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

John Duffy† Ed Hopkins‡ Tatiana Kornienko§

January 19, 2026

Abstract

Cooperation in repeated interactions is important for much socio-economic activity. In this paper, we put subjects in the simplest dynamic setting that can rationalize cooperative behavior while eliminating confounding factors such as multiple equilibria, strategic uncertainty, and other-regarding concerns: subjects play against a computer programmed to follow the Grim strategy, and this fact is known to subjects. Across all supergames, only 1-3% of subjects behave perfectly consistently with rational choice predictions, and only 3-8% behave consistently with the theory at least 95% of the time, in two standard student subject pools and in a more representative online subject pool. The majority make dominated choices resulting in money left on the table and a substantial minority engage in "end-timing", or gambling on the end to a supergame.

Keywords: Cognition, Cooperation, Rational Inattention, Prisoner's Dilemma, Repeated Games, Robot Players, Experimental Economics.

JEL Codes: C72, C73, C92.

*For helpful comments and suggestions, we thank Maria Bigoni, Gabriele Camera, Tim Cason, Subhasish Chowdhury, David Cooper, Guillaume Fréchette, Eugenio Proto, Yaroslav Rosokha, Yefim Roth, and Alistair Wilson, as well as numerous seminar and conference audiences. We are grateful to Patrick Julius, Nidhi Pandey and Yaoyao Xu for assistance programming the experiment. 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 # 2015-2082. The reported experiment is 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 and ISER Osaka University, duffy@uci.edu

‡King's College London, Ed.Hopkins@kcl.ac.uk

§King's College London, Tatiana.Kornienko@kcl.ac.uk

1 Introduction

Cooperation in repeated interactions is important for much socio-economic activity. However, despite an extensive experimental literature, it is still unclear what exactly determines cooperative behavior in prisoner's dilemma settings (see Dal Bó and Fréchette (2018) for a survey). 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 and the continuation probability is sufficiently high. However, rather than interpreting either behavior as mistaken, the literature instead has focused on many possible confounds, including diverse beliefs about opponents' strategies, heterogeneous risk attitudes, social preferences and cognitive limitations.

In an effort to simplify the analysis, this study introduces a novel experimental design that deliberately reduces or eliminates at least three confounding factors - multiple equilibria, strategic uncertainty, and social preferences. The experiment involves subjects playing a series of indefinitely repeated prisoner's dilemma (IRPD) games with different continuation probabilities δ against a robot opponent known to play the Grim trigger strategy.[1]. This enables a focus on the cognitive task of trading off present gains against future rewards, relying on basic dynamic programming arguments, effectively converting a strategic problem into an individual decision-making problem. Here the optimal policy is simple in theory: a subject should cooperate in each round if and only if the continuation probability δ is above a critical level, here 0.5. And since opponent behavior is perfectly predictable here, we can observe whether subjects implement the optimal policy over the *entire supergame*, rather than only in the first round as in much existing experimental analysis of repeated games.

In the baseline experiment, subjects faced a set of continuation probabilities δs skewed towards lower values (and thus shorter duration supergames), where it is theoretically optimal to always cooperate in two thirds of the supergames and always defect in the remaining one

[1]The Grim Trigger strategy prescribes cooperation until the opponent's first defection, after which the player defects permanently for the remainder of the supergame. While Tit-for-Tat (TFT) is a natural alternative, it has several disadvantages; see Section 3 for a discussion

third. This baseline was faced by subjects from two pools: a traditional group of university students (referred to as "Lab" subjects) and a more representative population of Amazon Mechanical Turk (AMT) workers (referred to as "Mturkers"). As a robustness check, a different group of university subjects faced a high-δ treatment involving a set of continuation probabilities δs skewed towards higher values (and thus longer duration supergames), comparable to prior repeated-game literature.[2] There, it is theoretically optimal to always cooperate in two thirds of the supergames and always defect otherwise.

Despite our simple setup, we find that all three samples of subjects behave in a manner that is strikingly different from the predictions of the rational choice framework used to explain cooperative behavior. Overall, across all supergames, only 2% (Lab), 1% (AMT), and 3% (high-δ) of subjects behave perfectly consistently with the rational choice predictions; only 5% (Lab), 3% (AMT), and 8% (high-δ) do that at least about 95% of the time. On average, subjects cooperate too much in the first round, with the share of decisions exceeding theoretically optimal by about 15% (Lab), 19% (AMT), and 21% (high-δ).

These very low frequencies of rational play across all treatments and subject pools indicate that the rational choice framework for explaining cooperative behavior could be less empirically relevant than commonly assumed. Admittedly, our design does not involve *human opponents*, but it is difficult to imagine that subjects would make fewer errors when confronted with the additional complexities of strategic uncertainty, the existence of multiple equilibria and uncertainty regarding others' preferences and beliefs, that feature in the standard setting.

Using subjects' choices in the first rounds of each supergame as a clear upper bound for the overall measure of rational play, we find that only 2% (Lab), 3% (AMT), and 7% (high-δ) of subjects behave perfectly consistently with the rational choice predictions at the start of each supergame. We further find that first round cooperation is strongly increasing in the continuation probability δ, from 10% (Lab), 26% (AMT), and 15.4% (high-δ) when

[2] In the high-δ design, continuation value δ was fixed for four consecutive supergames to promote equilibrium learning, and subjects faced the standard laboratory conditions.

$\delta = 0.1$ to 76% (Lab) and 74% (AMT) when $\delta = 0.7$ to 77.7% (high-δ) when $\delta = 0.85$. This responsiveness to the continuation probability δ (particularly in the Lab pool) is much greater than estimates based on Dal Bó and Fréchette (2018) in standard subject versus subject experiments, which suggests that our design is successful in reducing strategic uncertainty.

As noted, our methodology allows for a detailed examination of behavior in rounds beyond the initial one. Our analysis reveals significant deviations from the optimal strategy. First, 52% (Lab), 54% (AMT), and 40% (high-δ) of subjects cooperate at least once after already having defected in a supergame (and, thus, after triggering the defection by the robot opponent for the rest of the supergame), a strictly dominated behavior that is difficult to rationalize. Further, 24% (Lab), 30% (AMT), and 34% (high-δ) of subjects make this type of mistake repeatedly, in at least 3 out of 13 relevant supergames. While the frequency of such behavior diminishes with experience, it never completely disappears.

Second, subjects commonly defect after having begun the play of a supergame by cooperating. In contrast, the theory suggests that, given initial cooperative play, a player should continue cooperating for the duration of the supergame. Although the supergames have an unknown random end, subjects appear to be engaging in what we call "end-timing", which is gambling on the end of a supergame, defecting in the round they guess will be the last round of the supergame.[3] This "end-timing" strategy is similar to "sniping" in auctions (see Roth and Ockenfels (2002)), and we find that such "end-timing" behavior increases with experience. We are able to identify these behaviors only because of our novel, single-person experimental design, and our findings offer an alternative interpretation of results from other repeated game experiments involving matched pairs of human participants.[4][5]

[3] As is standard in indefinitely repeated games, we employ a constant termination probability of $1 - \delta$, where δ is fixed and known to subjects. Subjects may nonetheless hold the non-Bayesian belief that the probability of termination increases with the expected length of a supergame. Alternatively, Mengel et al. (2022) find that past *realized* supergame lengths matter for subjects' decisions in subsequent supergames.

[4] This "end-timing" behavior was an unexpected finding - see pre-registration of the project on the AEA RCT registry, https://doi.org/10.1257/rct.6318-1.0.

[5] Romero and Rosokha (2018) and Cooper and Kagel (2023) also report decreasing cooperation rates in indefinitely repeated prisoner's dilemma game experiments. However in those settings, the decrease in cooperation may be caused by beliefs that it is the cooperation by opponents which may be about to end.

Our results thus show that subjects do respond to the payoffs – but only partially, frequently making significant mistakes. And, behind these aggregate results, there is considerable heterogeneity - some subjects never cooperate while others always do. To explain this diversity of rationality and biases, we propose and test a novel model of behavioral inattention which predicts responsiveness to payoffs while allowing for inclinations and a substantial error rate, and where rational behaviour arises as a special case. We adapt the approach of Gabaix (2019) and assume that there is an unknown payoff associated with an unknown state of the world, which the agent seeks to match with her action. Yet she might be inattentive to the payoff-generating process, including the continuation probabilities δ, the opponent's strategy, design features, and so on. Our formulation of inattention is convenient as it directly implies a probit choice rule with two predictions. First, subjects will only noisily respond to expected payoffs, with the error rate decreasing in their cognitive ability. Second, individuals with less precise grasp of the decision problem will be more influenced by their own "default" prior payoff which is typically obtained in comparable situations outside the laboratory.

Consistently with the first prediction of this simple model, we find that subjects with higher test scores from a cognitive reflection test (CRT) tend to earn higher payoffs, make fewer errors of cooperating after triggering the opponent's irreparable defection, behave closer to theoretical predictions, and, indicatively, engage in end-timing. The second prediction may explain why, in all three samples, choices to cooperate in the initial round of a supergame (before any response by the computer opponent) are correlated with subjects' intrinsic level of patience (elicited and interpreted as a default prior for cooperation) - but only for those subjects with lower, not higher, cognitive test scores. In contrast, the prediction of the standard theory of repeated games is independent of individual characteristics (beyond risk and/or social preferences).

Experimental economists have used robot players for control purposes in a number of studies.[6] Two prior studies, by Roth and Murnighan (1978) and Murnighan and Roth (1983),

[6]For instance, Houser and Kurzban (2002) and Johnson et al. (2002) used robot players to remove the influence of social preferences in applications involving finitely repeated games. For surveys of experiments

are most closely related in having a population of subjects play the repeated PD game against a fixed strategy, as well as being the first studies to run experiments on supergames with an uncertain end. However, subjects in those studies faced some strategic uncertainty as they were *not informed* of the strategy they faced or that their opponent was in fact the experimenter.[7] By contrast, in this experiment we instruct subjects that they are playing against a programmed opponent who *plays the Grim trigger strategy*, and precisely what this means. 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. Also related is Andreoni and Miller (1993) who study play of *finitely repeated* PD games when there is a known probability that the opponent could be a robot tit-for-tat player (and not another human subject). They find that cooperation increases with the probability of facing such a robot player.

Recently, subjects' individual characteristics have been found to be correlated with cooperation in the indefinitely repeated prisoner's dilemma game (e.g., Davis et al. (2016), Kölle et al. (2020) and Gill and Rosokha (2022)). Proto et al. (2019) (see also Proto et al. (2022)) also find higher cognitive ability players being more cooperative, making fewer mistakes and earning higher payoffs. However, in all of these studies involving human vs. human subject pairings, there is no unique optimum policy as there is in our study. Consequently, any errors must 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, our experiment reveals that such behavior may represent an attempt to guess the final round of a supergame. Further, as noted, with our design we can also identify the dominated behavior of cooperating after one has previously defected.

---

combining human subjects and robot players see March (2021) and Bao et al. (2022).

[7]In Roth and Murnighan (1978), subjects "were told that they played a programmed opponent, but were not told what strategy he would be using" (p.194). The programmed opponent was in fact an experimenter playing the Tit for Tat (or "matching") strategy. In Murnighan and Roth (1983), 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]" (p.289). Roth and Murnighan (1978) explain that such design choices were made to "control for differences in subjects' behavior due to differences in their opponents" (p.194).

Recently and independently, the repeated play by human subjects against robot strategies was also explored experimentally, albeit with different designs and research objectives. In Reverberi et al. (2021), subjects played repeated games against robot opponents but the game and the strategy faced could change randomly over time. They then look at the interaction between the complexity of the strategy subjects face and their cognitive ability in determining the frequency of mistakes. Normann and Sternberg (2023) investigate whether algorithmic pricing facilitates collusion, in a design involving three and four firm oligopolies where one firm's prices are determined by an algorithm, and point out that interactions between humans and algorithms or robots may have increasing practical importance. Finally, Blonski et al. (2025) has human subjects play against two types of computer algorithm. However, in all of these studies subjects were not informed exactly what the algorithmic strategies were and so were subject to strategic uncertainty.

Surprisingly, despite the fundamental importance of dynamic optimization in contemporary economic theory, only a few experiments, Noussair and Olson (1997) and Carbone and Duffy (2014), explored subjects' capacity for dynamic optimization and also found deviations from optimal behavior. In our study, the problem at hand is even simpler as it does not involve some changing, continuous state variable such as capital or wealth, instead focusing on an arguably more intuitive trade-off between the immediate and future gains from cooperation.

Our methodology shares similarities with the approach of Charness and Levin (2009), who experimentally demonstrate the winner's curse in a single person bidding task. In both studies, simplifying the environment allows identification of cognitive failures that are harder to detect in more complex strategic settings. In our case, the failure is an inability to maintain a consistent strategy in a stationary environment, leading to suboptimal behavior absent strategic uncertainty. A key distinction (aside from the different game) is our within-subject design, where subjects face both cooperative and non-cooperative optimal conditions. Thus, our study of repeated interactions builds on the methodology of Duffy et al. (2021) and Charness et al. (2021), who also used experimental designs with contrasting environments.

2 Theory and Hypotheses

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

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

where X (Y) denote the cooperate (defect) actions.

2.1 Insights from the Theory of Repeated Games

The main theoretical prediction tested in our experiment derives from the Folk Theorem for repeated games, which states that if players are sufficiently patient, then any pure-action profile yielding payoffs that strictly dominate the pure-action minimax is a subgame-perfect equilibrium of the repeated game in which this action profile is played every period (Mailath and Samuelson, 2006, p.69). In laboratory indefinitely repeated games, "patience" is proxied by a high continuation probability, δ. In our setup, one player (the computer) follows a Grim Trigger strategy, reducing the environment to a single-person decision problem with a unique optimal policy. Assuming risk neutrality, the optimal strategy is to cooperate (defect) in all rounds if δ is above (below) the critical threshold $\delta^* = 0.5$ for our parameterization (1).

To see this, note that since the computer plays a fixed Grim Trigger strategy, it begins each supergame by cooperating, and continues to do so as long as the human also cooperates. It switches permanently to defection after any human defection. Given the fixed continuation probability δ, the expected payoff from always cooperating (X) is

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

while the expected payoff from defecting once in the first period (Y) is

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

Simple algebra shows that cooperation yields a higher payoff than defecting when $\delta > 0.5$, so the critical continuation probability is $\delta^* = 0.5$.

This assumes risk neutrality. Risk aversion could change these calculations. For example, for CRRA preferences, one can verify numerically that the critical value δ^* is increasing in the degree of risk aversion, so potentially defection becomes optimal for values of δ for which cooperation is optimal under risk neutrality. However, to change any predictions for our current parameters, the requisite CRRA risk-aversion parameter is implausibly high. For example, to make defection optimal when $\delta = 0.67$ (the lowest value of δ we use for which cooperation is optimal under risk neutrality), a CRRA parameter of more than 1.1 is needed.[8] Thus, we do not expect risk attitudes to affect play in our games.

Since the continuation probability δ is constant over time, the decision problem is *stationary*: if cooperation is optimal in period 1, it is optimal in every subsequent period. Hence, it cannot be optimal to switch within a supergame from cooperate to defect.[9] Moreover, given the computer's fixed Grim Trigger strategy, once a subject defects and triggers permanent punishment, it is optimal to keep defecting rather than switch back to cooperation. This yields the following simple hypothesis.

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

Our design removes three confounding factors from the standard two-player repeated Prisoner's Dilemma. First, it eliminates *multiple equilibria*. When δ is sufficiently high to

[8] Andersen et al. (2008) estimate the average CRRA risk-aversion parameter using controlled experiments with field subjects in Denmark as 0.741, and with estimated standard deviation of only 0.056. Hence, 1.1 is more than six standard deviations above the mean.

[9] In contrast, if one were playing a finitely repeated dilemma with known length against a known Grim opponent, the optimal strategy is to cooperate until the penultimate round and defect in the last. This very different prediction is one of the reasons why we concentrate on uncertain length supergames here.

support cooperation, the standard game admits infinitely many equilibria, creating difficult coordination problems for subjects. With a Grim Trigger computer opponent, the equilibrium set collapses to a singleton: always cooperate or always defect, depending on δ.

Second, our design minimizes *strategic uncertainty*, which arises in standard experiments because subjects do not know their opponent's strategy. Indeed, a simplification used by Dal Bó and Fréchette (2018), following Blonski et al. (2011), is to suppose strategies are limited to Grim Trigger and Always Defect, and show that there exists a $\delta^{RD} > \delta^*$ such that Grim Trigger (and thus initial cooperation) is risk-dominant (RD) if and only if $\delta > \delta^{RD}$. Thus, while cooperation is an equilibrium for $\delta > \delta^*$, strategic uncertainty can impede cooperation unless $\delta > \delta^{RD}$, a higher hurdle.

Third, our design should remove other-regarding *social preferences* as a driver of behavior. Social preferences often matter in repeated-game experiments (see, e.g., Camerer (2003); Chaudhuri (2008)). In the repeated Prisoner's Dilemma, Bernheim and Stark (1988) and Duffy and Muñoz-García (2012) show that altruism lowers the critical continuation probability δ^*, allowing cooperation even when $\delta < \delta^*$. Moreover, if subjects believe others have social preferences, even self-interested players may cooperate more (Andreoni and Samuelson, 2006), creating an interaction with strategic uncertainty. In our design, however, the opponent is a computer known to play Grim Trigger, which both delivers a unique optimal response (eliminating strategic uncertainty) and makes altruistic motives or beliefs about the opponent's altruism unlikely. Thus, multiple equilibria, strategic uncertainty, and social preferences—and their interactions—are minimized, if not eliminated, by our design.

2.2 A Simple Cognitive Model

The above theory predicts an optimal policy for the repeated game that is *independent* of individual characteristics. In this section, we present a simple model that can be used to explain heterogeneity in deviations from optimality in single-person decision problems. Importantly, this model generates the novel and testable prediction that subjects with lower cognitive ability both make more mistakes and are also more influenced by irrelevant non-cognitive factors.

We assume that subject i faces a single-person decision problem. Based on prior experience, she expects an initial default payoff d_i, interpreted as a typical payoff to cooperation outside the laboratory (a subject's non-cognitive characteristic). In the lab, faced with a supergame with continuation probability δ, she seeks to determine the optimal action by introspection, and updates the payoff to cooperation to fit the circumstances. Our innovation is to model this by assuming she receives a signal the informativeness of which depends on her cognitive ability, which varies across individuals. As in Enke and Graeber (2023), we further assume that she recognizes her cognitive limitations and places greater weight on the default payoff the higher is her cognitive uncertainty.

Let the true relative return to cooperation (the payoff to cooperation minus the payoff to defection) be $\pi(\delta)$, where $\pi(\cdot)$ is strictly increasing and, given our parameters, $\pi(\frac{1}{2}) = 0$. Adapting Gabaix (2019)'s simple Gaussian framework,[10] we assume subject i subjectively estimates $\pi(\delta)$ as normally distributed with mean d_i and variance σ^2, i.e., $\pi(\delta) \sim N(d_i, \sigma^2)$. Thus, her initial evaluation is influenced by her outside default d_i, which varies across subjects (while, for simplicity, σ^2 is common).

With further cognitive introspection, subject i can obtain a noisy payoff signal s_i equal to the true value $\pi(\delta)$ plus noise $\varepsilon_i \sim N(0, \sigma_i^2)$, or $s_i(\delta) = \pi(\delta) + \varepsilon_i$. The noisiness of the signal varies across individuals through σ_i^2; higher cognitive ability implies lower σ_i^2 and thus a more precise signal. Following Gabaix (2019) and standard Bayesian updating, the posterior estimate of $\pi(\delta)$ is a weighted average of the signal and the prior, or $P_i(\pi(\delta) \mid s_i) = \lambda_i s_i(\delta) + (1 - \lambda_i) d_i$, with weight determined by relative variances, $\lambda_i = \frac{\sigma^2}{\sigma^2 + \sigma_i^2}$. Note that as cognitive noise σ_i^2 goes to zero, the weight λ_i goes to one and the posterior estimate P_i is closely clustered around the true payoff π. However, for a subject with high σ_i^2, the posterior estimate is instead very close to the individual's default payoff d_i.

[10]Our model leads to the choice of cooperation as being characterized by a probit choice rule, which is similar to the rational inattention model of Matějka and McKay (2015) which results in a logit choice rule.

Recall that, given our definition of π as the relative payoff to cooperation, the individual estimates that cooperation is preferable to defection if the posterior estimate $P_i > 0$.[11] Thus, when facing a decision problem with continuation probability δ, the subject's probability of cooperation C_i is the probability that P_i is positive:

$$C_i(\delta) = \Pr\left(s_i + \frac{1-\lambda_i}{\lambda_i} d_i > 0\right) = \Phi\left(\pi(\delta) + \frac{\sigma_i^2}{\sigma^2} d_i\right) \tag{4}$$

where Φ is the normal CDF of ε_i with variance σ_i^2. Thus, the individual's actions follow a probit in which the probability of cooperation depends on both the true payoff and the individual-specific default payoff.

Further, individuals with higher cognitive ability, and thus lower cognitive noise σ_i^2, place less weight on the default payoff d_i and more weight on the true payoff π. This has two implications. First, because $\pi(\delta)$ is increasing in the continuation probability δ, those with higher cognitive ability/lower noise should be more sensitive to δ in their cooperation choices.[12] Second, the probability of cooperation C will be closer to its optimal value, resulting in higher ability (lower noise) individuals earning higher payoffs.

Hypothesis 2. *For high noise (low cognitive ability) subjects, the probability of cooperation is more influenced by their default payoff value to cooperation (their non-cognitive characteristics) and less influenced by the true payoff to cooperation (and, thus, the current continuation probability δ), resulting in lower average payoffs, than for low noise (high ability) subjects.*

3 Experimental Design

The main experimental task consisted of 24 indefinitely repeated prisoner's dilemma games, or "supergames," played against a computer program known to use the Grim trigger strategy.

[11] We depart slightly from Gabaix (2019) at this point, because he considers a continuous action space rather than the discrete choice of cooperation versus defection here.

[12] Specifically, $\partial C/\partial\delta$ is proportional to $\Phi'(\cdot)$, which is decreasing in σ_i around the critical point $\pi(\delta) = 0$, by properties of the normal distribution.

We represented the stage game in normal form, with both players' payoffs common knowledge, to ensure comparability with human-versus-human studies of the prisoner's dilemma that use that same format. The payoff matrix for the stage game was identical across all treatment conditions and is shown in (1). Subjects were told that the rows referred to their actions, the columns to the computer's actions, and that the first number in each cell (in bold) was their payoff in points, while the second (in italics) was the computer's payoff.

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 δ 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 six continuation probabilities, $\delta \in \{0.1, 0.25, 0.33, 0.4, 0.67, 0.7\}$, was selected according to several criteria. First, under specification (1), the expected total theoretical payoff averaged across these values of δ is identical for subjects who are biased toward always cooperating or always defecting. Second, the expected payoff from always following the theoretically optimal strategy, relative to either of these fully biased strategies, is substantial, ensuring a clear payoff distinction. Finally, because the threshold continuation probability for sustaining cooperation in the stage game (1) is $\delta^* = 0.5$, we sought to avoid a setting in which the simple heuristic of cooperating in 50% of the supergames coincides with the optimal policy. Under our design, optimal play instead entails cooperating in only 8 of the 24 supergames (those with $\delta = 0.67$ or 0.7) and defecting in the remaining 16. Thus, unlike most existing studies, our distribution of continuation probabilities is intentionally weighted toward shorter expected supergame durations. As discussed in Section 6, we later replicate our experiment using higher values of δ to assess robustness.

We ran our experiment with Grim Trigger as the only programmed strategy. While Tit-for-Tat (TFT) may seem a reasonable alternative to Grim, it provides a much weaker restriction on optimal strategies: cooperation after defection is not necessarily an error against TFT while it is against Grim. Thus, the identification of optimal play is significantly more difficult if the robot player plays TFT rather than Grim, and for this reason we elected to consider only the Grim strategy.

The experiment was computerized using oTree (Chen et al. (2016)) software and conducted online. Subjects received written instructions for the 24 IRPD games (referred to neutrally as "sequences") they would play, then completed a comprehension quiz testing their understanding of payoff outcomes, the Grim trigger strategy used by the computer program, and how the continuation probability affected game duration (see Appendix A.1 for the instructions and quiz). Although the quiz was not incentivized, subjects could not proceed to the main task until all questions were answered correctly. Example screenshots appear in Appendix A.3.

After subjects were presented the list of all continuation probabilities δ, but before play of the fist supergame, each subject was asked to provide their belief as to the proportion of times they would choose the cooperative action (referred to neutrally as action "X") in each of the *first* rounds of the 24 sequences (supergames) that they would face, given knowledge that they would face 4 supergames for each of the 6 different δ values (details in Appendix A.2). After this "Prediction" belief was elicited, they played the 24 supergames against the computer opponent. We collected this prediction belief to measure how much/little attention subjects paid to the structure of the game as per the experimental instructions, and indeed, we use this prediction data later on in the evaluation of our model of inattentive behavior.[13]

Half the subjects faced a sequence of randomly drawn continuation probabilities δ, four supergames for each of six δ values, total of 24 supergames.[14] For the other half, the order of supergames was reversed.[15] (See Tables A1 and A2 in Appendix A.6 for the sequence of randomly drawn continuation probabilities δ, realized durations and their expected durations.)

