value= 0.278). The corresponding mean (st.dev.) overall counts are, respectively, 25.52 (9.29) and 22.66 (12.83) (out of 48), with the difference remaining insignificant ($t = 1.28$, Kolmogorov-Smirnov one-sided p-value= 0.198).

As for the optimal choices, the first round counts are higher in the long treatment, with mean (st.dev.) being, respectively, 16.2 (3.49) and 17.42 (3.91), but this difference is marginally significant only according to the Kolmogorov-Smirnov test (one-sided p-value= 0.034), but not according to t-test ($t = 1.65$, p-value= 0.051). The overall optimal choice counts are, again, higher in the long order treatment (with mean (st.dev.) of 32.54 (8.16) in long order, and 29.56 (7.76) in reverse), but this is marginally significant only according to t-test ($t = 1.87$, p-value= 0.032), but not according to Kolmogorov-Smirnov test (one-sided p-value= 0.135. The order effect disappears in mixed effects panel regressions in Table 1.

C More on Subject Heterogeneity

As Figures C1 and C2 show, there is a significant heterogeneity is subjects' behavior, without any clear "representative" pattern. The top right panels of each figure provide two-dimensional distributions of the cooperative and optimal choices, where the possible choice combinations are restricted to the polygons delineated by the dashed lines. The shape of the polygon for the overall choices in Figure C2 is due to the possibility of dominated choices).

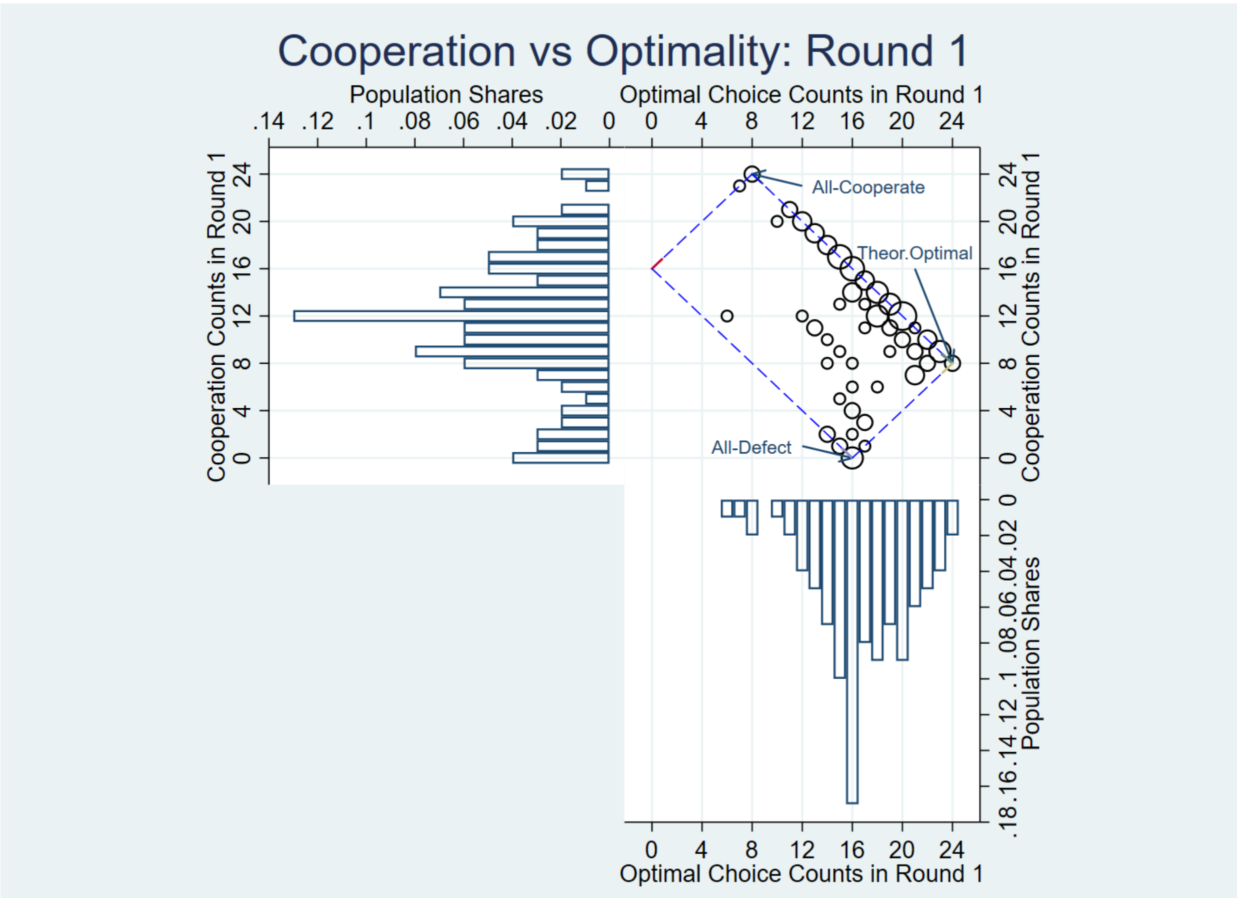

Figure C1: Distributions of counts of cooperation (top left) and of optimal choices (bottom right), in the first rounds across all 24 supergames. The top right panel combines these two distributions, with bubble size proportional to the share of subjects, 100 subjects total.

As Figure C1 shows, subjects tend to excessively cooperate in the first round of each supergame, with the largest number of subjects choosing to start half of the supergames by cooperating. This early excessive cooperation in the first round is followed by the subsequent defection, represented in Figure C2 (top left and bottom right panels). As the top right panel of Figure C2 shows, there are three equally sized clusters at each of the three corners of the polygon. There are only two subjects at the far right corner, who behaved fully theoretically optimally (all 48 rounds) and three subjects who made only 2 theoretically suboptimal choices. In the top corner, there are two subjects who always cooperated, and two subjects who defected only once and three times, respectively. In the bottom corner, there is a single subject who always defected, and one, two, and one subjects who cooperated once, twice, and three times, respectively. The presence of strategic errors (CaD) complicates the interpretation of the remaining 85% of subjects, most of whom are located away from the boundaries, in the center of the figure. Many of those observations represent the overall early excessive cooperation in the first rounds followed by the subsequent defections within a supergame, possibly due to some form of a sniping strategy.