Subjects were told that at the end of the session, six supergames — one for each δ —

[13]While one could argue that such elicitation may anchor subsequent behavior, as we will show later, consistently with the inattention model, this "Prediction" variable is correlated with the behavior of only a certain subset of subjects.

[14]Supergame lengths were drawn using a random number generator, and subjects were informed of this procedure. To reduce noise, the same supergame lengths were used for all participants.

[15]See Appendix B.3 for a discussion of order effects.

would be randomly selected for payment. During play of the 24 supergames, subjects were always informed of the probability that each 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 the *same* decision screen where they made their own action choice (X or Y) for that same round (see the screenshots in Appendix A.3 Figures A1-A2). Thus, any strategic uncertainty regarding the play of the computer player should have been eliminated.

3.1 Subject Pools and Experimental Earnings

We recruited two gender-balanced subject samples, university students (Lab) and Amazon Mechanical Turk workers (AMT), using the same experimental program for both groups.

The first pool (henceforth, Lab subjects) consisted of 100 undergraduates (52% female) recruited via Sona Systems from the Experimental Social Science pool at the University of California, Irvine. The mean age was 21.5 years (range 18–34). Subjects represented diverse majors: 36 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). Sessions with these lab subjects were conducted online via Zoom in the aftermath of the Covid pandemic (late 2020) where the experimenter verified each participant's identity, monitored activity, and remained available to answer questions throughout the session, and thus was continuously present to ensure attention and supervision comparable to conventional lab settings. (In Section 6, we report similar results from standard, *in-person* lab sessions conducted in 2025.)

We also recruited 149 AMT subjects residing in the United States (henceforth, AMT subjects). There were equal numbers of males and females, with one subject preferring not to disclose gender (coded as 0.5). The AMT subjects were considerably older than the Lab sample, ranging from 21–75 years with a mean of 39.8 years. (Appendix Table B2 formally compares the two samples.) Their educational attainment was more varied, including two participants without a high school diploma.

In both subject pools, subjects were informed that their total point earnings from six randomly chosen supergames (one for each δ value) would be multiplied by USD $0.01 to determine their monetary earnings from the repeated PD games. Thus, both groups faced identical variable earnings, allowing direct comparison. Fixed show-up payments differed—$7 for Lab and $1 for AMT subjects—reflecting standard norms for laboratory versus online recruitment (see Rand (2012)). Lab subjects' total earnings averaged $17.90 for a one-hour extended session.[16] AMT subjects earned an average of $10.37 for a one-hour session.

Exactly half of the Lab subjects (50/100, or 4 of 8 sessions) and nearly half of the AMT subjects (74/149) faced the "long" order (first supergame $\delta = 0.67$), while the remainder faced the reverse order (starting with $\delta = 0.33$).

4 Results

Both subject populations completed 24 supergames, each lasting at least one round. Owing to random termination, 11 supergames ended after one round, while the remaining 13 lasted 2–5 rounds. Each subject therefore made 24 first-round and 24 subsequent-round choices, for a total of 48 (see Appendix Table A2). Given our parameters, the theoretically optimal strategy is to cooperate in all rounds of the 8 supergames with $\delta \in \{0.67, 0.7\}$ (26 cooperative choices) and to defect in all rounds of the remaining 16 supergames (22 defections).

4.1 The Headline Result

Figure 1 summarizes the full dataset, separately for the Lab and AMT samples. For each of the six δ values, it reports the shares of cooperate and defect actions pooled over all

[16] After completing the 24 repeated PD games, Lab subjects (only) were randomly paired for another two-player task, not reported here, in which they could earn an additional $1.00–1.70.

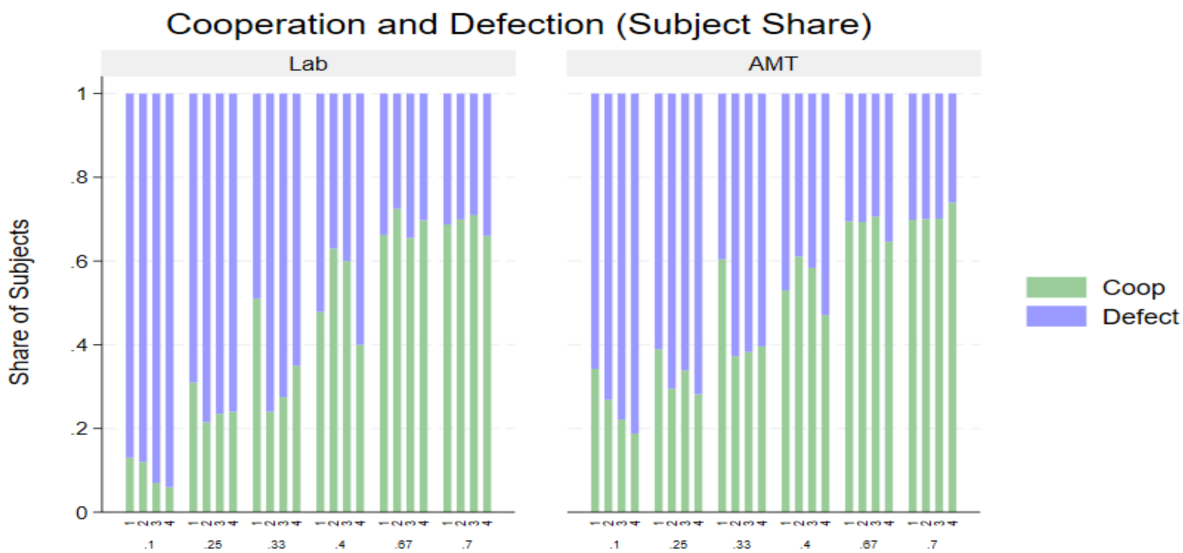

Figure 1: Cooperation and Defection rates across all rounds of the four supergames in chronological order (first horizontal axis) of each of the six δ values (second horizontal axis). Left panel: 100 Lab subjects. Right panel: 149 AMT subjects.

rounds within each of the four supergames (labeled 1–4 on the first horizontal axis), with the corresponding δ shown on the second horizontal axis. Three broad features stand out. First, cooperation is not driven to zero in the four supergames with $\delta < \delta^* = 0.5$, and cooperation is less than full for the two δ values above δ^*. Second, beginning with the second supergame, behavior changes little across the remaining supergames for a given δ, suggesting limited learning over time. Third, AMT participants are farther from the optimal policy than Lab participants, especially in the supergames with $\delta < \delta^*$. If we focus only on first round behavior, we find similar patterns of behavior (Appendix B, Table B1).

These aggregate patterns reflect substantial deviation from standard game-theoretic predictions (Hypothesis 1). As shown in Figure 2, only 2 of 100 Lab participants (2%) and 1 of 149 AMT participants (0.7%) behaved perfectly optimally in all rounds of all supergames. In both samples, fewer than 5% made three or fewer errors out of 48 choices (i.e., at least 45 optimal choices, or 93.75%). About 10% in each sample (10% Lab; 10.7% AMT) made at least 42 optimal choices (87.5%). Appendix B.4 reports the full distribution of individual counts of cooperative (and hence optimal) choices.

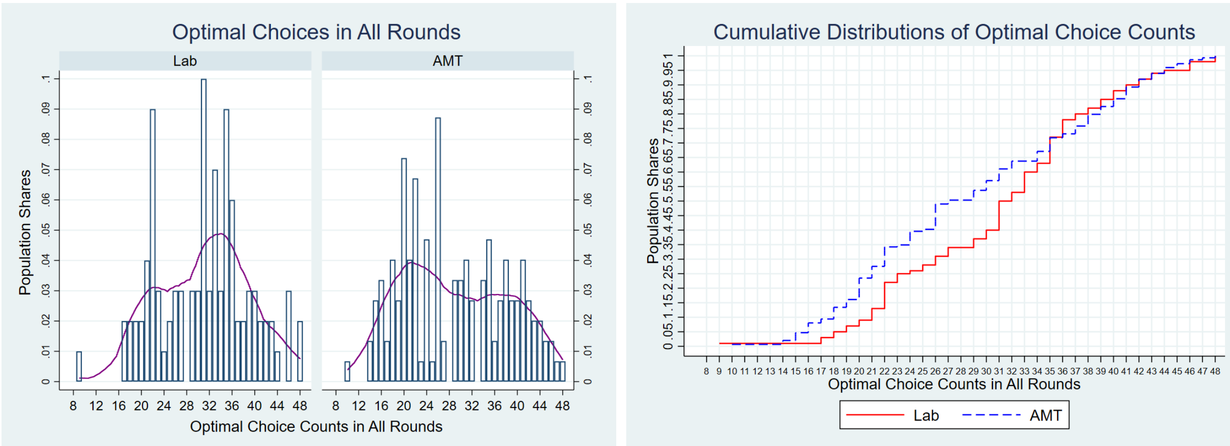

Figure 2: Left: Frequency distributions and kernel densities of per-subject counts of optimal choices across all supergames (48 decisions per subject): Lab (N=100) vs. AMT (N=149) subjects. Right: Cumulative distributions of optimal choice counts for both the Lab and AMT samples in the same graph.

Finding 1. *Across all 48 decisions in all 24 supergames, the fraction of subjects who behaved according to standard game-theoretic predictions (either perfectly or near-perfectly) does not exceed 5%. About 90% in both samples made fewer than 42 theoretically optimal choices out of 48 (87.5 %).*

As one would expect, Lab subjects behaved significantly closer to the theoretical optimum than AMT subjects (Kolmogorov–Smirnov one-sided test $D = 0.1799$, $p = 0.021$), though this difference narrows among those with higher rates of optimal play (see Figure 2).

4.2 First Round Behavior

Following the prior studies, we look at first-round choices made before any feedback to capture subjects' initial strategic tendencies. Given our design, the optimal strategy is to cooperate in all rounds of the eight supergames with $\delta = \{0.67, 0.7\}$ and defect in all others. Thus, perfectly optimal play implies 8 cooperative and 16 defecting first-round choices. By design, the theoretically optimal pattern is skewed toward defection in first rounds, as most δs are below 0.5, and toward cooperation in later rounds, when higher δ values make such cooperation optimal.

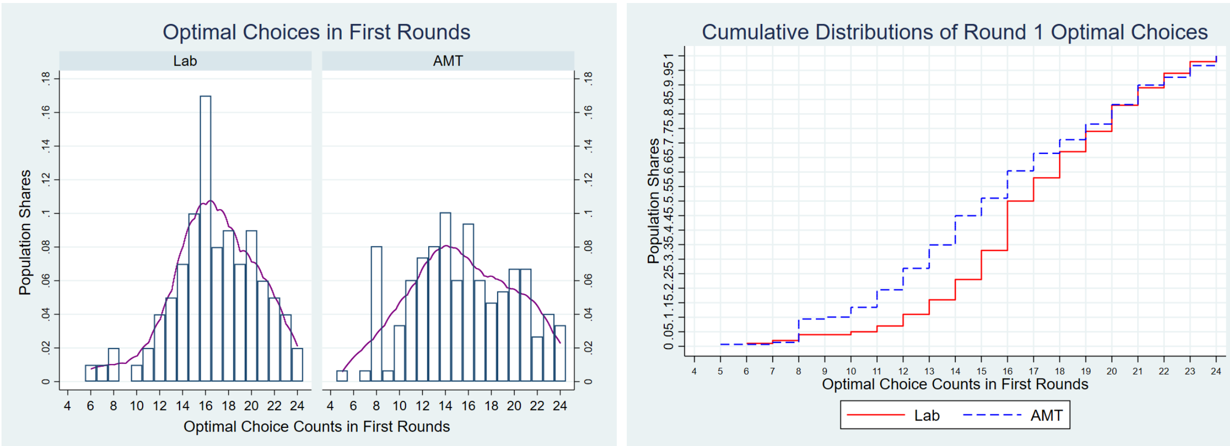

Figure 3: Left: Frequency distributions and kernel densities of per-subject counts of optimal choices in first round choices across all 24 supergames (24 decisions per subject): Lab (N=100) vs. AMT (N=149) subjects. Right: Both cumulative distributions on the same graph.

Figure 3 displays the distribution of optimal first-round choices across 24 supergames for both Lab and AMT samples. As shown, only a small fraction of subjects behave exactly in line with equilibrium predictions in these first rounds: specifically, 2 of 100 subjects (2%) in the Lab sample and 5 of 149 subjects (3.36%) in the AMT sample. Moreover, fewer than 17% of subjects in either sample made three or fewer mistakes (i.e., at least 21 optimal first-round choices out of 24, or 87.5%). While Lab subjects performed significantly closer to the theoretical benchmark than AMT subjects (Kolmogorov–Smirnov one-sided test $D = 0.2197$, $p = 0.003$), this advantage narrows among those achieving the highest rates of optimal play.

The left panel of Figure 4 displays the overall frequency of first round optimal play by continuation probability δ. First round optimal play is highest when δ is relatively low ($\delta \in \{0.1, 0.25\}$) and most subjects optimally defect. It is also high when δ is relatively high ($\delta \in \{0.67, 0.7\}$) and most subjects optimally cooperate. First-round optimal play is lowest at $\delta \in \{0.33, 0.4\}$, where subjects cooperated excessively despite these values lying below the (unknown) threshold $\delta^* = 0.5$. This tendency is especially strong at $\delta = 0.4$, where over half the subjects in both samples cooperated suboptimally.

This pattern is broadly consistent with a stochastic choice model such as the inattention model of Section 2.2 or a logit, where the frequency of optimal behavior is increasing in

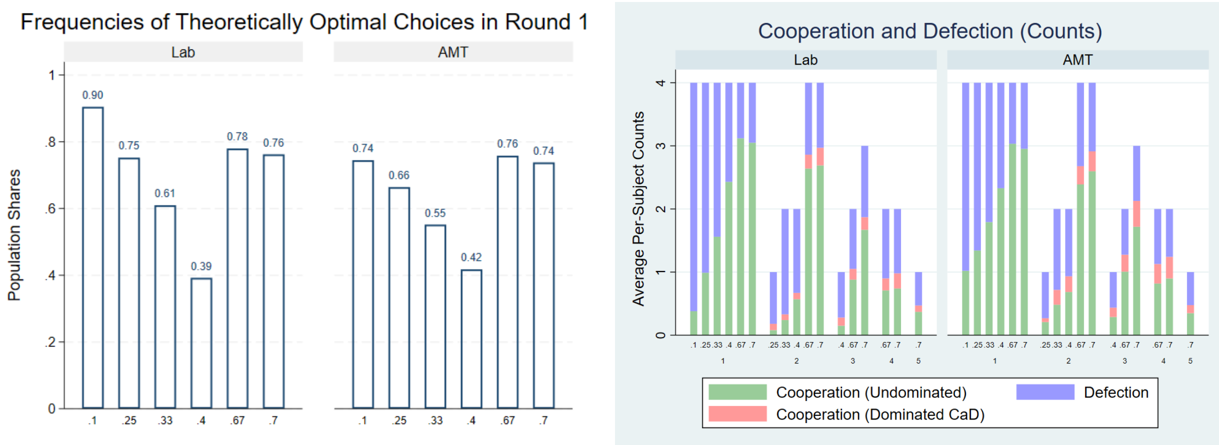

Figure 4: Patterns of optimality, cooperation and defection: Left panel: population shares of theoretically optimal choices in the first rounds of all supergames by the continuation probability δ (see also the leftmost set of bars in the right panel). Right panel: Average per-subject counts of cooperation versus defection, split by continuation probability δ (the first row of the horizontal axis scale) and by the round number of a supergame (the second row of the horizontal axis scale). Starting from round 2, a distinction is made between undominated cooperation and dominated *cooperation after defection* (CaD). The later rounds were never reached for some δ values (see Appendix Table A2). Lab: 100 subjects, 2,400 supergames, AMT: 149 subjects, 3,576 supergames.

payoff differences. However, as shown in Appendix B.2 while the cost of first round mistakes is indeed lowest (0) for a δ value equal to δ^*, and these costs rise as δ departs from δ^*, the increase in these costs is asymmetric, with relatively lower costs for δ values less than δ^* and relatively higher costs for δ values greater than δ^*. When cooperation is the optimal round 1 action, a single initial defection triggers permanent punishment, so the subject can never return to the cooperative path. By contrast, when defection is optimal, an initial incorrect cooperation choice need only delay convergence to the optimal outcome.

At the same time, the evidence that these incentive differences can account for the observed behavior is mixed. Figure 4 (and later on Figure 13) show that first round departures from the theoretical optimum are smallest for the *lowest* δ value of 0.10 and not for the highest δ value of 0.70, so the costs of mistakes alone cannot fully explain these deviations from theoretical predictions.

This excessive cooperation cannot be attributed to risk aversion either, since — as noted in Section 2.1 — risk aversion should instead lead to excessive defection. Rather, these apparent mistakes are better explained by the behavioral inattention theory of Section 2.2. In such binary stochastic choice models (or simple Probit or Logit models), the rate of optimal choice is lowest when the expected-payoff difference is smallest, near the critical value $\delta^* = 0.5$. This is *precisely* what we observe in the left panel of Figure 4.[17]

Finding 2. *In the first rounds of supergames, subjects in both samples cooperated in excess of the theoretical optimum—on average by 44.1% in the Lab and 55.9% in the AMT sample. Excessive cooperation is most pronounced at intermediate continuation probabilities,* $\delta \in 0.33, 0.4$*. Lab subjects were significantly more sensitive to changes in* δ *than AMT subjects.*

4.3 Behavior Over All Rounds: Cooperation After Defection (CaD)

Our design allows us to examine behavior beyond the first rounds of the supergames. Given our parameterization and random draws, 11 supergames ended after the first round and 13 continued (see Appendix Tables A1 and A2). The right panel of Figure 4 shows average per-subject counts of cooperation versus defection by continuation probability δ for the first round and for rounds 2–5 (no data exist beyond round 5). As shown, cooperation rises sharply with increases in δ, a finding confirmed later on in the mixed-effects probit results of Table 3, specifications (1)–(2). First-round cooperation rates range from 9.5% (Lab) and 25.5% (AMT) when $\delta = 0.1$, to 76.25% (Lab) and 73.83% (AMT) when $\delta = 0.7$. For the Lab subjects, this responsiveness far exceeds that observed in standard subject-versus-subject experiments.[18] In contrast, AMT subjects are less responsive, particularly at $\delta = 0.1$.

Since the robot opponent followed the Grim trigger strategy, any defection in a supergame induced defection by the opponent in all remaining rounds. Subjects were quizzed on their understanding of this strategy prior to play of the games. Further, subjects were explicitly informed of the strategy the robot would play in each round before they made their own

[17] Because payoffs are nonlinear, stochastic choice models predict an asymmetric mistake rate around $\delta^* = 0.5$, with optimal behavior more frequent at $\delta = 0.67$ than at $\delta = 0.33$.

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

decision. Thus, choosing to cooperate after defecting earlier in the *same* supergame (CaD) is *dominated* for any δ and constitutes a *strategic error*. In the right panel of Figure 4, such suboptimal cooperation (CaD) is distinguished from non-erroneous cooperation in the cooperation counts. Among the 13 supergames lasting more than one round, the share of CaD errors is 7.57% for Lab subjects and 11.47% for AMT subjects.

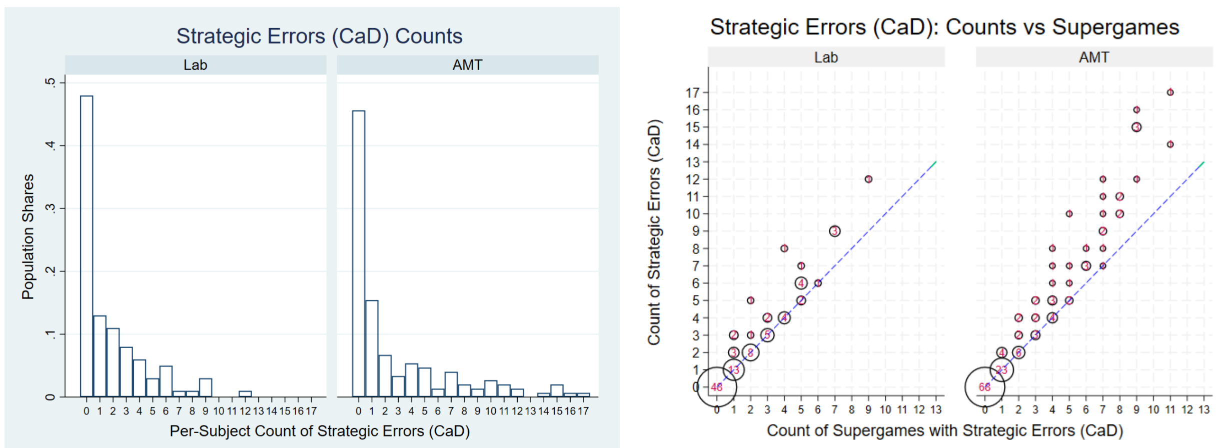

Figure 5: Strategic errors of dominated cooperation after defection (CaD) in 13 relevant supergames (i.e., those lasting longer than one round). Left: Distribution of per-subject counts of instances of cooperation after defection (CaD). Right: Per-subject counts of CaD instances vs. count of supergames with those instances. Bubble size is proportional to the share of subjects. (100 Lab subjects, 149 AMT subjects).

As the left panel of Figure 5 shows, slightly less than half of subjects (48% for Lab and 45.64% for AMT) never made the strategic error of CaD, and 20% of the Lab subjects and 28.86% of the AMT subjects made at least 4 dominated CaD choices. Some such choices could be intentional, e.g., due to a desire to verify the correctness of the displayed information about the computer opponent's behavior. Others could be due to a genuine "trembling hand" error of accidentally pressing the "defect" button without noticing it. In either case, an attentive payoff-maximizing subject would likely refrain from repeatedly making such dominated CaD choices in multiple supergames.[19]

Given the possibility of trembling-hand behavior, we consider the number of supergames in which the strategic error of CaD occurs at least once. The right panel of Figure 5 compares the total count of CaD errors per subject (vertical axis) with the number of supergames

[19]Kloosterman (2020) also finds that subjects in repeated, human vs. human prisoner's dilemma games, can surprisingly return to cooperation after defection.

containing such errors (horizontal axis). Most strategic errors occur only once per supergame (as shown by bubbles on the diagonal), yet their overall extent is substantial: 24% of Lab and 29.53% of AMT subjects made CaD errors in at least 3 of the 13 relevant supergames (those lasting more than one round). This suggests that some dominated CaD behavior may reflect inattention or limited strategic understanding (though recall that subjects had to pass a quiz before play). While the prevalence of CaD errors is relatively small, it nevertheless complicates interpretation of deviations from theoretically optimal behavior.

Finding 3. *Over all rounds of all supergames:*

(a) Cooperation (defection) increases (decreases) with the continuation probability δ.

(b) A majority of subjects (52% in Lab, 54.36% in AMT) made at least one strategic error by cooperating after previously defecting (CaD) within the same supergame, triggering a "grim" response. Such dominated cooperation accounts for 7.58% of relevant observations in the Lab and 11.47% in AMT, with 24% of Lab and 29.53% of AMT subjects making this error in at least 3 of the 13 supergames lasting more than one round.

4.4 Overall Point Totals

In this section we focus on total awarded points, which is the sum of points earned over all 48 decisions. This serves as a theoretical measure of the payoff consequences resulting from subjects' behavior. First, recall that the theoretically optimal policy of cooperating (defecting) in every round of every supergame when $\delta > (<) 0.5$ (see Hypothesis 1) is derived *ex ante*, before the lengths of each supergame are realised. One can calculate that, given the realization of random supergame terminations, the *ex ante* optimal play would result in an overall total of 4,050 points *ex post.*[20]

Figure 6 reports on subjects' point totals. As this figure shows, empirically overall point

[20]The theoretically optimal point total of 4,050 points is calculated as follows: A player earns 75 points in each round of supergames with $\delta \in \{0.67, 0.7\}$ (26 decisions) plus 120 points in the first rounds (16 decisions) and 30 points in the subsequent rounds (6 decisions) of the supergames with $\delta \in \{0.1, 0.25, 0.33, 0.4\}$.

totals range from $3,285$ to $4,185$ points for the Lab sample and from $3,195$ to $4,185$ for the AMT sample with a mean (st. dev.) of 3,835.05 (203.38) for Lab and 3,766.21 (240.58) for AMT. Note that the lowest *ex post* point total that is achievable here is quite substantial at 2,745 points.[21] This *ex post* lower bound can be seen as the "fixed" component of the overall point total, and any overall point total in excess of that amount can be seen as the "variable" component. It is easy to see that, on average, the Lab subjects earned only 83.5% and the AMT subjects only 78.3% of the "variable" component of the overall point totals that could be achieved by following the theoretically optimal policy.

As Figure 6 further reveals, some Lab subjects earned as little as 41.4% and some AMT subjects as little as 33.7% of the "variable" component achievable by following the optimal policy. Furthermore, 16% of Lab subjects and 28.86% of AMT subjects could have *increased* their total point earnings to 3,600 points by simply choosing either always to cooperate (All-C) or to always defect (All-D),[22] which is 65.5% of the "variable" component achievable by following the theoretically optimal policy. This observation suggests that, among other deviations, strategic errors (CaD) reduce overall point totals.

As Figure 6 further shows, the mode for the Lab sample is at the theoretically optimal point total of 4,050, and this point total is also relatively frequent among the AMT subjects. Yet, strikingly, the maximum point total is still higher in both samples. Overall, 19% of the Lab subjects and 16.11% of the AMT subjects were able to achieve at least the theoretically optimal point value of 4,050, despite only two subjects in the Lab sample and one subject in the AMT sample actually behaving in a way that was fully theoretically optimal (in all 48 choices). Thus, given the random realization of the supergames, some subjects were able to achieve at least as much as the theoretically optimal point total despite pursuing strategies

[21]This *ex post* theoretical minimum of 2,745 points arises from making the strategic mistake of cooperating after defecting and is calculated as follows. In the 11 supergames lasting only one round, one achieves the lowest payoff of 75 by cooperating (11 decisions). But in the supergames lasting longer than one round, the lowest payoff is obtained by defecting in the first round and earning 120 (13 decisions) and earning only 15 from cooperating thereafter (24 decisions).

[22]Recall that the expected payoffs to these two extremely biased strategies are the same by design. *Ex post*, by cooperating always, in every round of every supergame, one would earn 75 in each of 48 decisions; and by defecting always one would earn 120 in all first rounds (24 decisions) and 30 thereafter (24 decisions).

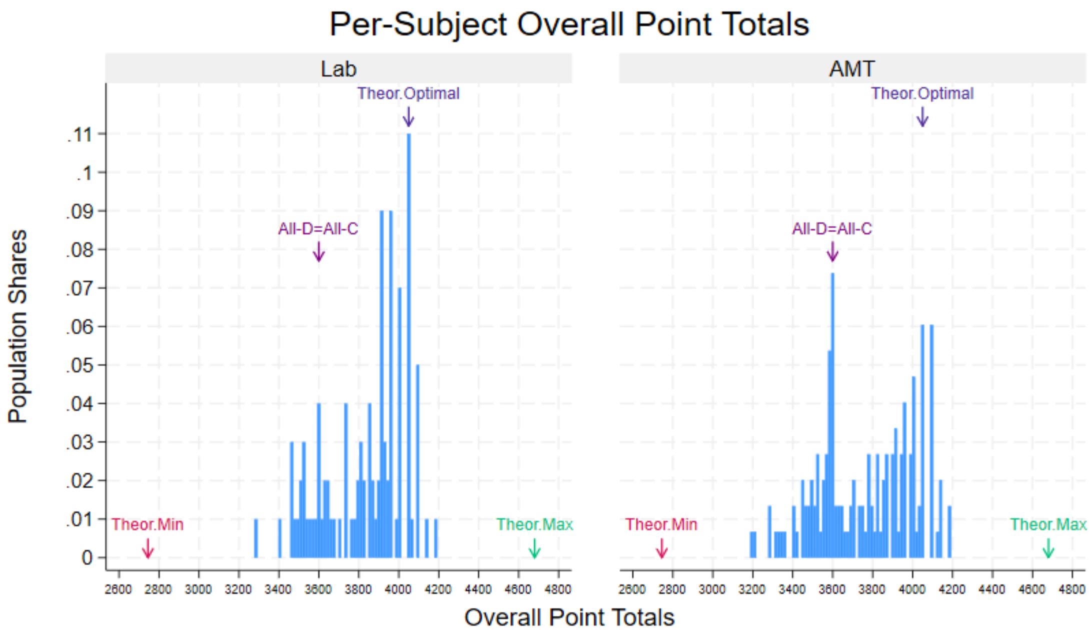

Figure 6: Distribution of overall point totals, or the sum of point earnings across all 48 decisions. *Ex post* theoretical point total from following the *ex ante* optimal policy is 4,050, while the ex post theoretical minimum and (omniscient) maximum point totals are 2,745 and 4,680, respectively.

that were not theoretically optimal. A potential explanation for this mystery is provided in the next Section 4.5.

Finding 4. *Lab subjects earned 83.5% and AMT subjects 78.3% of what could be achieved relative to the* ex post *theoretical minimum by following the* ex ante *optimal policy. Overall, 16% of Lab and 28.86% of AMT subjects earned less than they could have by always cooperating or always defecting. Notably, 17% of Lab and 15.43% of AMT subjects achieved totals at least as high as the theoretical optimum—without always following the optimal strategy.*

4.5 "End-Timing" (DaC)

Note that when playing against a robot that uses the Grim trigger strategy, one could earn the highest feasible payoff if one knew *exactly when* each supergame would end, by cooperating at first but defecting in the final round. Since subjects in our study did not know when a supergame would end, they could not execute this "end-timing" strategy perfectly.[23] Nevertheless, some subjects appear to be attempting this, perhaps forming subjective beliefs about the end of each supergame.[24] This is a form of "gambler's fallacy"—the belief that, because the supergame has not ended in earlier rounds, the probability that it ends in the current round has increased (see, e.g., Cowan (1969)), even though the termination probability in fact remains constant.

We formally define the "end-timing" strategy as *consistently* defecting after the earlier play of cooperation in the same supergame, or DaC for short. Such an end-timing strategy involves riding a cooperative wave and gambling on its end, and thus is *risky* as it is most profitable if the first defection happens in the final round of the supergame. It is thus possible that subjects employ a δ-specific "end-timing" strategy, believing that the supergame is highly likely to end at its expected length, even though the continuation probability δ in reality does not change. Note that risk attitudes of subjects cannot in themselves explain this behavior. As noted in Section 2.1, under CRRA preferences the threshold δ^* is increasing in risk aversion, so that risk averse agents would be less likely to cooperate in our experiment, and, if it were optimal for them to defect, they would begin defecting from the first round.[25]

Indeed, Figure 7 (left) shows that, for some continuation probabilities δ, 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.[26]

Further, as Figure 7 (right) shows, the share of supergames where subjects always defected

[23] Note that in both subject pools the maximum realised overall point total was 4,185 points – which is far below the (omniscient) maximum 4,680 points from perfect play of the end-timing strategy. This suggests that there was no information exchange/leakage across subjects in our experiment.

[24] Suppose a subject believes that the continuation probability is δ in the initial rounds but (incorrectly) believes that the experimenter will stop the supergame with probability one at some final round T. Then one can calculate that, when $\delta > \delta^* = \frac{1}{2}$, the optimal strategy is to cooperate in every round up to round T but defect at round T.

[25] Dal Bó and Fréchette (2018) report that existing experiments find no clear link between risk aversion and cooperation.

[26] Note that the expected duration of a sequence, $\frac{1}{1-\delta}$, as calculated from the perspective of round 1, as well as the average realized length of the sequences in the experiment, are both increasing with δ – see Table A2. Mengel et al. (2022) report that subjects respond to the *realized* supergame length, and are more likely to cooperate when they have experienced supergames of longer duration.

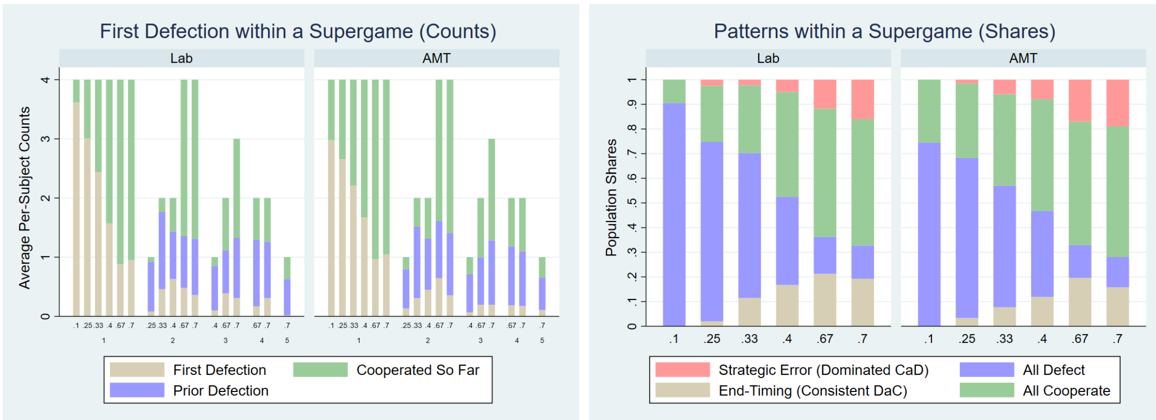

Figure 7: Patterns of cooperation and defection. Left: Average per-subject counts of first defection *within* a supergame by δ and round number. Right: The population shares of the behavioral patterns in a supergame, by δ value. By construction, the four strategies, CaD, DaC, All-D and All-C are mutually exclusive. (Lab: 100 subjects, 2,400 supergames; AMT: 149 subjects, 3,576 supergames.)

(All-D) declines as the continuation probability δ increases. However, this does not lead to greater use of the always cooperate (All-C) strategy as δ increases. Instead, as δ (and thus the expected duration of a supergame) rises, both the prevalence of strategic errors (CaD) and "end-timing" (DaC) strategies increase. Interpreting All-C strategies is complicated by attrition, as a subject may have *intended* to time their defection, but the supergame ended earlier than expected. Similarly, All-D strategies in low-δ supergames may not only reflect theoretically optimal behavior but may also be observationally equivalent to end-timing.

Indeed, if behavior were theoretically optimal, then in the mixed-effects probit regressions, the difference of the coefficients from the baseline of $\delta = 0.1$ would be insignificant for $\delta = \{0.25, 0.33, 0.4\}$ and significant for $\delta = \{0.67, 0.7\}$. Moreover, the round dummies would all differ insignificantly from the baseline of the first round. Instead, as Table 3 (specifications 1-2) reveals, subjects' tendency to choose cooperation increases with δ, but *decreases* significantly with the round number, consistent with the use of the end-timing strategy.

The left panel of Figure 8 shows that only 22% of Lab subjects and 30.9% of AMT subjects never engaged in end-timing, i.e., never switched from cooperation to defection within the same supergame (CaD). However, as the right panel of Figure 8 shows, some apparent end-timing behavior may be unintentional, reflecting "mistakes" by subjects who

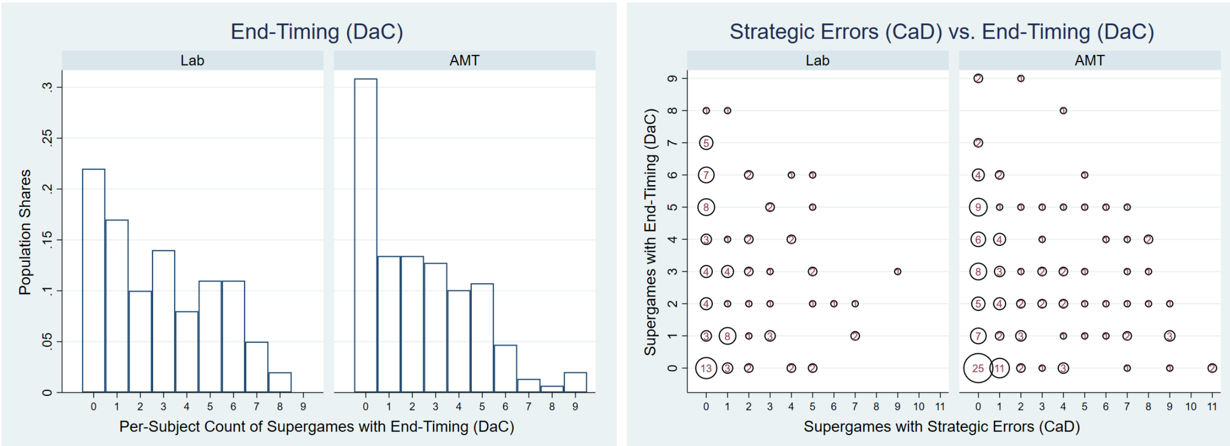

Figure 8: Left: Distribution of per-subject counts of supergames with "end-timing" (DaC) among 13 relevant supergames. Right: Per-subject counts of supergames with strategic errors (CaD) vs. those with end-timing (DaC). Bubble size is proportional to the share of subjects (100 Lab, 149 AMT).

make frequent strategic errors (CaD), i.e., those with higher counts on the horizontal axis of the scatterplots. Yet, a few subjects who rarely commit strategic errors (CaD) (those near zero on the horizontal axis) appear to engage in end-timing behavior.

Finding 5. *Some subjects appear to use a non-optimal "end-timing" strategy, attempting to time their first defection to the unknown final round of a supergame. Following this risky strategy enabled a few subjects, by chance, to earn more than the theoretically optimal payoff.*

4.6 Learning Over Time

As Table 1 shows, subjects in both pools respond to their experience, changing their play in the second half of the experiment (last 12 supergames) relative to the first (first 12). While they make fewer dominated CaD errors[27] and cooperate less per round, both subject pools show a clear increase in end-timing activity (DaC). Overall, the improvement in optimal play per round, though significant, remains modest — rising from 64% (59%) in the first half to 66% (62%) in the second for Lab (AMT) samples. Over time, only AMT subjects move toward theoretically optimal behavior within a supergame and earn slightly more points.

[27]As Figure B4 in the Appendix shows, while the incidence of dominated CaD errors declines over time, it does not disappear entirely, comprising 4% for Lab and 7% for AMT subjects in the second half of play.

Learning		1st Half Mean	1st Half StDev	2nd Half Mean	2nd Half StDev	t-stat	df	pvalue
Lab	Cooperate	0.52	0.50	0.48	0.50	2.63	4798	0.01***
Per Round	Optimal	0.64	0.48	0.66	0.47	-1.72	4798	0.09*
(N=4,800)	CaD	0.05	0.22	0.03	0.16	4.24	4798	0.00***†
	Optimal (All-D+All-C)	0.59	0.49	0.62	0.49	-1.33	2398	0.18
Lab	Optimal All-D	0.42	0.49	0.44	0.50	-1.44	2398	0.15
	Optimal All-C	0.17	0.38	0.17	0.38	0.16	2398	0.87
Per Supergame	Suboptimal All-D	0.04	0.20	0.05	0.22	-1.15	2398	0.25
	Suboptimal All-C	0.19	0.39	0.15	0.36	2.55	2398	0.01**
(N=2,400)	CaD	0.08	0.27	0.04	0.20	3.89	2398	0.00***†
	DaC (End-Time)	0.10	0.30	0.14	0.35	-3.11	2398	0.00***
	Point Total	158.20	71.33	161.40	71.46	-1.08	2398	0.28
AMT	Cooperate	0.57	0.50	0.54	0.50	2.07	7150	0.04**
Per Round	Optimal	0.59	0.49	0.62	0.49	-2.51	7150	0.01**
(N=7,152)	CaD	0.07	0.25	0.05	0.22	3.05	7150	0.00***
	Optimal (All-D+All-C)	0.52	0.50	0.57	0.50	-2.55	3574	0.01**
AMT	Optimal All-D	0.35	0.48	0.40	0.49	-2.87	3574	0.00***
	Optimal All-C	0.17	0.38	0.17	0.38	0.31	3574	0.76
Per Supergame	Suboptimal All-D	0.04	0.20	0.04	0.20	0.17	3574	0.87
	Suboptimal All-C	0.25	0.44	0.21	0.41	3.38	3574	0.00***†
(N=3,576)	CaD	0.10	0.30	0.07	0.25	3.54	3574	0.00***†
	DaC (End-Time)	0.08	0.27	0.12	0.32	-3.96	3574	0.00***†
	Point Total	154.70	71.72	159.20	72.50	-1.85	3574	0.06*

Table 1: The effect of learning, first half (first 12 sequences) vs second half (last 12 sequences): Means and standard deviations, and t-tests of differences between the two halves. DF stands for degrees of freedom or Satterthwaite's degrees of freedom in case of unequal variances for Age and Quiz Errors. Pvalue stands for $Pr(|T| > |t|) = 0$. (Significance * 0.10 ** 0.05 *** 0.01 ***† 0.001.)

Finding 6. *Over time, subjects learn to commit fewer dominated CaD errors, and move closer to theoretically optimal behavior. However, the play of dominated CaD does not disappear over time, and, moreover, the frequency with which subjects engage in end-timing behavior (DaC) increases over time.*

4.7 Classifying the Patterns of Play Within Each Supergame

In the existing IRPD experiments, it is common to classify subjects' play using more than a dozen of strategies of which Grim or Tit-For-Tat (TFT) are the most commonly known (Dal Bó and Fréchette, 2018, Section 2.5). Yet here the computer opponent is never first to defect, and never forgives, ruling out identifying these or more complex strategies, even if they could be optimal in repeated interactions among humans.

Instead, our two design innovations enable to not only distinguish between optimal and suboptimal stationary strategies All-C and All-D (by conditioning on δs) but also to document the prevalence of novel, non-stationary suboptimal strategies, end-timing and cooperate after defect, which do not fit existing classifications. Here, subjects' within-supergame behavior can be classified into only six *mutually exclusive* types: optimal All-C, optimal All-D, suboptimal All-C, suboptimal All-D, strategic errors (CaD), and end-timing (DaC).[28]

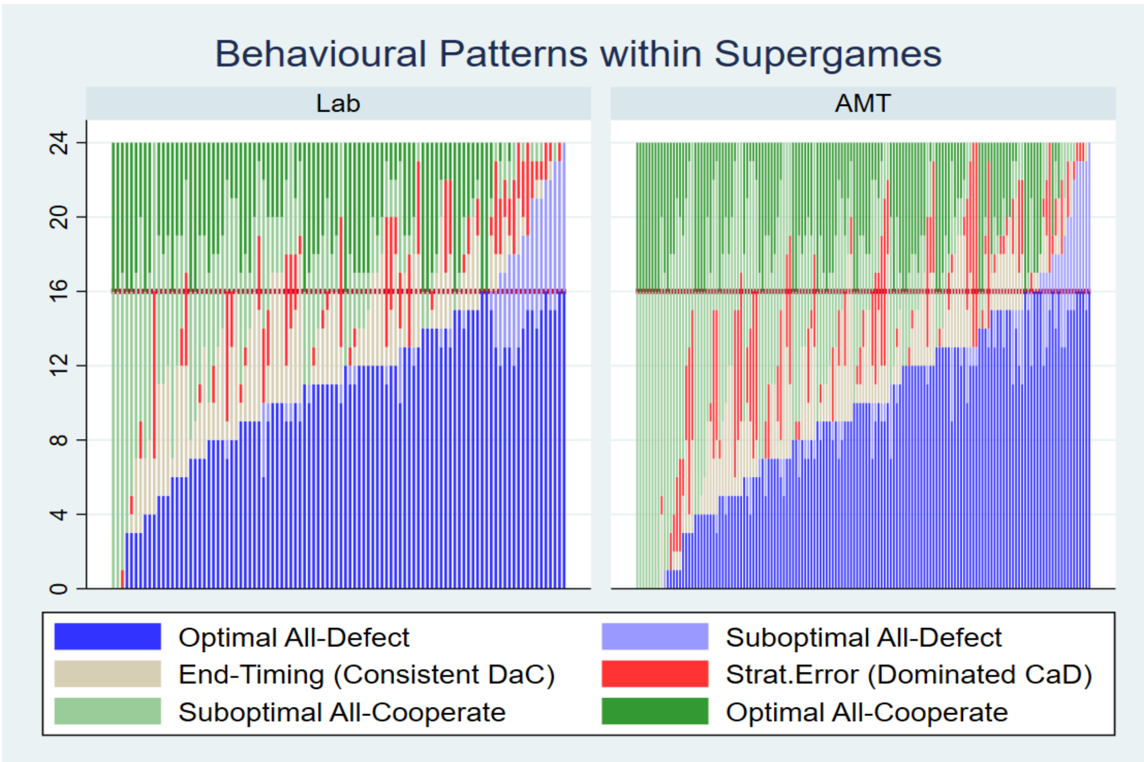

Figure 9: Subject heterogeneity in patterns of choices within 24 supergames, by subject, ordered by the count of supergames with (combined optimal and suboptimal) All-Defect choices (100 Lab subjects and 149 AMT subjects). The theoretically optimal strategy involves always defecting in 16 supergames and always cooperating in the remaining 8 (solid red horizontal line).

As Figure 9 shows, there is no dominant pattern in subjects' play. Both non-stationary play (CaD and DaC) and stationary play (All-C and All-D), whether optimal or not, tend to co-exist. The two (of 100) Lab subjects and eight (of 149) AMT subjects who always cooperated are shown by dark and light green bars meeting at the solid red line on the far left of each panel. Single subjects in each pool who always defected appear as dark and light blue bars meeting at the red line on the far right. The two Lab subjects and one AMT subject who always made theoretically optimal choices are represented by dark green and

[28]If the robot opponent instead played Tit-for-Tat, the number of possible play patterns would increase to at least eight, further complicating interpretation and analysis.

dark blue bars meeting at the red line toward the right end of each panel.

Finding 7. *There is notable heterogeneity in subjects' choices to cooperate or defect, without any representative pattern. Only two of 100 Lab subjects and one of 149 AMT subjects always followed the theoretically optimal strategy. Two (one) Lab subjects and eight (one) AMT subjects were fully biased toward cooperation (defection). The remaining subjects pursued the non-stationary suboptimal strategies of end-timing and cooperate after defect.*

5 Individual Differences and Behavioral Inattention

According to the behavioral inattention model of Section 2.2, subjects' cognitive abilities are a key determinant of their behavior. We use each subject's total score on seven Cognitive Reflection Test (CRT7) questions as a proxy measure of reflective cognitive style. The mean (st. dev.) CRT7 score was 3.78 (2.26) for Lab subjects and 3.58 (2.17) for AMT subjects, with a median of 4 for both (see Figure B6, left panel). There is no significant difference in CRT7 scores between the two subject pools (two-sided t-test = 0.72, $p = 0.48$).

5.1 The Effect of Cognitive Abilities

As Table 2 shows, consistent with Hypothesis 2 of Section 2.2, the CRT7 score is positively correlated with overall point totals (specifications 1-2) and negatively with the number of supergames involving strategic errors (dominated CaD) (specifications 3-4). For AMT subjects only, the count of theoretically optimal choices (specification 6) is positively correlated with CRT7. For Lab subjects (specification 5), the relationship is marginal, but the overall fit, as indicated by the F statistic, is poor.

For both subject pools, the count of apparent end-timing behavior (DaC) (specifications 7-8) is not correlated with the CRT7 score, or with any other characteristic. Yet, as discussed earlier in Section 4.5, choosing the theoretically optimal strategy can be observationally

	Overall Point Totals (OLS)		Dominated(CaD) (Tobit, ll=0)		Theor.Optimal (Tobit, ul=24)		End-Time(DaC) (Tobit, ll=0)		Th.Opt.+End-T(DaC) (Tobit, ul=24)	
	(1) Lab	(2) AMT	(3) Lab	(4) AMT	(5) Lab	(6) AMT	(7) Lab	(8) AMT	(9) Lab	(10) AMT
CRT7	26.37***	61.74***†	-0.51***	-0.94***†	0.32*	0.98***†	0.11	0.04	0.46**	1.06***†
	(9.32)	(7.05)	(0.16)	(0.17)	(0.19)	(0.16)	(0.14)	(0.12)	(0.19)	(0.16)
Female	-65.92	-81.49**	1.40**	2.15***	0.09	-1.05	-0.88	0.51	-0.52	-0.74
	(41.98)	(31.99)	(0.69)	(0.70)	(0.80)	(0.71)	(0.59)	(0.53)	(0.82)	(0.74)
Age	-4.85	-0.67	0.12	0.02	0.25	0.05	-0.24*	-0.04	0.07	0.02
	(9.17)	(1.50)	(0.14)	(0.03)	(0.18)	(0.04)	(0.14)	(0.03)	(0.18)	(0.04)
Order Long	57.11	-44.90	-0.22	-0.30	0.87	-1.16	0.34	-0.25	1.01	-1.42*
	(38.27)	(32.17)	(0.68)	(0.73)	(0.79)	(0.72)	(0.59)	(0.53)	(0.78)	(0.74)
Constant	3845.38***†	3635.03***†	-1.03	2.11	7.46*	8.58***†	7.47**	3.00**	13.78***	11.82***†
	(212.24)	(70.31)	(3.06)	(1.60)	(3.95)	(1.59)	(3.09)	(1.18)	(4.19)	(1.60)
F	5.47	25.26	4.52	10.11	1.44	14.89	1.80	0.76	2.72	15.18
p	0.00	0.00	0.00	0.00	0.23	0.00	0.14	0.56	0.03	0.00
Nobs.	100	149	100	149	100	149	100	149	100	149

Table 2: Individual differences in rationality. (Significance: * 0.10 ** 0.05 *** 0.01 ***† 0.001).

equivalent to an end-timing strategy. For example, when $\delta < 0.5$, immediately defecting is both optimal and consistent with end-timing. Yet, by construction, the theoretically optimal and DaC plays are *mutually exclusive*. To address this, we report, in specifications 9 and 10, a combined count of whether subjects follow either the theoretically optimal strategy (i.e., either All-D for $\delta < 0.5$ or All-C for $\delta > 0.5$), or engage in end-timing (i.e., consistent defection after cooperation, or DaC). This *is* positively correlated with the CRT7 score for both subject pools, in contrast to the above negative correlation with dominated CaD errors.