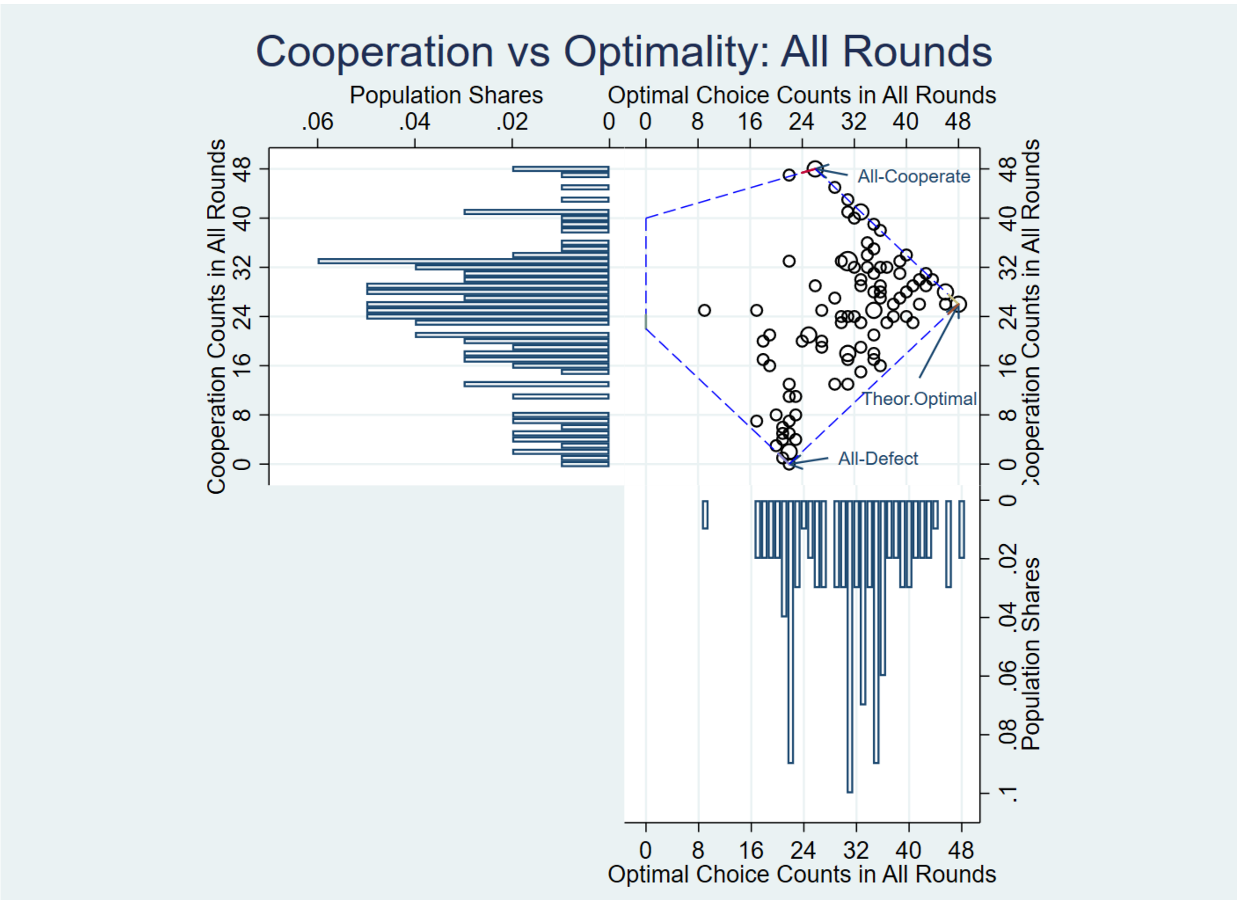

Figure C2: Distributions of counts of cooperation (top left) and of optimal choices (bottom right), in all rounds across all 24 supergames. The top right panel combines these two distributions, where the possible combinations of choices are restricted to the polygons delineated by the dashed lines, and the bubble size proportional to the share of subjects, 100 subjects total.

D Further Results

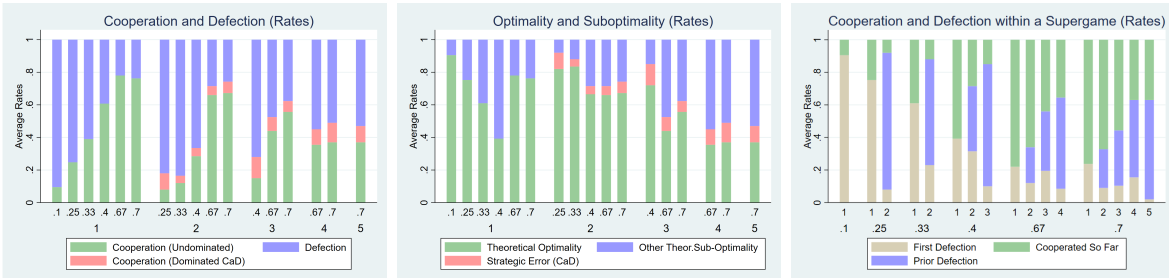

Figure D3: Patterns of cooperation and optimality (rates) (total 2,400 observations of supergames, for all 100 subjects). Shares of observations of cooperation vs defection (left) and optimality vs suboptimality (middle), split by round and δ. Right: average per-subject rates of first defection within a supergame for each round of each δ. (See Figure 1 for choice counts.)