In contrast, the CRT7 score alone has no effect on the choice to cooperate, as seen in Table 3 (specifications 3–4), which reports average marginals (dy/dx) from mixed-effects probit regressions of the choice to cooperate or defect, controlling for demographics, CRT7 score, and other personal characteristics.[29,30]

Finding 8. *Subjects with higher CRT7 scores earn higher payoffs, make fewer errors, and behave closer to theoretical predictions. They may also be more likely to engage in end-timing.*

[29] Note that Table 3 presents marginals, rather than odds, so that the same explanatory variable in different models can have different statistical significance despite similar coefficients and robust errors.

[30] Note that while the coefficient on the female dummy in specifications 3-4 of Table 3 is significantly negative, and the CRT7 score is negatively correlated with being female for the Lab subjects ($r = -0.2365, p = 0.0178$) but not for the AMT subjects ($r = -0.0625, p = 0.4497$), the coefficient on the CRT7 score remains insignificant if we exclude the age and gender demographic variables, or other individual characteristics (results available on request).

Cooperate	All				$CRT7 \leq 4$		$CRT7 > 4$	
(Marginals, dy/dx)	(1) Lab	(2) AMT	(3) Lab	(4) AMT	(5) Lab	(6) AMT	(7) Lab	(8) AMT
δ=0.25	0.23***†	0.08***†	0.22***†	0.08***†	0.24***†	0.08***	0.18***†	0.09**
	(0.03)	(0.02)	(0.03)	(0.02)	(0.04)	(0.03)	(0.04)	(0.03)
δ=0.33	0.31***†	0.18***†	0.30***†	0.18***†	0.31***†	0.14***†	0.28***†	0.24***†
	(0.03)	(0.03)	(0.03)	(0.03)	(0.05)	(0.03)	(0.05)	(0.04)
δ=0.4	0.47***†	0.28***†	0.46***†	0.28***†	0.44***†	0.21***†	0.46***†	0.37***†
	(0.04)	(0.03)	(0.04)	(0.03)	(0.05)	(0.04)	(0.06)	(0.04)
δ=0.67	0.66***†	0.42***†	0.65***†	0.42***†	0.61***†	0.29***†	0.68***†	0.60***†
	(0.04)	(0.03)	(0.04)	(0.03)	(0.06)	(0.04)	(0.06)	(0.04)
δ=0.7	0.69***†	0.46***†	0.68***†	0.45***†	0.62***†	0.33***†	0.73***†	0.63***†
	(0.04)	(0.04)	(0.04)	(0.04)	(0.06)	(0.04)	(0.06)	(0.05)
Round 2	-0.04	0.03	-0.04	0.03	-0.00	0.03	-0.09**	-0.02
	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.04)	(0.03)
Round 3	-0.12***†	0.04	-0.12***†	0.04	-0.06	0.06*	-0.20***†	-0.06
	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)	(0.05)	(0.04)
Round 4	-0.19***†	-0.02	-0.18***†	-0.02	-0.10*	0.04	-0.28***†	-0.16***†
	(0.04)	(0.03)	(0.04)	(0.03)	(0.05)	(0.04)	(0.06)	(0.04)
Round 5	-0.17***	-0.13***	-0.17***	-0.13***	-0.13*	-0.10*	-0.23***	-0.23***†
	(0.05)	(0.04)	(0.05)	(0.04)	(0.07)	(0.06)	(0.08)	(0.06)
Supergame	-0.06***	-0.06***	-0.06***	-0.06***	-0.09***	-0.04	-0.03	-0.09****
	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)
Order Long	0.05	0.03	0.06	0.04	0.07	0.06	0.04	0.04
	(0.04)	(0.04)	(0.04)	(0.04)	(0.06)	(0.06)	(0.05)	(0.06)
Prior Defect	-0.20***†	-0.23***†	-0.20***†	-0.23***†	-0.17***†	-0.19***†	-0.23***†	-0.22***†
	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)	(0.05)	(0.03)
CRT7			-0.00	0.01				
			(0.01)	(0.01)				
Prediction			0.24***	0.19***	0.19	0.08	0.23****	0.24***
			(0.09)	(0.06)	(0.15)	(0.08)	(0.07)	(0.09)
Female			-0.08**	-0.08**	-0.08	-0.11**	-0.07	-0.05
			(0.04)	(0.04)	(0.06)	(0.05)	(0.06)	(0.05)
Age			-0.01	-0.00	-0.01	-0.00	-0.01	0.00
			(0.01)	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)
Risk			-0.00	-0.01	-0.01	-0.02*	0.00	-0.00
			(0.01)	(0.01)	(0.02)	(0.01)	(0.02)	(0.02)
Patience			0.02	0.01	0.04***†	0.03*	-0.02	-0.01
			(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)
Punishment			0.00	-0.01	0.01	0.00	-0.01	-0.02**
			(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)
Altruism			-0.02	-0.00	-0.02	-0.01	-0.02	-0.00
			(0.01)	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)
Reciprocity			0.02	-0.01	0.01	-0.02	0.05	0.00
			(0.02)	(0.01)	(0.02)	(0.01)	(0.04)	(0.01)
Retribution			-0.01	0.01	0.00	0.01	-0.02**	0.01
			(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)
Trust			0.00	0.01	-0.01	0.00	0.02*	0.00
			(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
chi2	266.09	203.19	406.22	231.92	227.73	110.77	200.85	218.72
p	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
N	4800	7152	4800	7152	2688	4512	2112	2640

Table 3: Choices to cooperate: mixed-effects probit regressions, marginals (dy/dx), robust errors in parentheses. (See Table B4 for the corresponding odds.) “Supergame” is the supergame number in the sequence of supergames (scaled down by 24),“Order Long” is a dummy variable for whether the first supergame in the sequence had $\delta = 0.67$, “Prior Defection” is a dummy variable for whether the subject defected in prior rounds of a given supergame, “Prediction” is the subjects’ predictions of the share of their own cooperative choices in Round 1 across all 24 supergames (scaled down by 100). Chi2 and corresponding p-values are from the odds regressions (see Table B4). (Significance * 0.10 ** 0.05 *** 0.01 ***† 0.001.)

5.2 Behavioral Inattention

The less obvious prediction of Hypothesis 2 of Section 2.2 is that, when deciding whether to cooperate or defect, individuals with lower cognitive ability (lower attention) tend to rely more on default values, while those with higher cognitive ability (higher attention) respond more to the structure of the game. To test this hypothesis, we split both samples by the median CRT7 score (equal to 4). Table 3 reports average marginals (dy/dx) from mixed-effects probit regressions of the choice to cooperate or defect across all 48 rounds of the prisoners' dilemma (for odds, see Table B4 in the Appendix). As specifications 7–8 show, subjects with higher proxies for cognitive ability (CRT7>4) respond more strongly to the continuation probability δ, are less likely to cooperate after defection within the same supergame (i.e., make fewer CaD errors), and more frequently follow the end-timing strategy, as indicated by negative and significant round-dummy coefficients. In contrast, as specifications 5–6 show, lower-ability subjects (CRT7$\leq$4) exhibit no systematic sensitivity to the round number. These regression results are summarized visually in Figure 10.

These regressions further show that both lower and higher CRT7 groups each have a single significant factor *consistently* correlated with their decision to cooperate across both samples. In the lower CRT7 group (Table 3, specifications 5–6), those with higher self-reported *Patience* tend to cooperate more frequently (and, vice versa, those with higher self-reported impatience tend to defect more frequently)[31] — significantly among Lab subjects and marginally among AMT subjects.[32] Interestingly, Fehr and Leibbrandt (2011) find that patience measures are highly correlated with cooperative behavior in a field experiment. Patience has also been found to influence cooperation in laboratory experiments (Davis et al. (2016)). Other correlations in the lower CRT7 group are significant in only one subject pool: being female was significantly negatively correlated with cooperation for the AMT

[31] The proxy for Patience is taken from Falk et al. (2018): "How willing are you to give up something that is beneficial for you today in order to benefit more from that in the future?" (see Appendix A.4).

[32] Importantly, the Patience measure was marginally higher for the higher CRT7 group of the Lab subjects (one-sided $t = 1.3718$, $p = 0.0866$), and there was no difference between the two CRT7 groups of the AMT subjects (two-sided $t = 0.2802$, $p = 0.7798$).

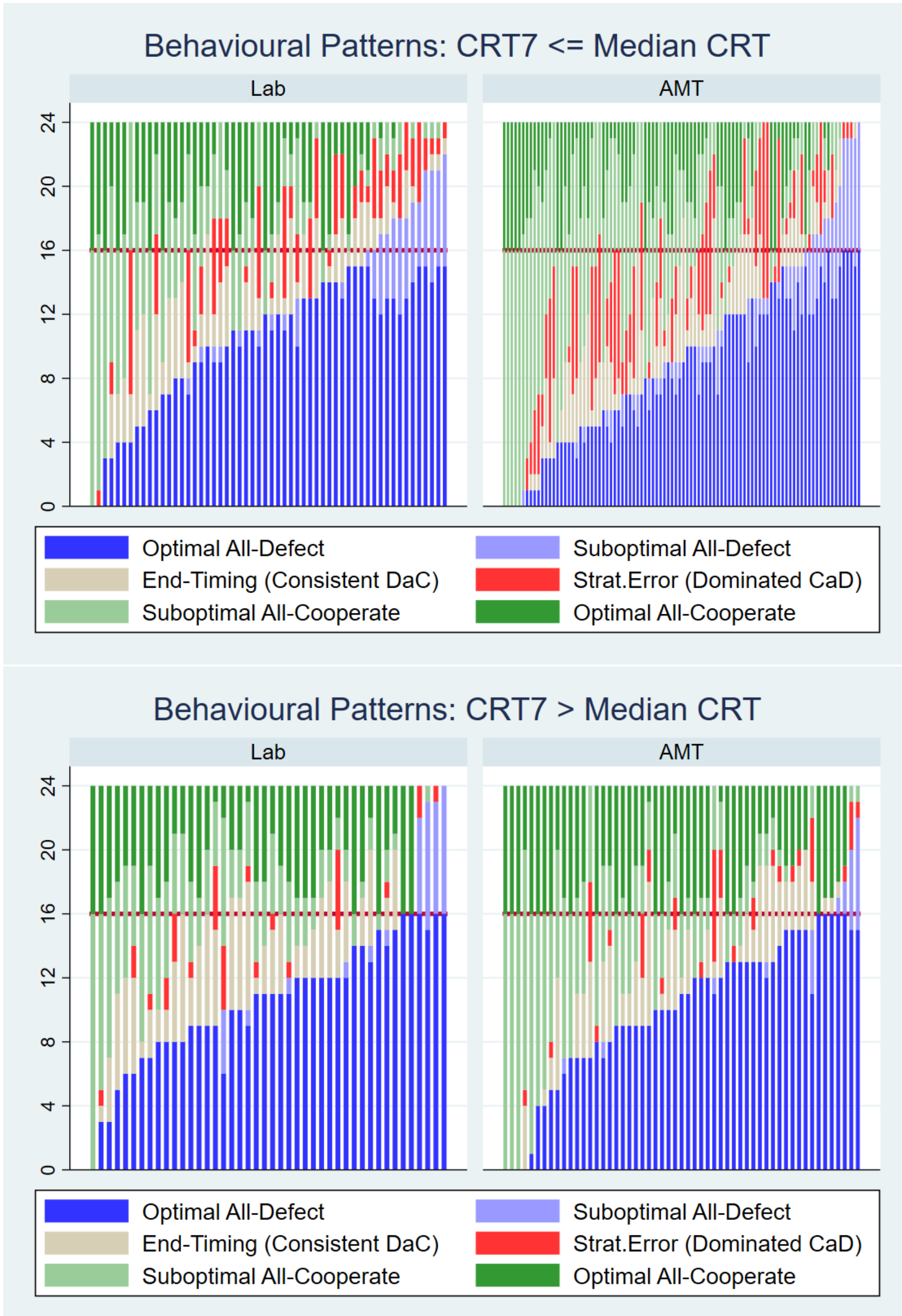

Figure 10: Inattention: each subjects' patterns of play within supergames (out of 24 supergames), split by median $CRT7$. Patterns are ordered by the count of supergames with combined optimal and sub-optimal All-Defect choices. The theoretically optimal strategy is represented by the horizontal line, with always defecting in 16 supergames and always cooperating in 8 supergames. (100 Lab and 149 AMT subjects.)

subjects only, and risk-prone AMT subjects only cooperated marginally less frequently.[33]

[33]The Risk measure is also from Falk et al. (2018): "In general, how willing are you to take risks?" (see

In contrast, in the higher CRT7 group (Table 3, specifications 7–8), no personal characteristics systematically explain subjects' cooperative choices, as predicted by behavioral inattention theory. (Cooperation was, however, significantly negatively correlated with Retribution for Lab subjects and with Punishment for AMT subjects.) Instead, Round 1 choices (the regression baseline) are strongly correlated with the "Prediction" variable, elicited before any choices were made, possibly reflecting subjects' understanding of the task.[34]

Finding 9. *Subject behavior is broadly consistent with a simple model of inattention. Cooperative choices by subjects with lower proxies for cognitive ability (higher attention) correlate with an elicited proxy for patience. In contrast, cooperative choices by those with higher proxies for cognitive ability (lower attention) depend more on the game's structure, correlating not with individual characteristics but with an elicited proxy for understanding that structure.*

6 Laboratory Robustness Check with Higher δ Values

We conducted additional sessions to address referees' concerns about the robustness of our results.[35] These sessions differed in four key respects. First, to ensure full experimental control, they were conducted in person at the UC Irvine laboratory: instructions were read aloud, and subjects completed all tasks individually at workstations. Second, we used higher continuation probabilities, $\delta \in \{0.1, 0.4, 0.67, 0.75, 0.8, 0.85\}$, producing longer supergames comparable to prior repeated-game studies. Third, δ values were implemented in blocks of

Appendix A.4). As Table 3 shows, there is no robust, significant relationship between this proxy for risk propensity and cooperation in the first rounds. This contrasts with the theoretical prediction in Section 2.1 that risk-averse subjects should cooperate less, and risk-prone subjects more.

[34] Before making any choices, subjects were asked what percentage of their Round 1 choices across all 24 supergames would be cooperative (see Section 3). The mean (st. dev.) of this "Prediction" variable for Lab subjects was 60.96 (29.57) with a median of 62, while for AMT subjects it was 66.64 (27.04) with a median of 72, with no significant difference between the two pools (two-sided t-test = 1.56, $p = 0.12$, see Figure B6, right panel). By comparison, the optimal choice is $\frac{1}{3}$ (33.33%). In a Tobit regression, this measure is significantly negatively correlated with "Altruism" ($p = 0.01$) and marginally positively correlated with "Retribution" ($p = 0.10$), but only for AMT subjects; for Lab subjects, there is no correlation (results available on request). Finally, the Prediction variable is significantly correlated with subjects' actual first-round choices ($r = 0.3517$, $p = 0.0003$ for Lab and $r = 0.2350$, $p = 0.0039$ for AMT).

[35] A pre-registration of the design of these new sessions can be found at the AEA RCT Registry `https://doi.org/10.1257/rct.15844`.

four supergames rather than randomly, allowing learning within each δ-block; instructions were modified accordingly. Finally, we added an incentivized risk-elicitation task following Holt and Laury (2002), as risk attitudes may influence optimal play (see Section 2.1); the original experiment used only an unincentivized Likert-scale measure. Implementation details appear in Appendix C.1–C.3.

We recruited a new, gender-balanced panel of 101 UC Irvine undergraduates distinct from the original sample. Average earnings were \$26.52—higher than in the initial sessions due to longer games under higher average δ and a \$3 increase in the show-up payment.

These modifications provide several robustness checks on our original findings. Despite the changes, behavior closely parallels the main treatment: error rates remain high, strategies are non-stationary, cooperation often follows defection, and many subjects exhibit end-timing. Elicited risk attitudes show no explanatory power.

Each subject completed 24 supergames, each lasting at least one round. Eleven ended after a single round, while the remaining 13 lasted 2–15 rounds, yielding 24 first-round and 75 subsequent-round choices (99 total decisions; see Appendix Table C2). Given the parameters, the optimal strategy is to defect in all rounds of the eight supergames with $\delta \in \{0.1, 0.4\}$ (11 decisions) and to cooperate in all rounds of the remaining sixteen supergames (88 decisions).

6.1 High Delta Sessions: The Headline Result

In the high-δ treatment, most δ values exceed $\delta^* = 0.5$ by design, so theoretically optimal choices are skewed toward cooperation both initially and later; hence, in this follow-up experiment, defection is often suboptimal. The theoretically optimal strategy here is to defect in all rounds of supergames with $\delta = \{0.1, 0.4\}$ and to cooperate in all others. Thus, perfectly optimal play entails 16 cooperative first-round choices (and 8 defections) and 72 cooperative choices in the remaining 75 rounds (and 3 defections), a total of 88 cooperative choices out of 99.

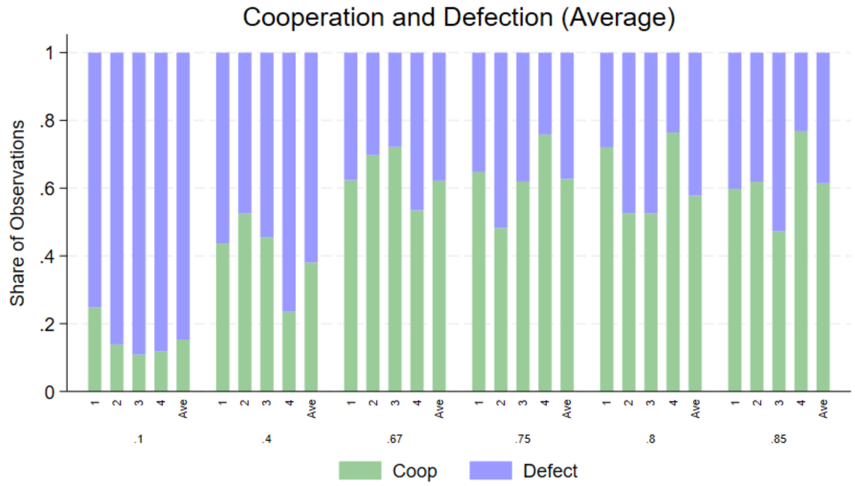

Figure 11: High δs: Cooperation and Defection rates across all rounds of the four supergames in chronological order (first horizontal axis) of each of the six δ values (second horizontal axis). (N=101 subjects.)

Figure 11 (the high-δ analogue of Figure 1 for the baseline treatment) summarizes behavior in the high-δ treatment. Overall, the pattern closely resembles the baseline: the rate of cooperative play is positive for the two δ values below the cutoff $\delta^* = 0.5$ and is less than full cooperation for all four δ values above δ^*, contrary to Hypothesis 1. Learning across supergames, while being potentially better than in the baseline, remains mixed, despite here the four supergames of each δ were played consecutively, potentially facilitating learning.

These aggregate patterns translate into low rates of fully optimal behavior. As Figure 12 (the high-δ analogue of Figure 2) shows, only 3 of 101 subjects (3%) played perfectly optimally. Only 8 subjects (7.9%) made at least 93 optimal choices out of 99 (93.9%), slightly above the share in the main experiment, while 17 subjects (16.8%) achieved at least 87 optimal choices (87.9%), versus roughly 10% in the main experiment. Overall, the mean (s.d.) number of optimal choices was 57.2 (26.41) with a median of 60, and subjects made significantly fewer optimal choices than predicted (one-sided t-test: $t = 15.8992$, $p < 0.001$).

Finding 10. *In high-δ Lab sample, across 99 decisions in 24 supergames, fewer than 8% of subjects behaved in line with standard game-theoretic predictions. About 83% made fewer than 87 of 99 theoretically optimal choices (87.9%).*

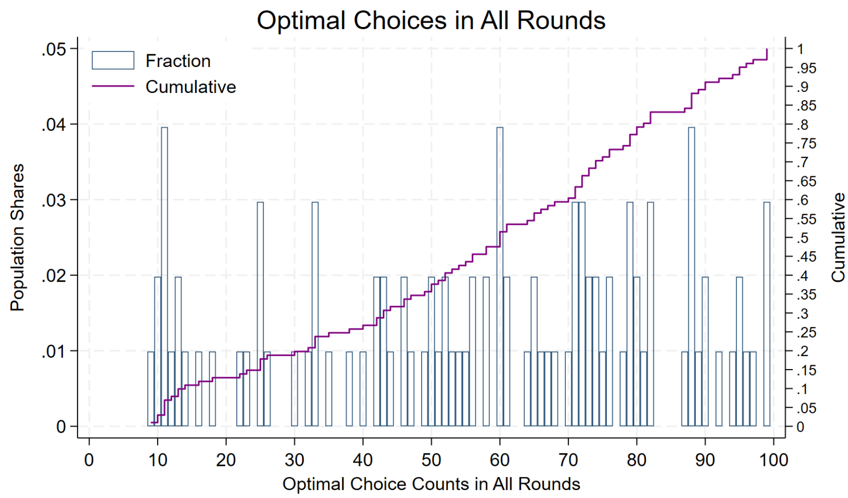

Figure 12: High δs: Frequency and cumulative distributions of per-subject counts of optimal choices across all supergames (99 decisions per subject, N=101 subjects).

6.2 High Delta Sessions: Response to δ

In the first rounds of supergames in the high-δ sessions, subjects' cooperation increases almost monotonically with the continuation probability δ, though this tendency flattens at higher values (similar to the baseline). This pattern, visible in the leftmost group of bars in the right panel of Figure 13, is confirmed by the mixed-effects probit regressions in Table C5, specifications 1-2. First-round cooperation rates are 15.4% when $\delta = 0.1$, 70.05% when $\delta = 0.67$, and 77.7% when $\delta = 0.85$, consistent with previous findings (see footnote 18). As in the main experiment, subjects respond less when continuation probabilities are intermediate - see the left panel of Figure 13 - and nearly half cooperate suboptimally in the first round when $\delta = 0.4 < \delta^* = 0.5$. Such excessive cooperation cannot be explained by risk aversion (further confirmed by our risk-aversion measure), while subjects again cooperate too little when $\delta > \delta^*$, particularly at $\delta = 0.67$ (right panel of Figure 13).

Finding 11. *In high-δ Lab sample, in every round of a supergame, the rate of cooperation (defection) tends to (weakly) increase (decrease) with continuation probability δ. However, there is a pronounced tendency toward excessive cooperation at the "intermediate" continuation probability $\delta = 0.4$.*

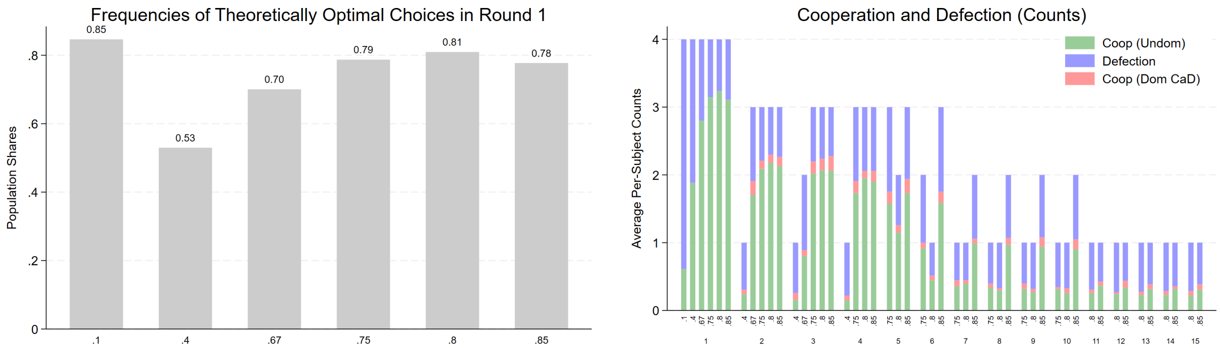

Figure 13: High δs: Patterns of optimality, cooperation, and defection: Left panel: Population shares of theoretically optimal first-round choices across all supergames, by continuation probability δ (see also the leftmost bars in the right panel). Right panel: Average per-subject counts of cooperation versus defection, by continuation probability δ (first row of the horizontal axis) and by supergame round number (second row of the horizontal axis). From round 2 onward, we distinguish between undominated cooperation and dominated *cooperation after defection* (CaD). Later rounds were not reached for some δ values (see Table C2). (N=101 subjects, 2,424 supergames.)

6.3 High Delta Sessions: First Round Behavior

Figures 13 and 14 show the subjects' first-round behavior in the high-δ session supergames (before any experience). Again, only a small fraction, 7 subjects out of 101 (6.9%), behaved perfectly as predicted and made 16 (8) choices to cooperate (defect) in the first rounds across all 24 supergames. In contrast to the baseline experiment, 30 out of 101 subjects (29.7%) made no more than 3 mistakes out of 24 first round choices (i.e., at least 21 optimal choices out of 24, or 87.5% of choices). However, given the mean (st.dev.) of optimal decisions in the first round of 17.8 (4.60) and the median of 19, the subjects made significantly fewer optimal decisions in the first round than predicted (one-sided t-test $t = 13.5346, p = 0.0000$).