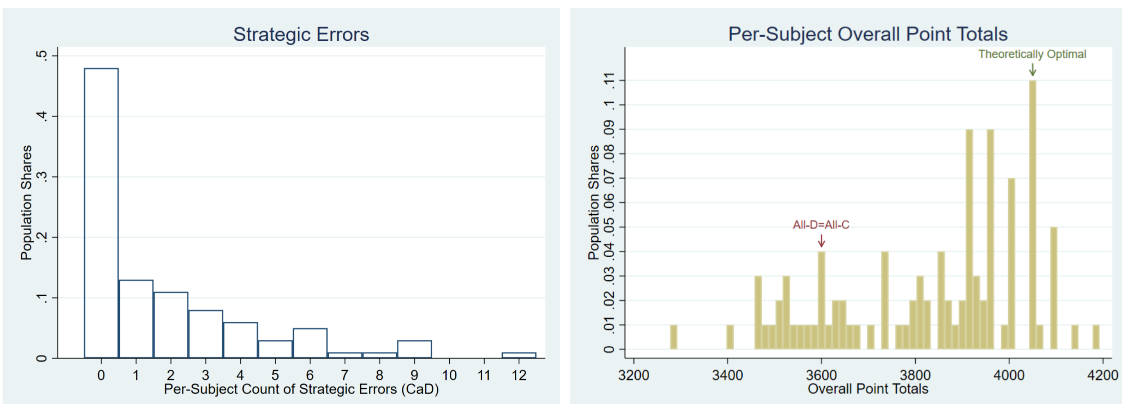

Figure D4: Left: Distribution of overall per-subject counts of instances of cooperation after defection (CaD). Right: Overall per-subject counts of instances of cooperation after defection (CaD) versus count of supergames (among 17 relevant supergames) where such strategic errors (dominated choices) happen. Bubble size is proportional to the share of subjects, number of subjects in the bubbles, 100 subjects total. Right: Distribution of overall payoffs. (See related Figure D4.)

Cooperate/Odds	All				$CRT7 \leq 4$		$CRT7 > 4$	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
δ=0.25	0.90***	0.90***	0.90***	0.90***	0.93***	0.93***	0.87***	0.88***
	(0.14)	(0.14)	(0.14)	(0.14)	(0.17)	(0.17)	(0.24)	(0.24)
δ=0.33	1.23***	1.23***	1.23***	1.23***	1.17***	1.17***	1.36***	1.36***
	(0.15)	(0.15)	(0.15)	(0.15)	(0.20)	(0.20)	(0.25)	(0.25)
δ=0.4	1.88***	1.88***	1.88***	1.88***	1.67***	1.67***	2.22***	2.23***
	(0.17)	(0.17)	(0.17)	(0.17)	(0.20)	(0.20)	(0.33)	(0.33)
δ=0.67	2.65***	2.65***	2.65***	2.65***	2.30***	2.30***	3.29***	3.30***
	(0.21)	(0.21)	(0.21)	(0.21)	(0.25)	(0.25)	(0.40)	(0.40)
δ=0.7	2.76***	2.76***	2.76***	2.76***	2.33***	2.33***	3.54***	3.54***
	(0.22)	(0.22)	(0.22)	(0.22)	(0.26)	(0.26)	(0.44)	(0.44)
Round 2	-0.14	-0.14	-0.14	-0.14	-0.02	-0.01	-0.43*	-0.43*
	(0.10)	(0.10)	(0.10)	(0.10)	(0.12)	(0.12)	(0.20)	(0.20)
Round 3	-0.47***	-0.47***	-0.47***	-0.47***	-0.22	-0.22	-0.97***	-0.97***
	(0.14)	(0.14)	(0.14)	(0.14)	(0.17)	(0.17)	(0.26)	(0.26)
Round 4	-0.74***	-0.74***	-0.74***	-0.74***	-0.38	-0.38	-1.38***	-1.37***
	(0.17)	(0.17)	(0.17)	(0.17)	(0.21)	(0.21)	(0.30)	(0.30)
Round 5	-0.69**	-0.69**	-0.69**	-0.69**	-0.48	-0.48	-1.09**	-1.09**
	(0.22)	(0.22)	(0.22)	(0.22)	(0.28)	(0.28)	(0.40)	(0.40)
Supergame	-0.01**	-0.01**	-0.01**	-0.01**	-0.01**	-0.01**	-0.01	-0.01
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)
Order Long	0.18	0.15	0.19	-0.01	0.18	-0.15	0.18	0.03
	(0.17)	(0.39)	(0.17)	(0.38)	(0.21)	(0.32)	(0.26)	(0.91)
Prior Defect	-0.81***	-0.81***	-0.80***	-0.81***	-0.65***	-0.65***	-1.13***	-1.13***
	(0.12)	(0.12)	(0.12)	(0.12)	(0.14)	(0.15)	(0.22)	(0.22)
Female			-0.47*	-0.60*	-0.41	-0.41	-0.60*	-0.73
			(0.18)	(0.26)	(0.26)	(0.41)	(0.30)	(0.54)
Age			-0.02	-0.01	-0.02	-0.02	-0.01	0.00
			(0.03)	(0.03)	(0.04)	(0.05)	(0.05)	(0.05)
CRT7			0.01	0.00	0.05	0.04	0.12	0.07
			(0.04)	(0.03)	(0.08)	(0.08)	(0.14)	(0.15)
Constant	-1.61***	-1.55***	-1.03	-0.78	-0.85	-0.78	-2.07	-1.67
	(0.19)	(0.39)	(0.66)	(0.73)	(1.02)	(1.21)	(1.21)	(1.28)
Controls	No	Yes	No	Yes	No	Yes	No	Yes
chi2	266.09	276.82	318.74	318.47	175.91	188.92	163.70	180.21
p	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
N	4800	4800	4800	4800	2688	2688	2112	2112

Table D2: Choices to cooperate: mixed-effects probit regressions, odds, robust errors in parentheses. "Supergame" stands for the number of the supergame in the sequence, "Order Long" stands for the order with the first supergame involving $\delta = 0.67$, "Ever Defected" stands for whether subject defected in prior rounds of a given supergame. Controls stands for session controls. Chi2 and corresponding pvalues are for the odds regressions. (Significance * 0.05 ** 0.01 *** 0.001.) (See the marginals in Table 1.)

E Appendix: Experimental Design

CRT questions

Subjects were asked to provide numerical answers to the following cognitive reflection test (CRT) questions.

1. The ages of Anna and Barbara add up to 30 years. Anna is 20 years older than Barbara. How old is Barbara?
2. If it takes 2 nurses 2 minutes to check 2 patients, how many minutes does it take 40 nurses to check 40 patients?
3. On a loaf of bread, there is a patch of mold. Every day, the patch doubles in size. If it takes 24 days for the patch to cover the entire loaf of bread, how many days would it take for the patch to cover half of the loaf of bread?
4. If John can drink one barrel of water in 6 days, and Mary can drink one barrel of water in 12 days, how many days would it take them to drink one barrel of water together?
5. A man buys a pig for $60, sells it for $70, buys it back for $80, and sells it finally for $90. How much profit has he made, in dollars?
6. Jerry received both the 15th highest and the 15th lowest mark in the class. How many students are in the class?
7. A turtle starts crawling up a 6-yard-high rock wall in the morning. During each day it crawls 3 yards and during the night it slips back 2 yards. How many days will it take the turtle to reach the top of the wall?

Repeated PD Game Instructions

You will participate in 24 sequences. Each sequence consists of one or more rounds.

In each round, you play a game.

Specifically, you will have to choose between action X or action Y. Your opponent also chooses between action X or action Y.

The combination of your action choice and that of your opponent results in one of the four cells shown in the payoff table below (which will be the same table in each round).

	X	Y
X	**75**, *75*	**15**,*120*
Y	**120**, *15*	**30**, *30*

In this table, the rows refer to your action and the columns refer to your opponent's actions. The first number in each cell (in bold) is your payoff in points and the second number in each cell (in italics) is your opponent's payoff in points. Thus for example, if you choose X and your opponent chooses Y, then you earn 15 points and your opponent earns 120 points.

In all 24 sequences, you will play this game against the computer. That is, your opponent is a computer program.

The rule the computer follows in choosing between action X or Y is this:

- In the first round of each sequence the computer will always choose X.
- Starting from the second round of each sequence, the computer's choice will be completely determined by your previous choices in that sequence:
  - If you have ever chosen Y in previous round(s) of the current sequence, the computer will choose Y in all remaining rounds of the current sequence.
  - Otherwise, the computer will choose X.

There is no randomness in the computer's choice, and its choice does not depend on your choices in any sequences other than the current one.

After choices are made by you and the computer, you learn the results of the round, specifically, your point earnings and those earned by the computer. A random number generator is used to determine whether the current sequence continues on with another round, or if the current round is the last round of the sequence.

Whether the sequence continues with another round or not depends on the probability (or chance) of continuation for the sequence. This continuation probability for a sequence is prominently displayed on your decision screen and remains constant for all rounds of a given sequence. However, this continuation probability can change at the start of each new sequence, so please pay careful attention to announcements about the continuation probability for each new sequence. Whether a sequence continues depends on whether at the end of a round the random number generator drew a number in the interval [1,100] that is less than or equal to the continuation probability (in percent).

For example, if the continuation probability in a sequence is 40%, then, after round 1 of the sequence, which is always played, there is a 40% chance that the sequence continues on to round 2 and a 60% chance that round 1 is the last round of the sequence. Whether continuation occurs depends on whether the random number generator drew a number from 1 to 100 that is less than or equal to 40. If it did, then the sequence continues on to round 2. If it did not, then round 1 is the final round of the sequence. If the sequence continues on to round 2, then after that round is played, there is again a 40% chance that the sequence continues on to round 3 and a 60% chance that round 2 is the last round of the sequence, again determined by the random number generator for that round. And so on.