Finding 12. *In high-δ Lab sample, in the first rounds, compared to the theoretical optimum, subjects behave suboptimally, cooperating too much (too little) for low (high) continuation probabilities.*

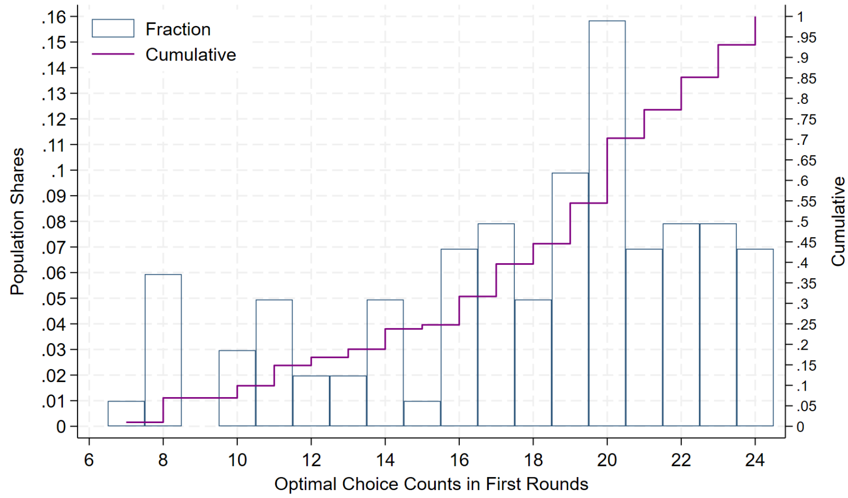

Figure 14: High δs: Left: Frequency and cumulative distributions of per-subject counts of optimal choices in first round choices across all 24 supergames. (N=101 subjects, 24 decisions each.)

6.4 High Delta Sessions: Cooperation after Defection (CaD)

Section 4.3 introduced the *strategic error* of *dominated* choice to cooperate after defecting earlier within the *same* supergame (CaD). In the right panel of Figure 13, among the relevant observations (in the 13 supergames lasting more than one round), the overall share of CaD errors is 5.7%, lower than in the baseline study. However, as Figure 15 (left) shows, the incidence of such errors per subject is greater than in the baseline, as fewer subjects, only 39.6%, never made the strategic error of CaD, and 34.65% of subjects made at least 4 dominated CaD choices. Figure 15 (right) shows that, again, at 33.7%, a greater share of subjects make errors in at least 3 out of the 13 relevant supergames than in the main experiment (see an example of subject 207 making CaD errors in 10 supergames in Table 4). Furthermore, only 36% of subjects made only one strategic CaD error per supergame.

The increased frequency of mistakes per subject in the follow-up experiment could be simply due to longer supergames in the follow-up study, with subjects making more than twice as many choices in the follow-up study (99 vs 48), increasing the chances of the "trembling hand" mistake. While the prevalence of CaD errors is relatively small, it nevertheless underscores the challenge of interpreting the deviations from the theoretically optimal behavior in the more complex environments of the classical RIPD experiments, as those involve relatively long supergames and thus greater number of choices.

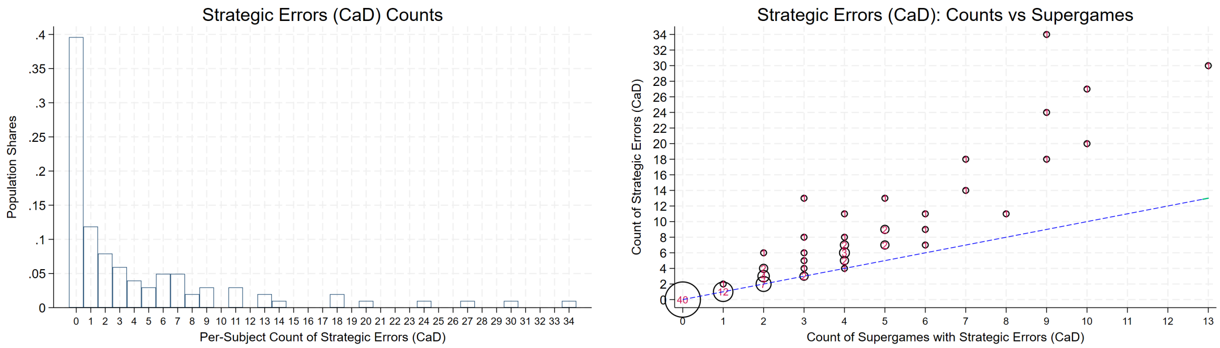

Figure 15: High δs: Strategic errors of dominated cooperation after defection (CaD) in 13 relevant supergames (i.e., those lasting longer than one round). Left: Distribution of per-subject counts of instances of cooperation after defection (CaD). Right: Per-subject counts of CaD instances vs. count of supergames with those instances. Bubble size is proportional to the share of subjects (N=101 subjects).

Finding 13. *In high-δ Lab sample, about three-fifths of subjects made at least one strategic error of cooperating after previously defecting (CaD) within the same supergame, and roughly one-third of subjects made dominated choices in at least 3 of the 13 supergames lasting more than one round. Such excessive cooperation accounts for 5.7% of relevant observations.*

6.5 High Delta Sessions: "End-Timing" (DaC)

Recall that the theoretically optimal strategy is *stationary*—either defect immediately when $\delta < \delta^*$ or never defect at all. However, against a Grim-trigger robot, a subject could do better by perfectly timing defection — cooperating in all but the final round of each supergame and defecting only in the last round to gain the temptation payoff without triggering any punishment. In Section 4.5, we defined the "end-timing" strategy as *consistently* defecting after the earlier play of cooperation in the same supergame, or DaC.

Indeed, Figure 16 (left) shows that, for all continuation probabilities $\delta > 0.1$, some subjects follow non-stationary strategies: they defect for the first time (thus triggering subsequent defection by the automated opponent) *later* in the supergame, rather than in the first round (if ever) as predicted by the theory. Furthermore, as Figure 16 (right) shows, as δ (and thus the expected duration of a supergame) increases, both the prevalence of strategic errors (CaD) and "end-timing" (DaC) strategies broadly increase.

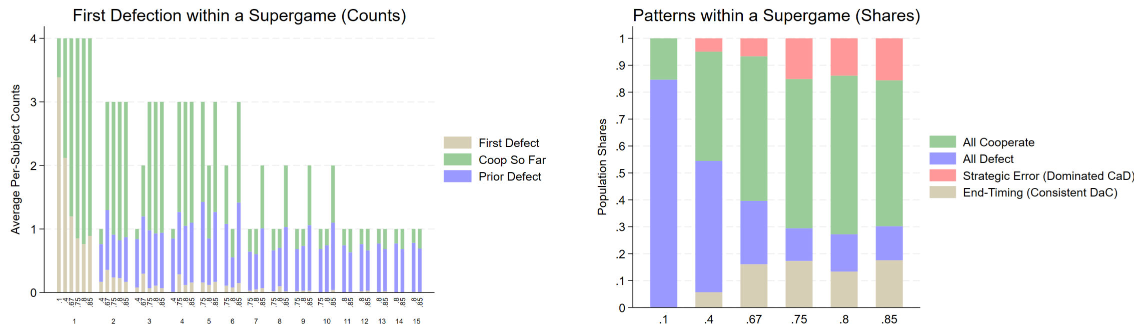

Figure 16: High δs: Patterns of cooperation and defection. Left: Average per-subject counts of first defection *within* a supergame by δ and round number. Right: The population shares of the behavioral patterns in a supergame, by δ value. By construction, the four strategies, CaD, DaC, All-D and All-C are mutually exclusive. (N=101 subjects, 2,424 supergames.)

Again, by looking at the mixed-effects probit regressions reported on in Table C5 (specifications 1-2), we observe that subjects' tendency to choose cooperation broadly increases with δ rather than change in a step-wise fashion predicted by the standard theory. The effect of the round number is not as clear cut as in the baseline treatment (Table 3), but there is also a tendency for cooperation to broadly *decrease* with the round number for relatively low round numbers, which is broadly consistent with the use of the end-timing strategy.

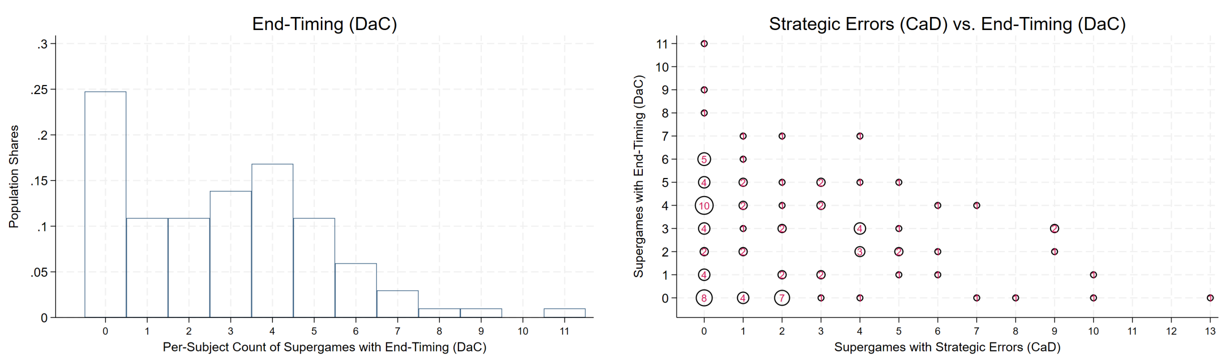

Figure 17: High δs: Left: Distribution of per-subject counts of supergames with "end-timing" (DaC), among 13 relevant supergames. Right: Per-subject counts of supergames with strategic errors (CaD) vs. supergames with end-timing (DaC). Bubble size is proportional to the share of subjects (N=101 subjects).

The left panel of Figure 17 shows that only 24.75% of subjects never used the end-timing strategy. As the right panel of Figure 17 shows, some potential end-timing behavior may be unintentional, amounting to "mistakes" by subjects who make frequent strategic errors (CaD), i.e., those with higher counts on the horizontal axis of the scatterplots. Yet a few subjects who never or almost never commit strategic errors (CaD) (those closer to zero on the horizontal axis indicating CaD errors) appear to be engaged in end-timing behavior.

Recall that the baseline treatment, by design, is biased toward lower continuation probabilities δ and thus shorter supergames on average. As a result, interpreting subjects' strategies there is complicated by attrition, as a subject might have *intended* to time their defection, but a supergame ended earlier than expected. In contrast, our high-δ treatment is biased toward higher continuation probabilities so supergames last longer on average and display greater variance in duration. Consequently, some lasted substantially longer (or shorter) than expected, allowing us to clearly detect when subjects used end-timing strategies. For example, subject 503 in Table 4 appears to have "gambled" by defecting in what they believed was the final round—though, given the realized draws, this strategy proved *ex post* unsuccessful (all subjects' play patterns are available upon request). As discussed earlier, such a strategy is less risky when continuation probabilities δ are relatively low (as in the baseline experiment) than when they are high, as in the follow-up. As shown in Appendix D, neither risk-attitude measure affects the likelihood of pursuing end-timing (see Table C4).

Finding 14. *In the high-δ Lab sample, some subjects appear to use risky "end-timing" strategy where they attempt to time their first defection to the unknown final round of a supergame. Given the higher variance of the duration of the long-horizon supergames, no subject was able to earn more than the theoretically optimal payoff in the long-horizon treatment.*

6.6 High Delta Sessions: Overall Point Totals

Given the realization of random supergame terminations in high-δ treatment, the *ex ante* optimal play would result in an overall total of 7,650 points earned over all 99 decisions

Supergame	δ	Duration	Subject 207 ($CRT7 = 0$)	Subject 503 ($CRT7 = 5$)
1	0.67	3	DDD============	CCD============
2	0.67	2	CC=============	CC=============
3	0.67	1	C==============	C==============
4	0.67	3	DDC============	CCD============
5	0.85	6	DDDDDD=========	CCCCCD=========
6	0.85	1	D==============	C==============
7	0.85	15	CCDDDCDDCCDCCDD	CCCCCDDDDDDDDDD
8	0.85	10	DCDDCDCDDD=====	CCCCCCDDDD=====
9	0.10	1	C==============	D==============
10	0.10	1	D==============	D==============
11	0.10	1	D==============	D==============
12	0.10	1	C==============	D==============
13	0.80	5	CDCDD==========	CCCCD==========
14	0.80	15	CCDCDCDCDCCDCDD	CCCCDDDDDDDDDDD
15	0.80	4	DCDD===========	CCCC===========
16	0.80	1	C==============	C==============
17	0.75	6	DCDDCD=========	CCCDDD=========
18	0.75	1	C==============	C==============
19	0.75	10	DDCDCDCDCD=====	CCCDDDDDDD=====
20	0.75	5	DCCDC==========	CCCDD==========
21	0.40	4	DDCD===========	CDDD===========
22	0.40	1	C==============	C==============
23	0.40	1	D==============	C==============
24	0.40	1	C==============	C==============

Table 4: High δs: Examples of subjects engaging in non-stationary play. Subject 207: dominated cooperation after defection (CaD) vs. subject 503: end-timing (DaC). Both subjects faced Long order.

ex post.[36] As in the baseline study, the lowest possible *ex post* point total (or the "fixed" component of the overall point total) is again quite substantial at 3,510 points.[37]

Figure 18 reports on subjects' total awarded points earned over all 99 decisions. Empirically overall point totals range from $4,620$ to $7,650$ points, with two modes (at 4 subjects each) at 7,425 (All-Cooperate) and slightly lower 7,155, while 3 subjects behaved perfectly theoretically optimal earning 7,650 points. The mean (st.dev.) is 6,428.8 (867.7) - higher than in the baseline study simply because of the longer supergames. On average, subjects earned less than in the baseline study, only 70.5% of the "variable" component of the overall point totals achievable by following the theoretically optimal policy.[38] The worst performing subject earned only 26.8% of the "variable" component achievable by following the optimal

[36] The theoretically optimal point total is a sum of 75 points in each round of supergames with $\delta \in \{0.67, 0.7, 0.8, 0.85\}$ (88 decisions) plus 120 points in the first rounds (8 decisions) and 30 points in the subsequent rounds (3 decisions) of the supergames with $\delta \in \{0.1, 0.4\}$.

[37] As in footnote 21, the *ex post* theoretical minimum of 3,510 points arises from CaD errors, with 75 points from cooperating in each of the 11 supergames lasting only one round (11 choices); plus 120 points from defecting in the first round (13 choices) and 15 points from cooperating thereafter (75 choices) in each of the 13 supergames lasting longer than one round).

[38] As in the baseline study, this is calculated as the share of the average point totals in excess of the theoretical minimum relatively to the 4,140 theoretically optimal point totals in excess of the minimum.

policy (again, quite miserably relatively to the baseline study). Moreover, 11.9% of subjects could have *increased* their total point totals to the level of 5,130 points (39.1% of the optimal "variable" component) by simply choosing to always defect (All-D). And, strikingly, 85.6% of subjects could have earned 7,425 point totals (94.6% of the optimal "variable" component) by choosing to always cooperate (All-C)! (Recall that the follow-up experiment has higher deltas, and thus is biased by design towards cooperation.) These observations suggest that, among other deviations, strategic errors (CaD) reduce overall point totals.

Finally, if subjects were omniscient and perfectly knew when will be the final round of a supergame, the end-timing (DaC) strategy in our high-δ treatment yields an *ex post* maximum point total of 8,505 — about 11% higher than the theoretical maximum. Yet, here, no subject earned more than the theoretically optimal payoff, confirming that the end-timing behavior is not profitable in a riskier environment of high continuation probabilities (with a higher chance of both very short and very long supergames).[39]

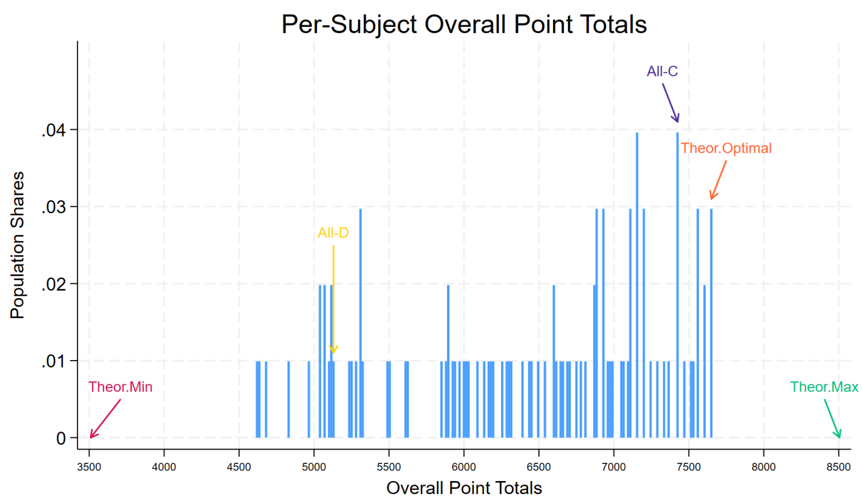

Figure 18: High δs: Distribution of overall point totals, or the sum of point earnings across all 99 decisions. *Ex post* theoretical point total from following the *ex ante* optimal policy is 7,650, while the ex post theoretical minimum and (omniscient) maximum point totals are 3,510 and 8,505, respectively. (N=101 subjects.)

Finding 15. *In high-δ Lab sample, on average, subjects earned only 70.5% of what could be achieved relative to the* ex post *theoretical minimum by following the* ex ante *optimal policy.*

[39] Again, this also suggests there was no information exchange or leakage between participants, or that any such information was not used to maximize payoffs.

11.9% of subjects achieved lower point totals than what they could have achieved by always defecting. No subject was able to achieve more than the theoretically optimal point totals.

6.7 High Delta Sessions: Learning Over Time

Do subjects learn to play the theoretically optimal strategy and make fewer dominated errors (such as CaD) as they gain experience in the high-δ sessions, where supergames with the same δ value are played consecutively in blocks? Here, as in the baseline treatment, there is clear evidence of learning over time, from the first to the second halves of the sessions (see the top panel of Table 5, which is comparable to Table 1), though interpretation is complicated by the order of δ-blocks (see Appendix Figure C5). Subjects make significantly fewer dominated CaD errors as they gain experience, although such errors do not vanish entirely, occurring in 5.7% of supergames in the second half of play. They also cooperate more often, shifting away from both optimal and suboptimal All-D strategies and toward All-C strategies. Given that the high-δ treatment naturally favors cooperative play, this behavioral shift results in a higher overall frequency of optimal choices.

Unlike in the baseline treatment, subjects here faced each continuation probability δ in blocks of four supergames. The bottom panel of Table 5, considers learning *within these δ-blocks*, between the first and last two supergames of the same δ-value. We find that subjects make fewer optimal decisions and use "end-time" (DaC) strategies significantly more often in the second half of these δ-blocks. This pattern is reflected in fewer optimal and more suboptimal All-C choices. Interestingly, the frequency of strategic (CaD) errors remains statistically unchanged within δ-blocks.

Finding 16. *In high-δ Lab sample, from the first to the second halves of sessions, subjects make fewer dominated CaD errors and move closer to the theoretical optimum. However, dominated CaD play persists, and end-timing (DaC) behavior increases—significantly so only within a δ-block.*

Other findings for the high-δ treatment closely mirror those reported earlier for the base-

Learning: Within Session		1st Half (SG 1-12)		2nd Half (SG 13-24)		t-stat	df	pvalue
		Mean	StDev	Mean	StDev			
Per Round	Cooperate	0.51	0.50	0.64	0.48	-13.13	9997	0.00****
	Optimal	0.52	0.50	0.64	0.48	-12.19	9997	0.00****
(N=9999)	CaD	0.06	0.24	0.03	0.16	7.69	9997	0.00****
Per Supergame	Optimal (All-D+All-C)	0.56	0.50	0.63	0.48	-3.60	2422	0.00****
	Optimal All-D	0.25	0.43	0.20	0.40	2.89	2422	0.00***
	Optimal All-C	0.31	0.46	0.43	0.50	-6.19	2422	0.00****
(N=2424)	Suboptimal All-D	0.13	0.34	0.08	0.27	4.48	2422	0.00****
	Suboptimal All-C	0.07	0.26	0.11	0.32	-3.50	2422	0.00****
	CaD	0.13	0.34	0.06	0.23	6.25	2422	0.00****
	DaC (End-Time)	0.11	0.31	0.13	0.33	-1.20	2422	0.23
	Point Total	259.10	227.20	276.60	249.60	-1.80	2422	0.07*
Learning: Within Delta Blocks		**1st Half (first 2 SG)**		**2nd Half (last 2 SG)**		**t-stat**	**df**	**pvalue**
		Mean	StDev	Mean	StDev			
Per Round	Cooperate	0.58	0.49	0.57	0.50	1.44	9997	0.15
	Optimal	0.60	0.49	0.56	0.50	4.68	9997	0.00****
(N=9999)	CaD	0.04	0.20	0.04	0.20	-0.16	9997	0.87
Per Supergame	Optimal (All-D+All-C)	0.63	0.48	0.55	0.50	4.02	2422	0.00****
	Optimal All-D	0.24	0.43	0.21	0.41	1.81	2422	0.07*
	Optimal All-C	0.40	0.49	0.35	0.48	2.53	2422	0.01**
(N=2424)	Suboptimal All-D	0.09	0.29	0.11	0.32	-1.67	2422	0.10*
	Suboptimal All-C	0.08	0.27	0.11	0.31	-1.96	2422	0.05**
	CaD	0.09	0.29	0.10	0.30	-0.77	2422	0.44
	DaC (End-Time)	0.10	0.30	0.13	0.34	-2.09	2422	0.04**
	Point Total	268.20	233.90	267.50	243.70	0.08	2422	0.94

Table 5: High δs: Top panel: Means, standard deviations, and t-tests comparing the first (12) and second (12) supergames of each session. Bottom panel: Within-δ-block comparisons (4 supergames per block), first vs. second half (2 supergames each). (Significance * 0.10 ** 0.05 *** 0.01 ***† 0.001.) (N=101 subjects.)

line treatment and are reported in Appendix C. As before, subjects fall into six types with no dominant pattern: only three of 101 always chose the theoretically optimal action, and higher CRT scores align with more payoff-maximizing behavior. Overall, behavior is consistent with our simple inattention model: cooperative choices by lower-ability (more attentive) subjects correlate with patience, while those by higher-ability (less attentive) subjects depend more on game structure than individual traits.

7 Conclusion

We report an experiment testing fundamental aspects of the standard game-theoretic model of repeated interactions, using both student and AMT subjects. In the repeated prisoner's dilemma, subjects played against a robot programmed with the Grim trigger strategy, converting the game into a single-person decision problem with a unique optimal strategy and eliminating confounds such as strategic uncertainty, social preferences, and multiple equilibria. Subjects completed many supergames with varying continuation probabilities, allowing classification of within-supergame play into distinct behavioral patterns and linking deviations from the theoretical optimum to individual characteristics, especially cognitive ability.

Only 1–2% of subjects behaved fully consistently with rational choice predictions, and just 3–5% did so more than 95% of the time. These low frequencies suggest that strict rational choice models may have limited empirical relevance for cooperative behavior. Behavioral models that incorporate errors or stochastic choice, such as logit or quantal response equilibrium, appear more suitable. In particular, the inattention model introduced here accounts for systematic biases, connects them to cognitive ability, and accommodates heterogeneity across individuals.

Our simplified individual-choice design pinpoints subjects' mistakes and shows how individual characteristics explain them. A majority (52–54%) made at least one strategic error of cooperating after defection, and some followed an end-timing strategy—defecting after initially cooperating (DaC)—which can yield higher payoffs than the theoretical optimum. Such behavior extends beyond standard theory and highlights the value of experimental evaluation. These patterns correlate with cognitive ability, and differences between high- and low-ability subjects align with a simple model of inattention.

We hope our findings help refine theoretical and empirical work on repeated strategic interaction and clarify the boundary between deliberate strategies and errors in such environments.

References

Ackerman, R. (2014): "The diminishing criterion model for metacognitive regulation of time investment." *Journal of Experimental Psychology: General*, 143, 1349–1368.

Agranov, M. and P. Ortoleva (2017): "Stochastic choice and preferences for randomization," *Journal of Political Economy*, 125, 40–68.

Andersen, S., G. W. Harrison, M. I. Lau, and E. E. Rutström (2008): "Eliciting risk and time preferences," *Econometrica*, 76, 583–618.

Andreoni, J. and J. H. Miller (1993): "Rational cooperation in the finitely repeated prisoner's dilemma: Experimental evidence," *The Economic Journal*, 103, 570–585.

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

Arechar, A. A., S. Gächter, and L. Molleman (2018): "Conducting interactive experiments online," *Experimental Economics*, 21, 99–131.

Bao, T., E. Nekrasova, T. Neugebauer, and Y. E. Riyanto (2022): "Algorithmic Trading in Experimental Markets with Human Traders: A Literature Survey," *Handbook of Experimental Finance*, 302–321.

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

Blonski, M., P. Ockenfels, and G. Spagnolo (2011): "Equilibrium selection in the repeated prisoner's dilemma: Axiomatic approach and experimental evidence," *American Economic Journal: Microeconomics*, 3, 164–92.

Blonski, M., A. von Schenk, V. Klockmann, S. Eibelshäuser, and H. Hegemann (2025): "Human and Algorithmic Cooperation: Parameters Matter!" *Available at SSRN 4955708*.

Camerer, C. F. (2003): *Behavioral Game Theory: Experiments in Strategic Interaction*, Princeton University Press.

Carbone, E. and J. Duffy (2014): "Lifecycle consumption plans, social learning and external habits: Experimental evidence," *Journal of Economic Behavior and Organization*, 106, 413–427.

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 (2023): "Using team discussions to understand behavior in indefinitely repeated prisoner's dilemma games," *American Economic Journal: Microeconomics*, 15, 114–145.

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.

DAVIS, D., A. IVANOV, AND O. KORENOK (2016): "Individual characteristics and behavior in repeated games: An experimental study," *Experimental Economics*, 19, 67–99.

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 and Organization*, 126, 55–74.

ENKE, B. AND T. GRAEBER (2023): "Cognitive uncertainty," *Quarterly Journal of Economics*, 138, 2021–2067.

FALK, A., A. BECKER, T. DOHMEN, B. ENKE, D. HUFFMAN, AND U. SUNDE (2018): "Global Evidence on Economic Preferences," *Quarterly Journal of Economics*, 133, 1645–1692.

FEHR, E. AND A. LEIBBRANDT (2011): "A field study on cooperativeness and impatience in the tragedy of the commons," *Journal of Public Economics*, 95, 1144–1155.

GABAIX, X. (2019): "Behavioral inattention," in *Handbook of Behavioral Economics: Applications and Foundations 1*, Elsevier, vol. 2, 261–343.

GILL, D. AND Y. ROSOKHA (2022): "Beliefs, learning, and personality in the indefinitely repeated prisoner's dilemma," *working paper*.

HOLT, C. A. AND S. K. LAURY (2002): "Risk aversion and incentive effects," *American economic review*, 92, 1644–1655.

HOUSER, D. AND R. KURZBAN (2002): "Revisiting kindness and confusion in public goods experiments," *American Economic Review*, 92, 1062–1069.

Johnson, E. J., C. Camerer, S. Sen, and T. Rymon (2002): "Detecting failures of backward induction: Monitoring information search in sequential bargaining," *Journal of Economic Theory*, 104, 16–47.

Kloosterman, A. (2020): "Cooperation in stochastic games: a prisoner's dilemma experiment," *Experimental Economics*, 23, 447–467.

Kölle, F., S. Quercia, and E. Tripodi (2020): "Social preferences under the shadow of the future," *SSRN Working Paper 3622125*.

Mailath, G. J. and L. Samuelson (2006): *Repeated Games and Reputations: Long-run Relationships*, Oxford University Press.

March, C. (2021): "Strategic interactions between humans and artificial intelligence: Lessons from experiments with computer players," *Journal of Economic Psychology*, 87, 102426.

Matějka, F. and A. McKay (2015): "Rational inattention to discrete choices: A new foundation for the multinomial logit model," *American Economic Review*, 105, 272–98.

Mengel, F., S. Weidenholzer, and L. Orlandi (2022): "Match Length Realization and Cooperation in Indefinitely Repeated Games," *Journal of Economic Theory*, 200, 105416.

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.

Normann, H.-T. and M. Sternberg (2023): "Human-Algorithm Interaction: Algorithmic Pricing in Hybrid Laboratory Markets," *European Economic Review*, 152, 104347.

Noussair, C. and M. A. Olson (1997): "Dynamic decisions in a laboratory setting," *Southern Economic Journal*, 978–992.

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

——— (2022): "Intelligence, Errors and Cooperation in Repeated Interactions," *Review of Economic Studies*, 89, 2723–2767.

Rand, D. G. (2012): "The promise of Mechanical Turk: How online labor markets can help theorists run behavioral experiments," *Journal of Theoretical Biology*, 299, 172–179.

Reverberi, C., D. Pischedda, M. Mantovani, J.-D. Haynes, and A. Rustichini (2021): "Strategic Complexity And Cognitive Skills Affect Brain Response In Interactive Decision-Making," *Scientific Reports*, 12, 1–18.

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.

Roth, A. E. and A. Ockenfels (2002): "Last-Minute Bidding and the Rules for Ending Second-Price Auctions: Evidence from eBay and Amazon Auctions on the Internet," *American Economic Review*, 92, 1093–1103.

Snowberg, E. and L. Yariv (2021): "Testing the waters: Behavior across participant pools," *American Economic Review*, 111, 687–719.

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

Appendices (For Online Publication Only)

A Appendix: Experimental Instructions and Questions

This Appendix contains the instructions, quiz questions, belief elicitation, and post-experimental questions in the order in which subjects encountered these in our original, baseline study.

A.1 Repeated PD Game Instructions and Comprehension Quiz

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

A.2 Belief Elicitation

After a subject had successfully completed all quiz questions, they were asked to provide their belief as to the proportion of times they would choose action X (the cooperative action) in each of the first rounds of the 24 sequences (supergames) that they would play. Prior to making this choice they were told that they would play 4 supergames for each of the 6 different delta values. After submitting their belief regarding their overall play of the cooperative action, the experiment proceeded on to the first indefinitely repeated PD game.

A.3 Repeated PD Games: Screenshots

For each indefinitely repeated PD game (referred to as a "sequence") subjects were clearly instructed about the continuation probability for that repeated game. E.g., the screenshots shown in Figures A1-A2 provide an illustration of the screens that subjects faced in the first

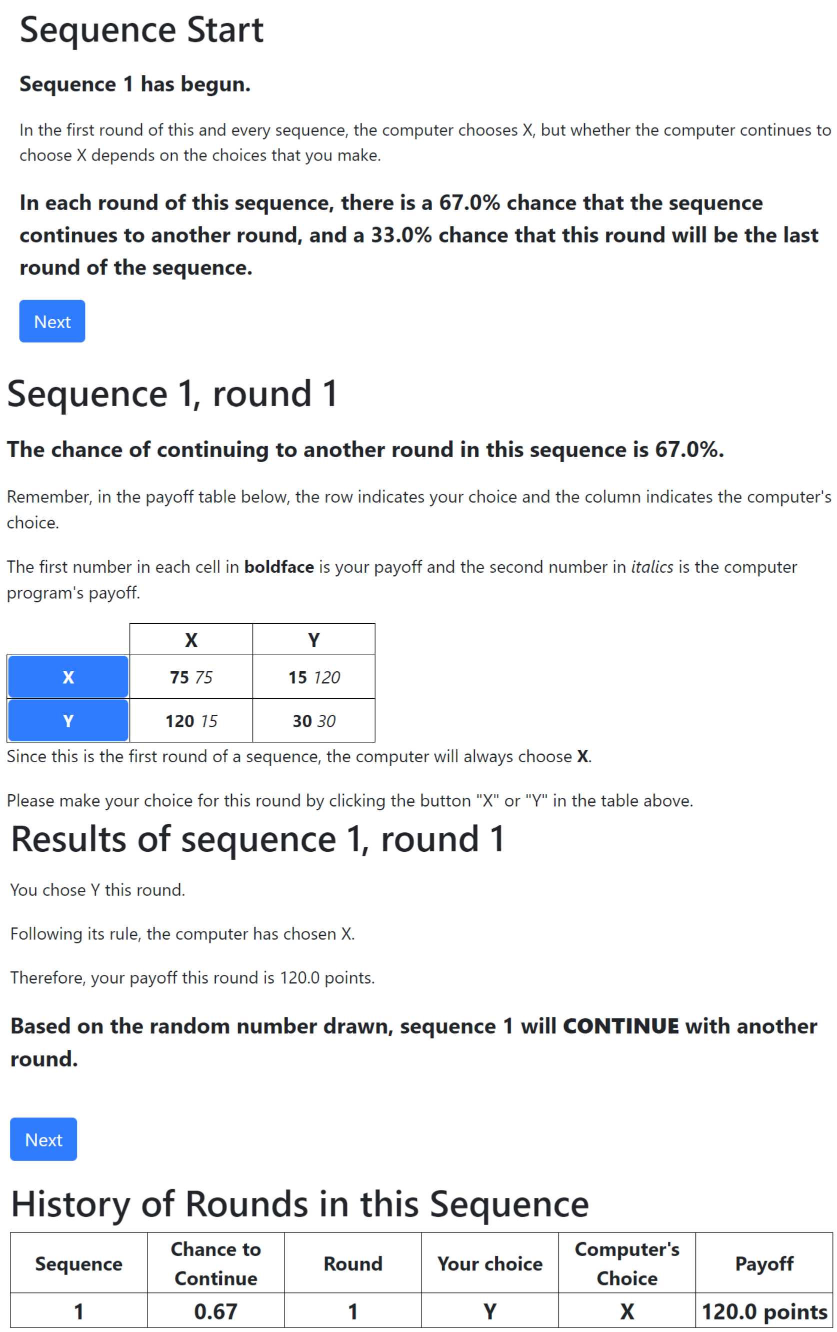

Figure A1: (Top) Start screen for a new sequence. (Middle) Main decision screen for a period in the sequence. (Bottom) Results screen for a period in the sequence.

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

Results of sequence 1, round 4

You chose Y this round.

Following its rule, based on your choices in previous rounds, the computer has chosen Y.

Therefore, your payoff this round is 30.0 points.

Based on the random number drawn, sequence 1 has ENDED.

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
1	0.67	2	Y	Y	30.0 points
1	0.67	3	Y	Y	30.0 points
1	0.67	4	Y	Y	30.0 points

Figure A2: (Top) Decision screen for a continuation period in the sequence, noting what the robot player will do, based on the history of play. (Bottom) Screen for the final period of a sequence noting that based on the random drawn, the sequence has ended.

round (Figure A1) and in continuation rounds (Figure A2) of the first supergame of the "orderlong" treatment, which had a continuation probability of 0.67 and lasted for 4 rounds.

In this illustration, the subject chooses Y (defect) in all 4 rounds and the computer program responds accordingly. Note that subjects were always informed in advance about the computer opponent's decision for each round based on the round number, the history of play and the prescriptions of the Grim trigger strategy. For instance, in round 1 (Figure A1) the subject is instructed: "Since this is the first round of a sequence the computer will always choose X." After the subject chose Y in the first round of Sequence 1, in the second round of the sequence (Figure A2) the subject is instructed: "Based on your choices in previous rounds of this sequence the computer will choose Y".

A.4 Personality Questions

After the main task, subjects were asked to complete the following "questionnaire" by clicking on radio buttons from 0,1,2,..10 to report their answers to each question.[40]

Questionnaire

We now ask for your willingness to act in a certain way in 2 different areas. Please indicate your answer on a scale from 0 to 10, where 0 means you are "completely unwilling to do so" and a 10 means you are "very willing to do so". You can also use any numbers between 0 and 10 to indicate where you fall on the scale, like 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10.

1. In general, how willing are you to take risks?
2. How willing are you to give up something that is beneficial for you today in order to benefit more from that in the future?
3. How willing are you to punish someone who treats you unfairly, even if there may be costs for you?
4. How willing are you to give to good causes without expecting anything in return?

How well do the following statements describe you as a person? Please indicate your answer on a scale from 0 to 10. A 0 means "does not describe me at all" and a 10 means "describes me perfectly". You can also use any numbers between 0 and 10 to indicate where you fall on the scale, like 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10.

5. When someone does me a favor I am willing to return it.
6. If I am treated very unjustly, I will take revenge at the first occasion, even if there is a cost to do so.

[40] Taken from Falk et al. (2018).

7. I assume that people have only the best intentions.

A.5 CRT questions

Subjects were asked to provide numerical answers to the following cognitive reflection test (CRT) questions.[41]

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?

A.6 Continuation Probabilities and Realizations

Tables A1 and A2 report on the continuation probabilities δ for each of the 24 sequences along with the actual number of rounds played for the two treatment orders.

[41] Based on Toplak et al. (2014) and Ackerman (2014).

	OrderShort		OrderLong	
Sequence	δ	No.Rounds	δ	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 A1: Continuation probabilities δ and the number of rounds played for each of the 24 sequences, both treatment orders (one order is just the reverse of the other).

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 A2: The distribution of the supergames, split by continuation probabilities δ. That is, out of 24 supergames, 11 lasted only 11 rounds, 7 only two rounds, and so on. The average theoretical and realized supergame durations are 1.99 rounds and 2 rounds, respectively.

B Further Results Baseline Treatment

B.1 Evolution of first-round Behavior

Analogous to Figure 1, which reports average cooperation and defection rates across all rounds of supergames 1–4 for each δ, Figure B1 reports the corresponding rates for the *first round only*. Specifically, it shows first-round cooperation and defection rates for each of the four supergames at each of the six baseline-treatment δ values. The results are similar to

Figure 1: again, cooperation is not driven to zero in the four games with $\delta < \delta^* = 0.5$, and is less than full for the two δ values above δ^*. There is again not much evidence of learning after the second supergame, and AMT participants are farther from the optimal policy than Lab participants.

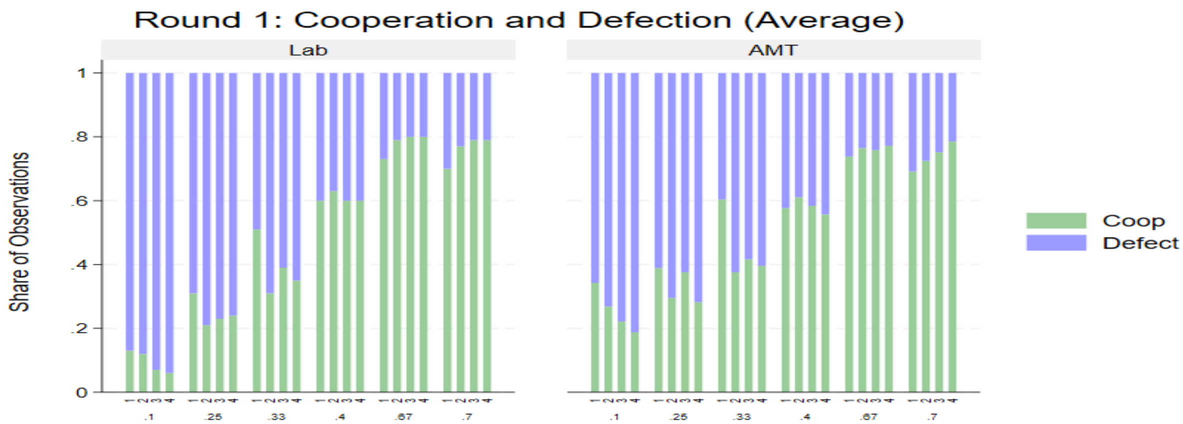

Figure B1: Cooperation and Defection rates in first rounds of the four supergames (first horizontal axis) of each of the six δ values (second horizontal axis) Left: Lab (N=100), Right: AMT (N=149).

B.2 Cost of first-round Mistakes

For each supergame, we quantify the ex-ante expected payoff loss from the observed round-1 action relative to the expected-payoff maximizing benchmark for the given continuation probability δ. Let $a_1 \in \{C, D\}$ denote the subject's actual round-1 choice and let $a_1^*(\delta)$ denote the benchmark round-1 choice. We focus on round 1 because it is observed for every subject in every supergame. To isolate the marginal payoff consequence of an incorrect round-1 classification, we evaluate a one-shot deviation loss: we switch only the round-1 action and then assume optimal continuation thereafter, given the grim-trigger state induced by the round-1 outcome.

Define the ex-ante (EA) loss as

$$\mathcal{L}^{EA}(\delta) \equiv \mathbb{E}[\Pi(a_1^*(\delta)) - \Pi(a_1) \,|\, \delta] \,,$$

where the expectation integrates over the stochastic termination process. The loss is zero when $a_1 = a_1^*(\delta)$ and increases with the foregone continuation value.

Because the benchmark action switches at $\delta^* = 0.5$, the relevant mistake differs across regions. If $\delta > \delta^*$, the benchmark prescribes cooperation; a round-1 defection triggers permanent punishment under grim trigger, implying

$$\mathcal{L}^{EA}(\delta) = \frac{75}{1-\delta} - \left(120 + \frac{30\delta}{1-\delta}\right) = 45 \cdot \frac{2\delta - 1}{1-\delta} = 90 \cdot \frac{\delta - \delta^*}{1-\delta}, \qquad \delta > \delta^*.$$

If $\delta < \delta^*$, the benchmark prescribes defection; a round-1 cooperation choice delays defection by one period and the optimal recovery is to defect from round 2 onward, yielding

$$\mathcal{L}^{EA}(\delta) = \left(120 + \frac{30\delta}{1-\delta}\right) - \left(75 + 120\delta + \frac{30\delta^2}{1-\delta}\right) = 45(1-2\delta) = 90(\delta^* - \delta), \qquad \delta < \delta^*.$$

At $\delta = \delta^* = 0.5$, the two actions are payoff-equivalent and $\mathcal{L}^{EA}(\delta^*) = 0$.

Table B1 reports the implied losses (in points) for the δ values used in the baseline study and in the high-δ treatment discussed in section 6. The loss is mechanically small near the cutoff where $\delta = \delta^*$ and is asymmetric across regions: when cooperation is optimal ($\delta > \delta^*$), a round-1 defection irreversibly shifts play to the punishment path, whereas when defection is optimal ($\delta < \delta^*$), a round-1 cooperation choice only delays implementation of the defection path by one period, if players understand their mistake and correct it.

Table B1: Ex-ante loss from a round 1 mistake (in points)

	δ								
Treatment	0.10	0.25	0.33	0.40	0.67	0.70	0.75	0.80	0.85
Baseline treatment	36.0	22.5	15.3	9.0	46.36	60.0			
High-δ treatment	36.0			9.0	46.36		90.0	135.0	210.0

B.3 Order Effects

As Mengel et al. (2022) documented, early exposure to relatively long sequences could affect subsequent behavior in the prisoner's dilemma, potentially leading to an order effect.

For the Lab subject pool, while the mean (st. dev.) first-round per-subject counts of cooperation in the reverse and long 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 $p = 0.278$). The corresponding mean (st.dev.) overall counts are, respectively, 25.52 (9.29) and 22.66 (12.83) (out of 48), with the difference still insignificant ($t = 1.28$, Kolmogorov-Smirnov one-sided $p = 0.198$). 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 only significant based on the Kolmogorov-Smirnov test (one-sided $p = 0.034$), and only marginally based on t-test ($t = 1.65, p = 0.051$). 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 based on t-test ($t = 1.87, p = 0.032$), but not based on Kolmogorov-Smirnov test (one-sided $p = 0.135$).

For the AMT subject pool, while the mean (st. dev.) first-round per-subject counts of cooperation in the reverse and long orders are 11.96 (5.96) and 12.99 (6.16), respectively (out of 24), this difference is insignificant ($t = 1.03$, Kolmogorov-Smirnov one-sided $p = 0.217$). The corresponding mean (st.dev.) overall counts are, respectively, 25.96 (11.21) and 27.38

(11.97) (out of 48), with the difference still insignificant ($t = 0.75$, Kolmogorov-Smirnov one-sided $p = 0.410$). As for the optimal choices, the first-round mean (st. dev.) counts are, respectively, 16.45 (4.45) and 14.53 (4.40), thus – in contrast to the Lab – significantly lower in the long treatment ($t = 2.68, p = 0.0042$, Kolmogorov-Smirnov one-sided $p = 0.011$). The overall optimal choice counts are, again, lower in the long order treatment (with mean (st.dev.) of 27.12 (9.03) in long order, and 30.71 (9.21) in reverse), and this is significant ($t = 2.40, p = 0.0089$, Kolmogorov-Smirnov one-sided $p = 0.038$.

Importantly, for both subject pools, once one controls for subjects' individual differences, the order effect is not discernible in mixed effects panel regressions in Table 3.

Finding 17. *There is no consistent order effect in either subject sample.*

B.4 More on Cooperation and Optimality

In Section 4.1 we showed the distributions of subjects' optimal choices at the very beginning of each supergame and overall, respectively. We now look at these choices in more details.

We start with the decisions in Round 1 of each supergame. For each subject pool, Figure B2 presents the frequency distributions of choices to cooperate (bottom left panels) and of optimal choices (top panels). The bottom right panels further 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. As the histograms show, subjects in both pools tend to excessively cooperate in the first round of each supergame, far above the theoretical prediction of 8. The mean (st. dev.) count of cooperative choices is 11.53 (5.73) for Lab and 12.47 (6.06) for AMT (with no significant difference across the two pools, see Table B2). As a result, the mean (st.dev.) count of theoretically optimal choices per subject is 16.81 (3.74) for Lab, which is significantly higher than 15.50 (4.52) for AMT. Both pools are prominently short of the theoretical prediction of 24.

Turning to the overall choice counts, Figure B3 presents the two-dimensional distributions of cooperative and optimal choices and the corresponding marginal distributions for all 48 choices in all 24 supergames. The mean (st. dev.) of the overall count of cooperative choices is only 24.09 (11.24) for Lab, which is significantly lower than the theoretical prediction of 26 (one-sided $t = 1.670, p = 0.046$). Interestingly, the AMT subjects' choice to cooperate are 26.66 (11.58), and thus on average are not significantly different from the predicted value of 26 (two-sided $t = 0.701, p = 0.4847$). However, for both subject pools, cooperative choices are often sub-optimal – as depicted by the two-dimensional distributions in the bottom right panels for each subject pool in Figure B3. (The shapes of the polygons for the overall choices in the bottom right panels are due to the possibility of dominated CaD choices, described in Section 4.3.) Indeed, for both subject pools, the overall optimal choice counts are significantly short of the theoretical prediction of 48, with a mean (st. dev.) of 31.05 (8.06) for Lab, which is marginally greater than 28.93 (9.27) for AMT (see Figure B2).

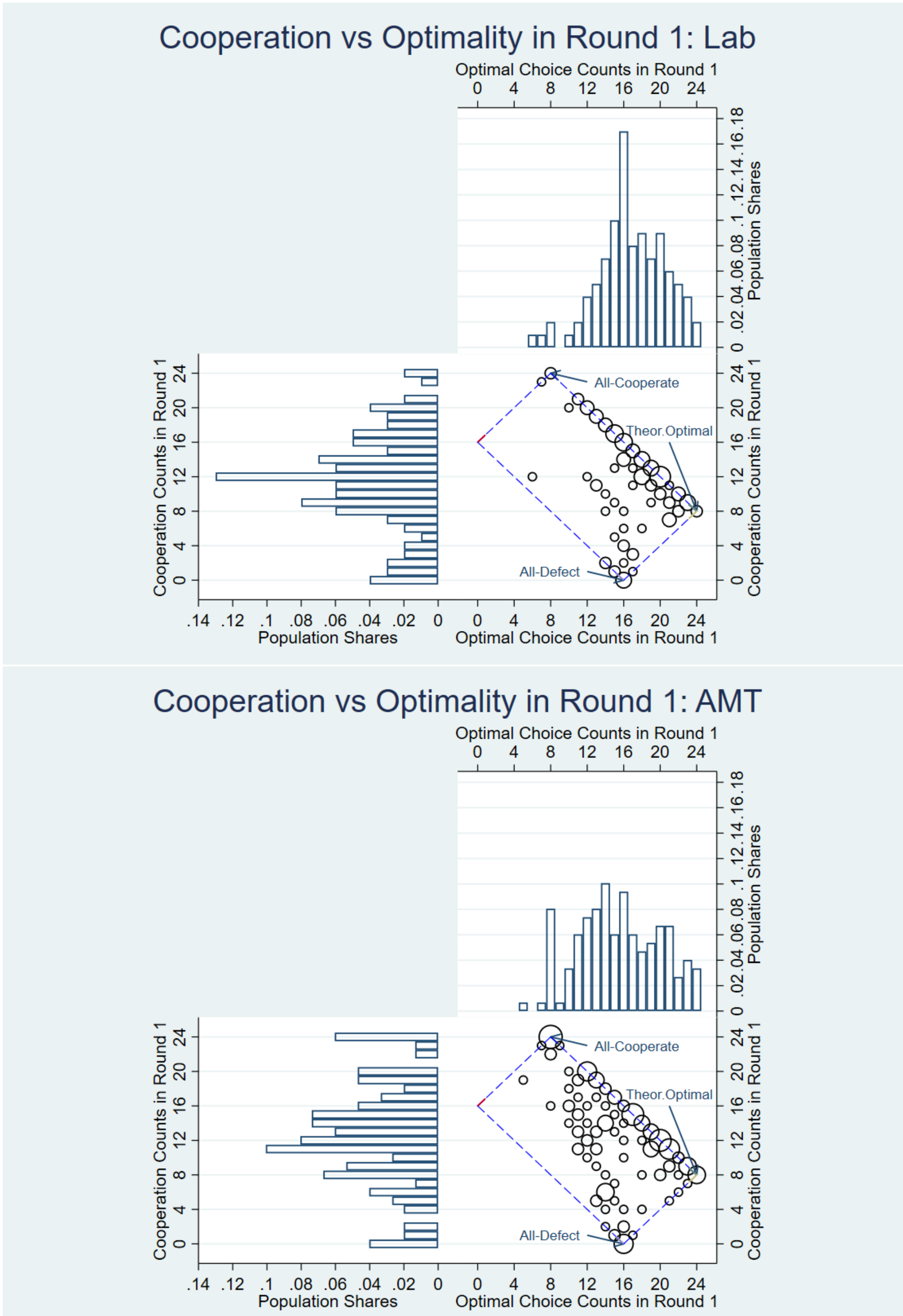

Figure B2: Choices in the first rounds each supergame: Lab (N=100) vs. AMT (N=149) subjects. Two-dimensional distributions of per-subject counts of cooperation and of optimal choices across all 24 supergames, together for the frequency distributions of cooperative (bottom left) and optimal (top right) choices. Bubble size is proportional to the share of subjects.