Thus, the higher is the continuation probability (chance), the more rounds you should expect to play in the sequence. But since the continuation probability is always less than 100%, there is no guarantee that any sequence continues beyond round 1.

At the end of the experiment, you will be paid your point earnings from six sequences, randomly selected so that each selected sequence has a different continuation probability. Each point you earn over all rounds in each of the 6 randomly selected sequences is worth $0.01 in US dollars, that is, the greater are your point earnings, the greater are your money earnings.

Comprehension quiz

Now that you have read the instructions, before proceeding, we ask that you answer the following comprehension questions. For your convenience, we repeat the payoff table below, which you will need to answer some of these questions. In this table, the rows indicate your choice and the columns indicate the computer's choices.

	X	Y
X	**75**, *75*	**15**,*120*
Y	**120**, *15*	**30**, *30*

The first number in each cell (in bold) is your payoff in points and the second number in each cell (in italics) is the computer's payoff in points.

Questions

1. If, in a round, you chose X and the computer program chose X, what is your payoff in points for the round? What is the computer program's payoff?

2. If, in a round, you chose Y and the computer program chose X, what is your payoff in points for the round? What is the computer program's payoff?

3. If, in a round, you chose Y and the computer program chose Y, what is your payoff in points for the round? What is the computer program's payoff?

4. If you have chosen Y in any prior round of the current sequence, what will the computer program choose in the current round of the sequence? Choose: X or Y

5. True or false: At the start of each sequence, you will know exactly how many rounds will be played in the sequence. Choose: True or False

6. True or false: If, in a sequence, the continuation probability is 75%, then you can expect that there will be more rounds in that sequence, on average, than in a sequence with a continuation probability of 25%. Choose True or False

Repeated PD Games

After a subject had successfully completed all quiz questions, the experiment proceeded on to the first indefinitely repeated PD game. For each such game, subjects were instructed clearly about the continuation probability for that repeated game. For instance, Figures E5-E9 show illustrative screenshots from the first indefinitely repeated game of the "orderlong" treatment. Table E3 reports on the continuation probability for each of the 24 sequences along with the actual number of rounds played for the two treatment orders.

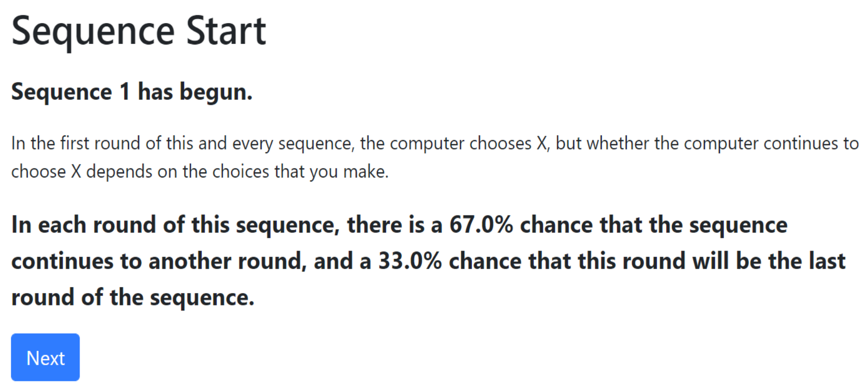

Figure E5: Start screen for a new sequence

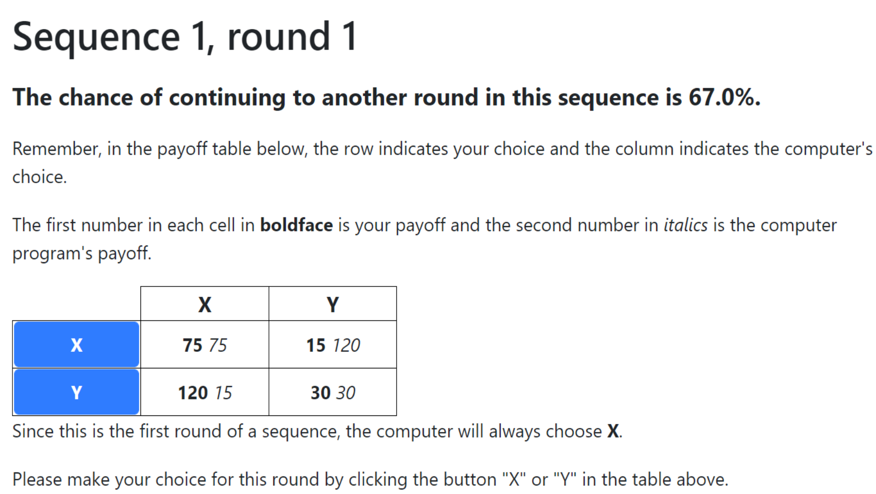

Figure E6: Main decision screen for a period in the sequence

Results of sequence 1, round 1

You chose Y this round.