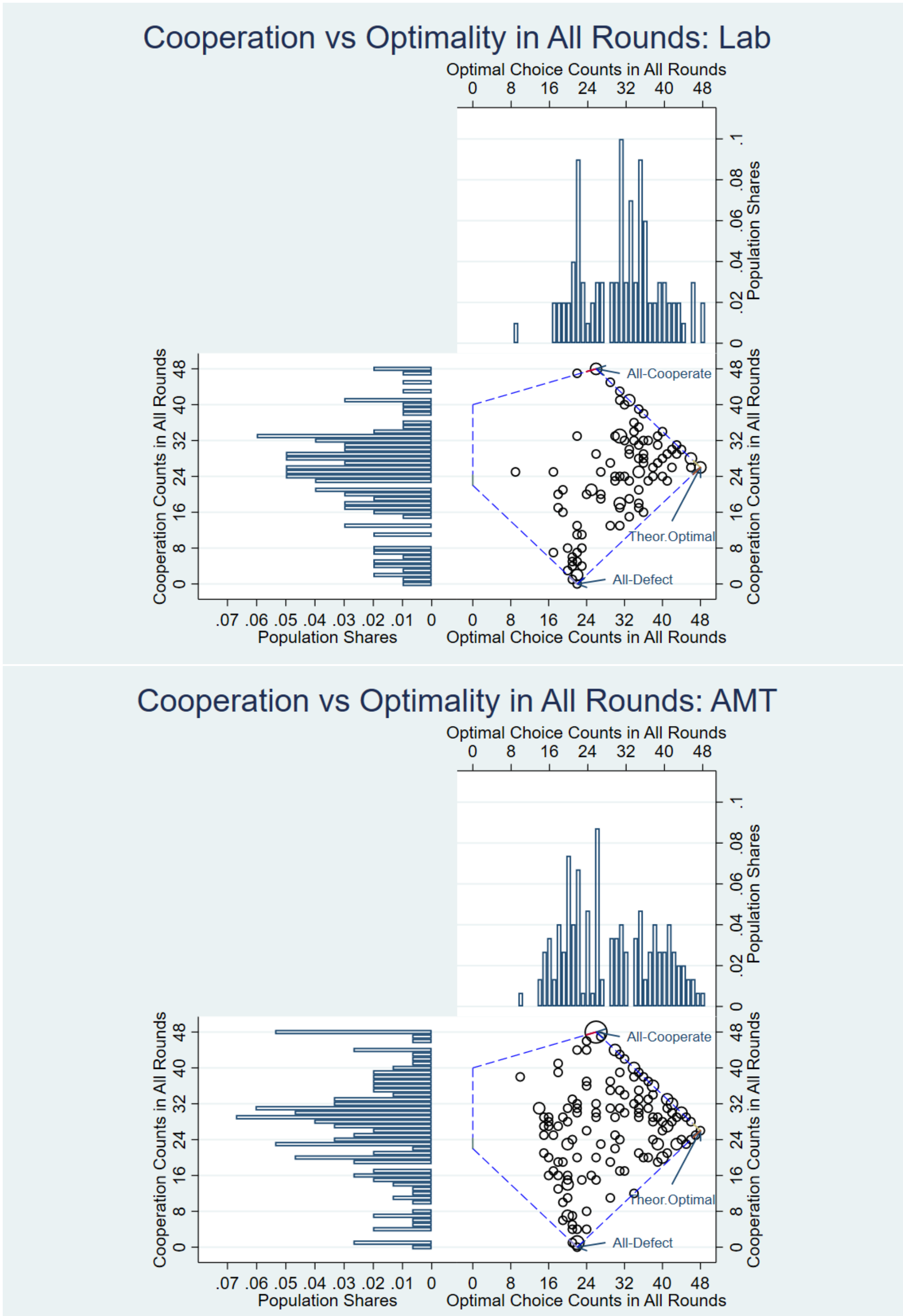

Figure B3: Choices in all rounds: Lab (N=100) vs. AMT (N=149) subjects. Two-dimensional distributions of per-subject counts of cooperation and of optimal choices across all 24 supergames, together for the frequency distributions of cooperative (bottom left) and optimal (top right) choices. Bubble size is proportional to the share of subjects.

In other words, early excessive cooperation in the first rounds is followed by the subsequent defection. 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) for Lab and 14.19 (6.58) for AMT, both significantly lower than the theoretical prediction of 18 ($p = 0.000$). Note that this is despite the excessive cooperation of 1.82 counts per Lab subject and 2.75 counts per AMT subject on average due to strategic (CaD) errors.

Finding 18. *For both subject pools, compared with the 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* $\delta > 0.5$*, with only 64.69% for Lab and 60.27% for AMT of all choices being theoretically optimal.*

As Figures B2 and B3 show, there is a significant heterogeneity in subjects' behavior (particularly for AMT), without any clear "representative" pattern. The initial heterogeneity of play in the first rounds in Figure B2 (bottom right panels for each subject pool) is further amplified by the heterogeneous strategies employed by the subjects in the subsequent rounds, depicted in the corresponding panels in Figure B3.

As the bottom right bottom panels of Figure B3 for each subject pool show, there are three similarly sized clusters (about 3-7% of each subject pool) at each of the three corners of the polygon. Only two (out of 100) Lab and one (out of 149) AMT subjects made perfect theoretically optimal choices, in the far right corners of the corresponding polygons. Further only three and five such subjects, respectively, made up to three suboptimal choices.

In the top corner, two Lab subjects and eight AMT subjects always cooperated, and two further subjects in each pool defected up to three times. In the bottom corner, a single subject in each pool always defected, and further four subjects in each pool cooperated up to three times. The presence of strategic CaD errors complicates the interpretation of the remaining subjects, most of whom are located away from the boundaries, in the center of the figures. Many of those observations represent the overall early excessive cooperation in the first rounds followed by the subsequent defection within a supergame, possibly due to some form of previously under-reported "end-timing" strategies (see Section 4.5).

Finding 19. *In both subject pools, perfect and near-perfect theoretically optimal behavior is rare, with only 5% of the Lab and 4.69% of the AMT subjects making no more than 3 theoretically sub-optimal choices. These shares are of similar order of magnitude as the shares of subjects who defected no more than 3 times in both pools (4% Lab and 6.71% AMT), and who cooperated no more than 3 times in both pools (5% Lab and 3.36% AMT).*

B.5 Learning

In Figure B4, one can observe an increase in the "end-timing" activities in the both subject pools by comparing those in the first few supergames to that in the last few (further supported by the tests in Table 1). As this same Figure reveals, while the incidence of dominated CaD errors decline over time, they do not disappear entirely. Figure B5 further presents the patterns of intra-supergame play across all 24 sequences, split by the sequence order.

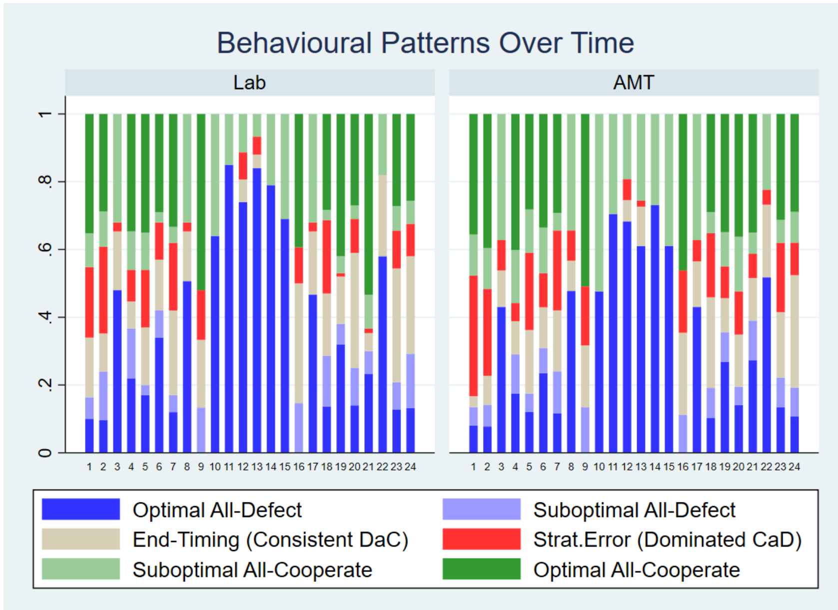

Figure B4: Subject behavior over the sequence of 24 supergames, by subject pool. Both supergame sequence orders are pooled together so a supergame with a given number could involve different continuation probabilities δ and corresponding optimal actions, depending on the sequence order.

B.6 Comparison of Lab and AMT Subjects

The regression results of Table 3 (and the corresponding odds in Table B4) are summarized visually in Figure 10, showing that the two subject groups differ in their patterns of play. Subjects in the lower CRT7 group (two top panels of Figure 10) make relatively more frequent strategic errors (CaD) and engage in suboptimal consistent defecting behavior (Suboptimal All-D). By contrast, subjects in the higher CRT7 group (bottom two panels) are closer to the theoretically optimal policy and engage in end-timing behavior (DaC) more often.

Figure 10 further reveals differences in patterns of behavior between the two subject pools (see also Table B2). The top two panels showing behavior by subjects with CRT7 scores less than or equal to the median score are consistent with the earlier insights of Arechar et al. (2018) and Snowberg and Yariv (2021) that the Lab subjects are less prone to pro-social behaviour and less likely to make mistakes as compared with the AMT subjects. However, the bottom two panels, showing behavior by subjects whose CRT7 scores are strictly above the median suggest that there is hardly any difference in patterns of behavior across the two subject pools, despite apparent individual differences in non-cognitive characteristics. Indeed, this visual observation is further confirmed by formal t-tests in Table B3.

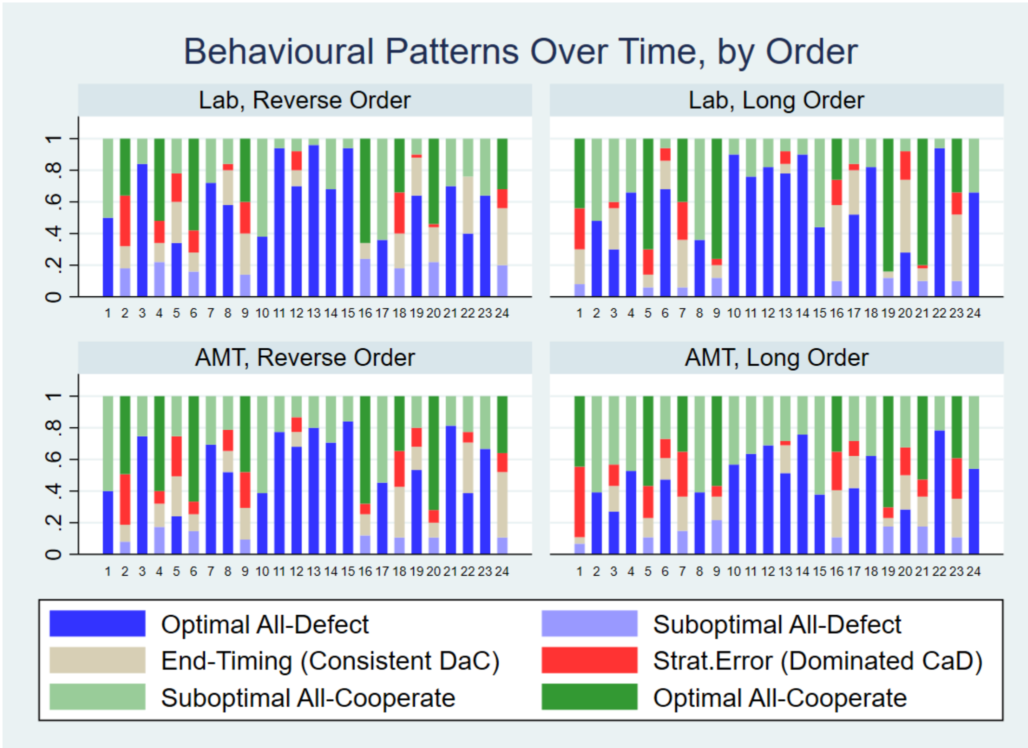

Figure B5: Subject behavior over the sequence of 24 supergames, by sequence order and subject pool.

Finding 20. *The behavior of subjects with relatively high cognitive costs (CRT7 scores weakly below the median) differs markedly for Lab and AMT subject pools. In contrast, there is little difference in the behaviour of subjects with relatively low cognitive costs (CRT7 scores above the median) across the two pools, despite differences in their non-cognitive characteristics.*

Figure B6 presents the distributions of the two key variables for the inattention model, $CRT7$ and Prediction variable. As Table B2 shows, the two subject pools do not differ significantly in the means of these two variables.

Lab vs. AMT	Lab (N=100)		AMT (N=149)		t-stat	df	pvalue
	Mean	StDev	Mean	StDev			
Female	0.52	0.50	0.50	0.50	0.31	247	0.76
Age*	21.49	2.46	39.74	10.49	-20.42	171.45	0.00***†
CRT7	3.78	2.26	3.58	2.17	0.71	247	0.48
Risk	6.28	1.87	5.15	2.69	3.64	247	0.00***†
Patience	7.72	1.97	7.62	1.84	0.42	247	0.67
Punishment	4.84	2.51	4.07	2.97	2.14	247	0.03**
Altruism	7.10	2.26	7.44	2.50	-1.08	247	0.28
Reciprocity	9.08	1.23	8.38	1.93	3.20	247	0.00***
Retribution	3.42	2.32	3.15	2.99	0.77	247	0.44
Trust	4.48	2.22	5.73	2.58	-3.97	247	0.00***†
Prediction	60.96	29.57	66.64	27.04	-1.56	247	0.12
Quiz Errors*	1.40	3.38	2.69	7.34	-1.86	222.91	0.06*
Points Total	3835.10	203.40	3766.20	240.60	2.35	247	0.02**
Round 1: Cooperate	11.53	5.73	12.47	6.06	-1.23	247	0.22
Round 1: Optimal	16.81	3.74	15.50	4.52	2.40	247	0.02**
Total: Cooperate	24.09	11.24	26.66	11.58	-1.74	247	0.08*
Total: Optimal	31.05	8.06	28.93	9.27	1.87	247	0.06*
Total: CaD	1.82	2.56	2.75	4.09	-2.03	247	0.04**
Supergames: Optimal	14.44	3.92	13.07	4.88	2.34	247	0.02**
Supergames: Optimal All-D	10.31	3.96	8.95	4.67	2.40	247	0.02**
Supergames: Optimal All-C	4.13	2.83	4.13	2.96	0.01	247	0.99
Supergames: Suboptimal All-D	1.14	2.18	1.02	1.85	0.47	247	0.64
Supergames: Suboptimal All-C	4.09	3.59	5.52	4.58	-2.63	247	0.01*
Supergames: CaD	1.50	2.05	2.05	2.82	-1.66	247	0.10*
Supergames: DaC (End-Time)	2.83	2.37	2.34	2.26	1.66	247	0.10*

Table B2: For each subject pool: Means and standard deviations of key variables, and t-tests of differences between the means for two pools (all equal variance tests except for Age and Quiz Errors). *df* stands for degrees of freedom or Satterthwaite's degrees of freedom in case of unequal variances for Age and Quiz Errors, *p*value stands for $Pr(|T| > |t|) = 0$. (Significance * 0.10 ** 0.05 *** 0.01 ***† 0.001.)

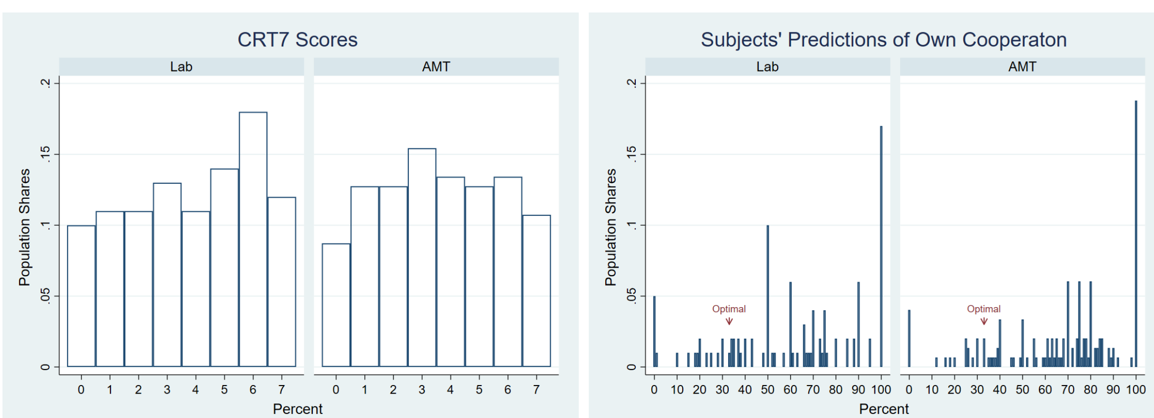

Figure B6: Frequency distributions of CRT7 scores (left panel) and of the "Prediction" variable (right panel).

Lab vs. AMT: $CRT7 > 4$	Lab (N=44)		AMT (N=55)		t-stat	df	pvalue
	Mean	StDev	Mean	StDev			
Female	0.41	0.50	0.46	0.50	-0.54	97	0.59
Age*	21.45	2.39	39.95	10.09	-13.14	61.45	0.00***†
CRT7	5.96	0.78	5.95	0.80	0.06	97	0.95
Risk	6.39	2.08	4.42	2.08	4.68	97	0.00***†
Patience	8.02	1.52	7.67	1.53	1.14	97	0.26
Punishment	5.11	2.54	3.71	2.39	2.83	97	0.01***
Altruism	6.55	2.54	7.42	2.23	-1.82	97	0.07*
Reciprocity	9.09	1.07	8.60	1.54	1.80	97	0.08*
Retribution	3.55	2.05	2.75	2.17	1.87	97	0.06*
Trust	4.16	1.88	5.42	2.28	-2.95	97	0.00***
Prediction	59.52	31.98	70.35	26.41	-1.84	97	0.07*
Quiz Errors	0.66	1.45	0.51	1.03	0.60	97	0.55
Points Total	3898.00	176.80	3917.50	183.80	-0.53	97	0.59
Round 1: Cooperate	12.32	5.26	13.02	5.34	-0.65	97	0.52
Round 1: Optimal	17.41	3.49	17.56	4.52	-0.19	97	0.85
Total: Cooperate	25.32	10.29	29.09	10.03	-1.84	97	0.07*
Total: Optimal	33.50	7.17	34.87	7.81	-0.90	97	0.37
Total: CaD	0.86	1.50	0.98	1.80	-0.35	97	0.73
Supergames: Optimal	15.16	3.95	15.51	4.55	-0.40	97	0.69
Supergames: Optimal All-D	10.43	3.88	9.98	4.54	0.52	97	0.60
Supergames: Optimal All-C	4.73	2.52	5.53	2.62	-1.54	97	0.13
Supergames: Suboptimal All-D	0.86	2.16	0.42	1.29	1.27	97	0.21
Supergames: Suboptimal All-C	4.05	3.29	4.84	4.52	-0.97	97	0.33
Supergames: CaD	0.71	1.23	0.82	1.44	-0.42	97	0.68
Supergames: DaC (End-Time)	3.23	2.61	2.42	2.37	1.62	97	0.11

Lab vs. AMT: $CRT7 \leq 4$	Lab (N=56)		AMT (N=94)		t-stat	df	pvalue
	Mean	StDev	Mean	StDev			
Female	0.61	0.49	0.52	0.50	1.02	148	0.31
Age*	21.52	2.53	39.63	10.77	-15.59	109.43	0.00***†
CRT7	2.07	1.40	2.19	1.35	-0.52	148	0.60
Risk	6.20	1.69	5.59	2.91	1.43	148	0.15
Patience	7.48	2.24	7.59	2.00	-0.29	148	0.77
Punishment	4.63	2.50	4.28	3.25	0.69	148	0.49
Altruism	7.54	1.94	7.45	2.65	0.22	148	0.83
Reciprocity	9.07	1.35	8.26	2.13	2.58	148	0.01**
Retribution	3.32	2.52	3.38	3.37	-0.12	148	0.91
Trust	4.73	2.44	5.92	2.74	-2.66	148	0.01
Prediction	62.09	27.77	64.47	27.30	-0.51	148	0.61
Quiz Errors*	1.98	4.25	3.96	8.98	-1.82	142.15	0.07*
Points Total	3785.60	210.70	3677.70	225.90	2.90	148	0.00***
Round 1: Cooperate	10.91	6.05	12.15	6.45	-1.16	148	0.25
Round 1: Optimal	16.34	3.89	14.30	4.07	3.02	148	0.00***
Total: Cooperate	23.13	11.93	25.24	12.22	-1.04	148	0.30
Total: Optimal	29.13	8.26	25.45	8.26	2.64	148	0.01***
Total: CaD	2.57	2.95	3.79	4.66	-1.75	148	0.08*
Supergames: Optimal	13.88	3.83	11.65	4.51	3.09	148	0.00***
Supergames: Optimal All-D	10.21	4.05	8.34	4.66	2.50	148	0.01**
Supergames: Optimal All-C	3.66	2.99	3.31	2.85	0.72	148	0.47
Supergames: Suboptimal All-D	1.36	2.19	1.37	2.04	-0.04	148	0.97
Supergames: Suboptimal All-C	4.13	3.83	5.93	4.59	-2.47	148	0.01**
Supergames: CaD	2.13	2.34	2.77	3.17	-1.31	148	0.19
Supergames: DaC (End-Time)	2.52	2.13	2.29	2.20	0.63	148	0.53

Table B3: For each subject pool: Means and standard deviations of key variables, and t-tests of differences between the means for two pools (all equal variance tests except for Age and Quiz Errors). *df* stands for degrees of freedom or Satterthwaite's degrees of freedom in case of unequal variances for Age and Quiz Errors, *p*value stands for $Pr(|T| > |t|) = 0$. (Significance * 0.10 ** 0.05 *** 0.01 ***† 0.001.)

Cooperate	All				$CRT7 \leq 4$		$CRT7 > 4$	
(Odds)	(1) Lab	(2) AMT	(3) Lab	(4) AMT	(5) Lab	(6) AMT	(7) Lab	(8) AMT
δ=0.25	0.90***†	0.31***†	0.89***†	0.31***†	0.92***†	0.28***	0.87***†	0.43**
	(0.14)	(0.08)	(0.13)	(0.08)	(0.17)	(0.10)	(0.23)	(0.18)
δ=0.33	1.23***†	0.68***†	1.22***†	0.68***†	1.17***†	0.50***†	1.36***†	1.21***†
	(0.15)	(0.10)	(0.15)	(0.10)	(0.20)	(0.11)	(0.25)	(0.20)
δ=0.4	1.88***†	1.04***†	1.87***†	1.04***†	1.67***†	0.73***†	2.22***†	1.86***†
	(0.17)	(0.12)	(0.17)	(0.12)	(0.20)	(0.13)	(0.33)	(0.24)
δ=0.67	2.65***†	1.56***†	2.65***†	1.56***†	2.30***†	1.00***†	3.29***†	3.03***†
	(0.21)	(0.15)	(0.21)	(0.15)	(0.25)	(0.16)	(0.40)	(0.29)
δ=0.7	2.76***†	1.70***†	2.76***†	1.70***†	2.32***†	1.15***†	3.53***†	3.15***†
	(0.22)	(0.15)	(0.22)	(0.15)	(0.26)	(0.16)	(0.44)	(0.33)
Round 2	-0.14	0.10	-0.14	0.09	-0.02	0.09	-0.43**	-0.10
	(0.10)	(0.08)	(0.10)	(0.08)	(0.12)	(0.10)	(0.20)	(0.15)
Round 3	-0.47***†	0.15	-0.47***†	0.15	-0.22	0.23*	-0.98***†	-0.29
	(0.14)	(0.10)	(0.14)	(0.10)	(0.17)	(0.12)	(0.26)	(0.18)
Round 4	-0.74***†	-0.06	-0.75***†	-0.07	-0.39*	0.14	-1.37***†	-0.83***†
	(0.17)	(0.12)	(0.17)	(0.12)	(0.21)	(0.13)	(0.30)	(0.23)
Round 5	-0.69***	-0.48***	-0.70***	-0.49***	-0.48*	-0.35*	-1.09***	-1.14***†
	(0.22)	(0.16)	(0.22)	(0.16)	(0.28)	(0.20)	(0.40)	(0.33)
Supergame	-0.26***	-0.23***	-0.26***	-0.23***	-0.33***	-0.15	-0.16	-0.47****
	(0.09)	(0.09)	(0.09)	(0.09)	(0.12)	(0.11)	(0.16)	(0.14)
Order Long	0.18	0.12	0.22	0.14	0.26	0.22	0.22	0.19
	(0.17)	(0.14)	(0.16)	(0.16)	(0.21)	(0.21)	(0.23)	(0.31)
Prior Defect	-0.81***†	-0.86***†	-0.80***†	-0.85***†	-0.64***†	-0.65***†	-1.13***†	-1.13***†
	(0.12)	(0.11)	(0.12)	(0.11)	(0.14)	(0.14)	(0.22)	(0.18)
CRT7			-0.00	0.03				
			(0.04)	(0.03)				
Prediction			0.97***	0.72***	0.70	0.27	1.12***	1.21**
			(0.36)	(0.24)	(0.54)	(0.29)	(0.37)	(0.48)
Female			-0.33*	-0.31**	-0.32	-0.39**	-0.35	-0.25
			(0.17)	(0.15)	(0.22)	(0.19)	(0.28)	(0.25)
Age			-0.03	-0.01	-0.03	-0.01	-0.05	0.01
			(0.02)	(0.01)	(0.04)	(0.01)	(0.04)	(0.02)
Risk			-0.01	-0.05	-0.03	-0.09*	0.02	-0.02
			(0.06)	(0.04)	(0.08)	(0.05)	(0.09)	(0.08)
Patience			0.07	0.03	0.14***	0.11*	-0.09	-0.04
			(0.04)	(0.05)	(0.04)	(0.07)	(0.09)	(0.10)
Punishment			0.01	-0.02	0.03	0.02	-0.03	-0.11*
			(0.05)	(0.04)	(0.07)	(0.04)	(0.06)	(0.06)
Altruism			-0.06	-0.02	-0.09	-0.04	-0.09	-0.01
			(0.05)	(0.03)	(0.06)	(0.04)	(0.08)	(0.07)
Reciprocity			0.07	-0.04	0.04	-0.06	0.26	0.02
			(0.07)	(0.03)	(0.08)	(0.05)	(0.18)	(0.06)
Retribution			-0.04	0.05	0.01	0.04	-0.10*	0.06
			(0.04)	(0.03)	(0.06)	(0.05)	(0.05)	(0.07)
Trust			0.01	0.02	-0.03	0.00	0.09	0.02
			(0.03)	(0.03)	(0.04)	(0.04)	(0.05)	(0.05)
Constant	-1.61***†	-0.73***†	-2.05***	-0.68	-1.67	0.11	-2.31	-1.85*
	(0.19)	(0.13)	(0.78)	(0.51)	(1.07)	(0.55)	(1.80)	(1.11)
chi2	266.09	203.19	406.22	231.92	227.73	110.77	200.85	218.72
p	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
N	4800	7152	4800	7152	2688	4512	2112	2640

Table B4: Choices to cooperate: mixed-effects probit regressions, odds, robust errors in parentheses. (See Table 3 for the corresponding marginals.) "Supergame" is the supergame number in the sequence of supergames (scaled down by 24), "Order Long" is a dummy variable for whether the first supergame in the sequence had $\delta = 0.67$, "Prior Defection" is a dummy variable for whether the subject defected in prior rounds of a given supergame, "Prediction" is the subjects' predictions of the share of their own cooperative choices in Round 1 across all 24 supergames (scaled down by 100). (Significance * 0.10 ** 0.05 *** 0.01 ***† 0.001.)