Following its rule, the computer has chosen X.

Therefore, your payoff this round is 120.0 points.

Based on the random number drawn, sequence 1 will CONTINUE with another round.

Next

History of Rounds in this Sequence

Sequence	Chance to Continue	Round	Your choice	Computer's Choice	Payoff
1	0.67	1	Y	X	120.0 points

Figure E7: Results screen for a period in the sequence

Sequence 1, round 2

The chance of continuing to another round in this sequence is 67.0%.

Remember, in the payoff table below, the row indicates your choice and the column indicates the computer's choice.

The first number in each cell in **boldface** is your payoff and the second number in *italics* is the computer program's payoff.

	X	Y
X	**75** *75*	**15** *120*
Y	**120** *15*	**30** *30*

Based on your choices in previous rounds of this sequence, the computer will choose **Y.**

Please make your choice for this round by clicking the button "X" or "Y" in the table above.

History of Rounds in this Sequence

Sequence	Chance to Continue	Round	Your choice	Computer's Choice	Payoff
1	0.67	1	Y	X	120.0 points

Figure E8: Decision screen for a continuation period in the sequence, noting what the robot player will do, based on the history of play

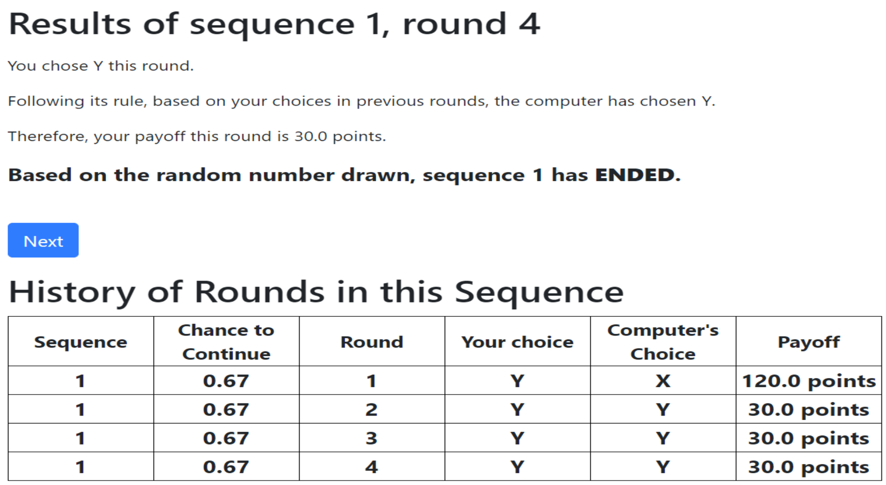

Figure E9: Screen for the final period of a sequence noting that based on the random drawn, the sequence has ended.

	OrderShort		OrderLong	
Sequence	$p =$	No.Rounds	$p =$	No. Rounds
1	0.33	1	0.67	4
2	0.7	4	0.33	1
3	0.1	1	0.4	2
4	0.67	2	0.25	1
5	0.4	3	0.7	3
6	0.7	2	0.33	2
7	0.25	1	0.7	5
8	0.33	2	0.4	1
9	0.67	4	0.67	2
10	0.4	1	0.1	1
11	0.1	1	0.25	1
12	0.25	2	0.1	1
13	0.1	1	0.25	2
14	0.25	1	0.1	1
15	0.1	1	0.4	1
16	0.67	2	0.67	4
17	0.4	1	0.33	2
18	0.7	5	0.25	1
19	0.33	2	0.7	2
20	0.7	3	0.4	3
21	0.25	1	0.67	2
22	0.4	2	0.1	1
23	0.33	1	0.7	4
24	0.67	4	0.33	1
Totals		48		48

Table E3: Continuation probabilities, p, and the number of rounds played for each of the 24 sequences, both treatment orders (one order is just the reverse of the other).