C High Delta Treatments Experimental Procedures and Further Results

C.1 Experimental Instructions and Questions

The Repeated PD Game Instructions and Comprehension Quiz are essentially the same as for the original design Appendix A. The only change to the instructions was in the following paragraph, which begins in the same way in the original instructions, but is modified to allow four sequences in a row of the same δ value: "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 *will change after every **four** sequences*, 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)."

C.2 Holt-Laury Risk Elicitation

The high delta treatments included an incentivized Holt–Laury risk-elicitation task (Holt and Laury, 2002) administered after the main PD game and before the personality and CRT questions. This task is explained in the following instructions given to subjects:

Additional money earning task

In this task you have the opportunity to earn an additional payment. You will choose between 2 different lotteries: lottery A or lottery B. You will choose between lottery A or lottery B ten different times. These ten choices are shown below in ten different rows.
The dollar amounts in lotteries A and B will always be the same. The probabilities in lotteries A and B will change from row to row. For both lotteries A and B, the probability of the bigger dollar amount starts at 1/10 (10%) and increases to 10/10 (100%). For both lotteries A and B, the probability of the smaller dollar amount starts at 9/10 (90%) and decreases to 0/10 (0%). You can choose either A or B for each lottery pair.

After you have made all ten choices, the computer program will randomly select one of your ten choices (one of the ten rows). All rows are equally likely to be chosen. Your lottery choice for the chosen row (A or B) will then be implemented. The computer program will choose a random integer uniformly from 1 to 100 inclusive and this number determines your payoff for your selected lottery. For example, lottery A in row 1 pays $2 if the randomly drawn integer is 1-10 and $1.60 if the randomly drawn integer is 11-100.

Note that your possible earnings from this lottery choice task are: $0.10, $1.60, $2.00 or

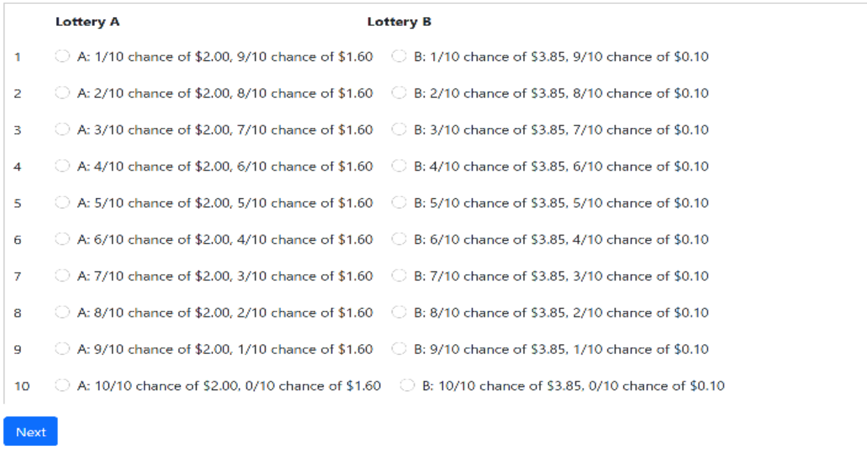

Figure C1: Screenshot of Holt Laury Lottery Choice Task

$3.85. These earnings will be added to your payoff from the main task.

C.3 Experimental Design: Continuation Probabilities and Realizations

Tables C1 and C2 report on the continuation probabilities δ for each of the 24 supergames (SG) of the high delta treatments along with the actual number of rounds played for the two treatment orders.

C.4 High Deltas: Order Effects

In the long order and reverse order treatments, respectively, there were 51 subjects (18 males and 33 females) and 50 subjects (18 males and 32 females); the mean (st.dev.) age was 20.53 (2.02) and 20.54 (2.30); the mean (st.dev.) CRT7 was 3.63 (2.38) and 3.60 (2.59).

The mean (st. dev.) first-round per-subject counts of cooperation (out of 24) is significantly higher in the long order 16.25 (5.50) vs. 13.30 (5.68) in the reverse one ($t = 2.66$, $p = 0.0046$, Kolmogorov-Smirnov one-sided $p = 0.001$). The corresponding mean (st.dev.) overall counts (out of 99) remain to be higher in the long order at 59.63 (22.57) than 53.88 (25.77) in the reverse order, but the difference is insignificant ($t = 1.19$, Kolmogorov-Smirnov one-sided $p = 0.271$). As for the optimal choices, the first-round counts are higher in the reverse treatment, with mean (st. dev.) being, respectively, 18.46 (4.79) and 17.16 (4.36), but the difference is insignificant ($t = 1.43$, Kolmogorov-Smirnov one-sided $p = 0.017$). Importantly, once one controls for subjects' individual differences, the order effect is not discernible in mixed effects panel regressions in Table C5.

Finding 21. *There is no consistent order effect in either subject sample.*

	OrderShort		OrderLong	
Supergame	δ	No.Rounds	δ	No. Rounds
1	0.4	1	0.67	3
2	0.4	1	0.67	2
3	0.4	1	0.67	1
4	0.4	4	0.67	3
5	0.75	5	0.85	6
6	0.75	10	0.85	1
7	0.75	1	0.85	15
8	0.75	6	0.85	10
9	0.8	1	0.1	1
10	0.8	4	0.1	1
11	0.8	15	0.1	1
12	0.8	5	0.1	1
13	0.1	1	0.8	5
14	0.1	1	0.8	15
15	0.1	1	0.8	4
16	0.1	1	0.8	1
17	0.85	10	0.75	6
18	0.85	15	0.75	1
19	0.85	1	0.75	10
20	0.85	6	0.75	5
21	0.67	3	0.4	4
22	0.67	1	0.4	1
23	0.67	2	0.4	1
24	0.67	3	0.4	1
Totals		99		99

Table C1: Continuation probabilities δ and the number of rounds played for each of the 24 supergames, both treatment orders (one order is the reverse of the other).

Delta	Number of Supergames (SGs) of Duration / Rounds																	Total
δ	Exp. $\left(\frac{1}{1-\delta}\right)$	Real. (Ave.)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	SGs
0.1	1.11	1.00	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4
0.4	1.67	1.75	3	0	0	1	0	0	0	0	0	0	0	0	0	0	0	4
0.67	3.03	2.25	1	1	2	0	0	0	0	0	0	0	0	0	0	0	0	4
0.75	4.00	5.50	1	0	0	0	1	1	0	0	0	1	0	0	0	0	0	4
0.8	5.00	6.25	1	0	0	1	1	0	0	0	0	0	0	0	0	0	1	4
0.85	6.67	8.00	1	0	0	0	0	1	0	0	0	1	0	0	0	0	1	4
Total SGs			11	1	2	2	2	2	0	0	0	2	0	0	0	0	2	24
Total Choices in SGs per Duration			11	2	6	8	10	12	0	0	0	20	0	0	0	0	30	99

Delta	Number of Choices / Round																Total
δ	Round →	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Choices
0.1		4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4
0.4		4	1	1	1	0	0	0	0	0	0	0	0	0	0	0	7
0.67		4	3	2	0	0	0	0	0	0	0	0	0	0	0	0	9
0.75		4	3	3	3	3	2	1	1	1	1	0	0	0	0	0	22
0.8		4	3	3	3	2	1	1	1	1	1	1	1	1	1	1	25
0.85		4	3	3	3	3	3	2	2	2	2	1	1	1	1	1	32
Total Choices		24	13	12	10	8	6	4	4	4	4	2	2	2	2	2	99

Table C2: The distribution of the supergames, split by continuation probabilities δ. That is, out of 24 supergames, 11 lasted only 1 round, 1 lasted only two rounds, and so on. The average theoretical and realized supergame durations are 3.58 and 4.13 rounds, respectively.

C.5 High Deltas: More on Cooperation and Optimality

In the high-delta treatment, Figure C2 furher presents the frequency distributions of cooperation (bottom left) and of optimal choices (top). The bottom right panel shows the two-dimensional distribution of cooperative and optimal choices, where the dashed lines delineate polygons restricting the possible choice combinations. As seen, subjects tend to cooperate insufficiently in the first rounds, with the mean (st.dev.) count of first-round cooperative choices, 14.79 (5.76), significantly below the theoretical prediction of 16 ($t = 2.11, p = 0.0187$). As a result, the mean (st.dev.) count of theoretically optimal first-round choices per subject is 17.80 (4.60), well short of the theoretical prediction of 24.

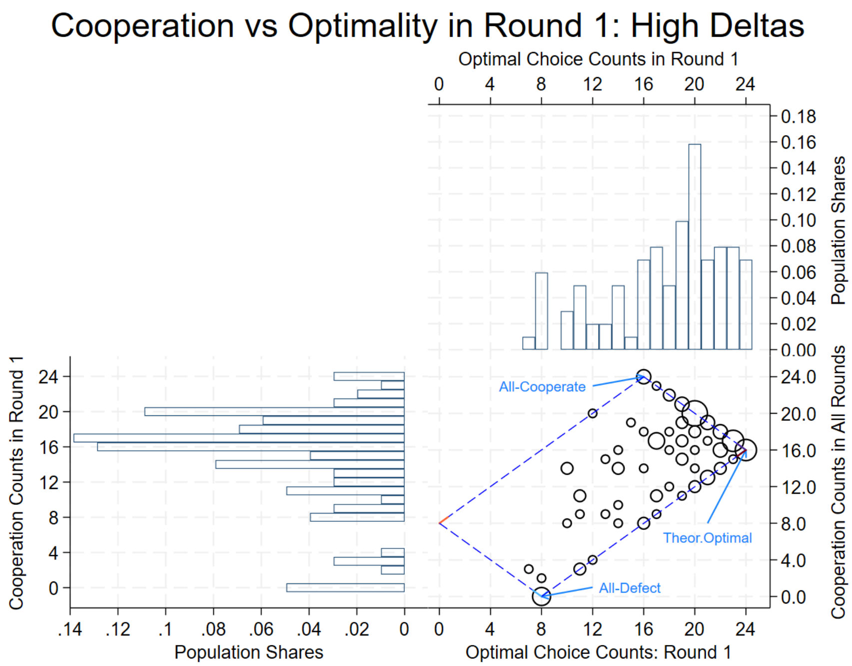

Figure C2: Choices in the first rounds each supergame: two-dimensional distributions of per-subject counts of cooperation and of optimal choices across all 24 supergames, and the frequency distributions of cooperative (bottom left) and optimal (top right) choices. Bubble size is proportional to the share of subjects (N=101).

Figure C3 further presents the two-dimensional distributions of cooperative and optimal choices and the corresponding marginal distributions for all 99 choices in all 24 supergames. The mean (st.dev.) of the overall count of cooperative choices is only 56.78 (24.26), prominently short of the theoretical prediction of 88. Moreover, cooperative choices are often sub-optimal – as depicted by the two-dimensional distribution in the bottom right panel in Figure C3 (where the shape of the polygon is due to the possibility of dominated CaD choices, see Section 6.4). Indeed, with a mean (st.dev.) of 57.22 (26.41), the overall optimal choice counts are prominently short of the theoretical prediction of 99.

In other words, subjects start by cooperating insufficiently in the first rounds, and thus can never "catch up" in the subsequent rounds. The suboptimal decisions are further aggravated by the excessive cooperation due to strategic (CaD) errors.

Finding 22. *Compared with the theoretical predictions, on average, subjects cooperate too little at the beginning of the supergames with $\delta > 0.5$, with only 57.80% of all choices being*

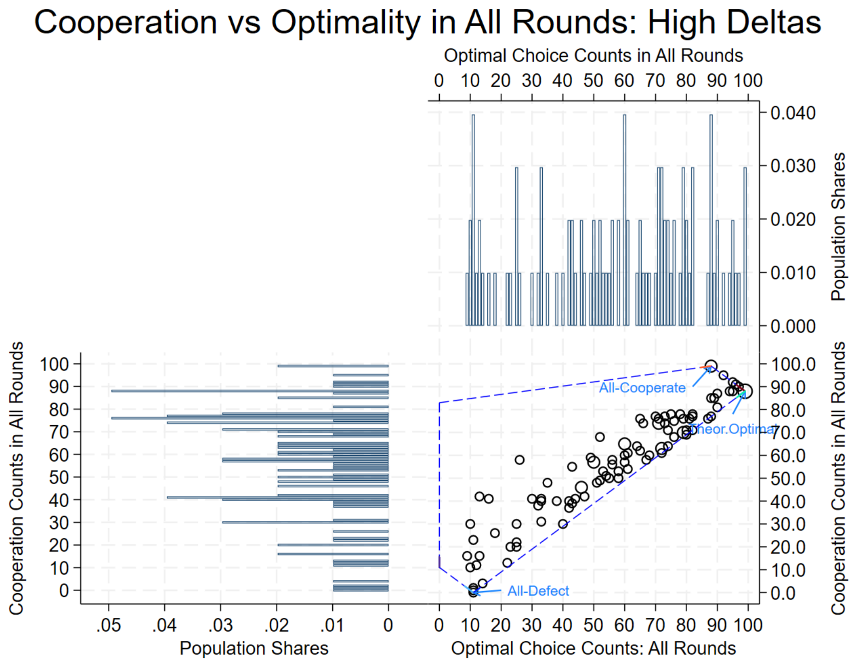

Figure C3: Choices in all 99 rounds: two-dimensional distributions of per-subject counts of cooperation and of optimal choices across all 24 supergames, and the frequency distributions of cooperative (bottom left) and optimal (top right) choices. Bubble size is proportional to the share of subjects (N=101).

theoretically optimal.

As Figures C2 and C3 show, there is a significant heterogeneity in subjects' behavior, without any clear "representative" pattern. The initial heterogeneity of play in the first rounds in Figure C2 (bottom right) is further amplified by the subjects' heterogeneous strategies in the subsequent rounds, depicted in the corresponding panels in Figure C3.

As the bottom right bottom panels of Figure C3 show, only three (out of 99) subjects made perfect theoretically optimal choices (in the far right corner of the polygon), and only two additional subjects made at most three suboptimal choices. In the top corner, two subjects always cooperated, and just one further subject defected four times. In the bottom corner, a single subject always defected, and further three subjects cooperated up to four times. The presence of strategic CaD errors complicates the interpretation of the remaining subjects, most of whom are located away from the boundaries, in the center of the figures, with some subjects also following "end-timing" strategies (see Section 6.5).

Finding 23. *In both subject pools, perfect and near-perfect theoretically optimal behavior is rare, with only 5% of the subjects making no more than 3 theoretically sub-optimal choices. These shares are of similar order of magnitude as the shares of subjects who defected no more than 4 times, and who cooperated no more than 4 times (3 and four, respectively).*

C.6 Random vs. Blocks of Deltas

Does experience playing supergames in blocks with the same δ facilitate learning across supergames? This comparison is also motivated by the deliberate-randomization hypothesis emphasized by Agranov and Ortoleva (2017), who study whether behavior changes when *identical* decision problems are repeated in a row versus repeated at a distance.

From the set of six δ values in the treatments of our experiment, three appear in both the baseline (random order) and high-δ (blocks) treatments: $\delta \in \{0.1, 0.4, 0.67\}$. As Figure C4 (student samples only) shows, there are differences between the random-order presentation of supergames (baseline) and the blocked presentation (high-δ), but these differences do not exhibit a clear pattern. In particular, the rate of optimal choices is higher under blocks only at $\delta = 0.4$ (the value closest to δ^*, where defection is optimal). This lack of a clear pattern is confirmed by the formal tests in Table C3: for $\delta \in \{0.1, 0.4, 0.67\}$, the rate of optimal choices under blocks is, respectively, lower, higher, and essentially unchanged relative to the baseline.

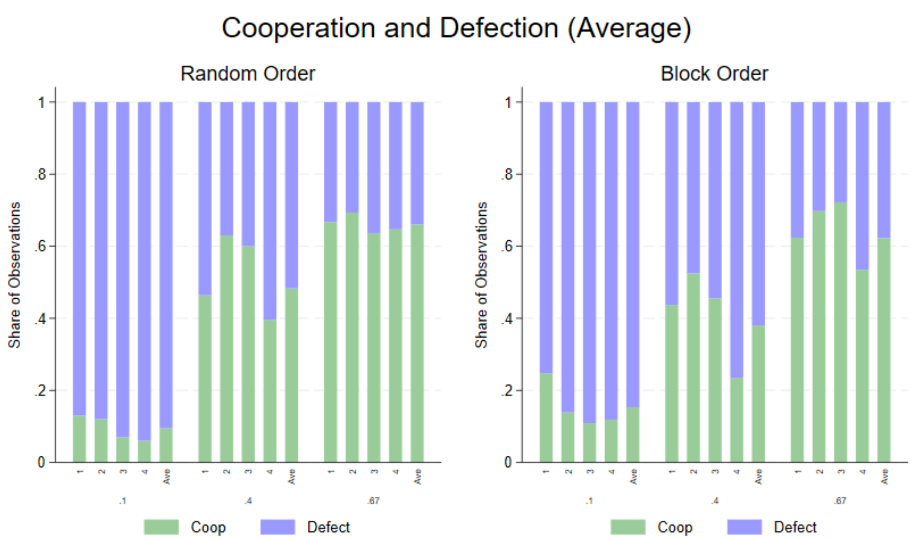

Figure C4: Cooperation and Defection rates across the supergames with $\delta \in \{0.1, 0.4, 0.67\}$ which appear both in the baseline and high-δ treatment (student samples only). Left: Lab sample (N=100), right: high-δ sample (N=101).

Interpreting error rates and payoffs is complicated by random realizations of supergame lengths (see Tables A2 and C2). For $\delta \in \{0.1, 0.4, 0.67\}$, the total number of decisions was $\{4, 7, 12\}$ in the baseline and $\{4, 7, 9\}$ in the high-δ treatment. As a result, subjects earned higher *total* points for $\delta = 0.67$ in the baseline simply because they made more decisions in that treatment. Moreover, some dominated errors—for example, cooperating after defection (CaD)—cannot arise in supergames that last only one round. Thus, the opportunity to make such mistakes was identical for $\delta = 0.1$ in both treatments (all supergames lasted one round), but it was greater in the baseline for both $\delta = 0.4$ and $\delta = 0.67$: for $\delta = 0.4$

(respectively, $\delta = 0.67$), only two (respectively, zero) supergames lasted one round in the baseline, compared to three (respectively, one) in the high-δ treatment. Indeed, in the supergames which lasted longer than a single round (see Table C3, last table subsection for $\delta \in \{0.4, 0.67\}$), the ranking of normalized point totals for $\delta = 0.4$ is reversed as, while, there is no significant difference in optimal choices and dominated CaDs, the rate of end-timing is marginally greater in the baseline (where, by chance, end-timing was profitable overall). And, for $\delta = 0.67$, there is no significant difference across any variable of interest in such longer supergames.

Random vs. Blocks		Random (Lab)		Blocks (high-δ)		t-stat	df	pvalue
		Mean	StDev	Mean	StDev			
δ=0.1 (N=804) Per Round	Cooperate	0.095	0.294	0.153	0.361	-2.52	802	0.012**
	Optimal	0.905	0.294	0.847	0.361	2.52	802	0.012**
δ=0.1 (N=804) Per Supergame	Optimal (All-D+All-C)	0.905	0.294	0.847	0.361	2.52	802	0.012**
	Point Total	115.7	13.21	113.1	16.24	2.52	802	0.012**
	Point Total Per Round	115.7	13.21	113.1	16.24	2.52	802	0.012**
δ=0.4 (N=1407) Per Round	Cooperate	0.483	0.500	0.380	0.486	3.89	1405	0.000***†
	Optimal	0.517	0.500	0.620	0.486	-3.89	1405	0.000***†
	CaD	0.033	0.178	0.034	0.181	-0.11	1405	0.91
δ=0.4 (N=804) Per Supergame	Optimal (All-D+All-C)	0.357	0.480	0.488	0.500	-3.76	802	0.000***†
	Point Total	138.8	54.85	132.4	66.22	1.51	802	0.13
	Point Total per Round	83.87	19.72	88.14	26.30	-2.61	802	0.009***
δ=0.4 (N=301) Per Supergame Lasting >1 round	Optimal (All-D+All-C)	0.330	0.471	0.426	0.497	-1.63	299	0.10
	CaD per Round	0.040	0.126	0.059	0.128	-1.25	299	0.21
	DaC (End-Time)	0.335	0.473	0.228	0.421	1.93	299	0.06*
	Point Total per Round	75.41	12.40	58.96	10.37	11.46	299	0.000***†
δ=0.67 (N=2109) Per Round	Cooperate	0.661	0.474	0.623	0.485	1.81	2107	0.07*
	Optimal	0.613	0.487	0.590	0.492	1.06	2107	0.29
	CaD	0.048	0.215	0.033	0.179	1.74	2107	0.08*
δ=0.67 (N=804) Per Supergame	Optimal (All-D+All-C)	0.520	0.500	0.537	0.499	-0.49	802	0.63
	Point Total	213.7	69.22	167.7	59.16	10.13	802	0.000***†
	Point Total per Round	72.62	10.38	76.78	14.42	-4.70	802	0.000***†
δ=0.67 (N=703) Per Supergame Lasting >1 round	Optimal (All-D+All-C)	0.520	0.500	0.469	0.500	1.34	701	0.18
	CaD per Round	0.043	0.127	0.036	0.121	0.65	701	0.51
	DaC (End-Time)	0.212	0.410	0.215	0.411	-0.07	701	0.95
	Point Total per Round	72.62	10.38	73.51	10.26	-1.14	701	0.25

Table C3: The effect of the order of supergames with the same δ value: Means, standard deviations, and t-tests comparing the baseline (random order) and high-δ (blocks), for each $\delta \in \{0.1, 0.4, 0.67\}$ which appeared both in baseline and high-delta treatments (student samples only). (Significance * 0.10 ** 0.05 *** 0.01 ***† 0.001.)

Finding 24. *For each $\delta \in \{0.1, 0.4, 0.67\}$ that appears in both the baseline and high-δ treatments, we find no systematic effect of presenting supergames in blocks with a common δ on either the frequency of optimal play or realized payoffs.*

C.7 High Deltas: Learning

In Figure C5 (top panel), one can observe an increase in "end-timing" activity by comparing the frequency of such DaC behavior in the first few supergames to that in the last few (and further supported by the tests in Table 5). Further, while the frequency of of dominated CaD errors declines over time, such behavior does not disappear entirely. Figure C5 (bottom panel) further presents the patterns of intra-supergame play across all 24 sequences, split by the treatment order (long first or the reverse).

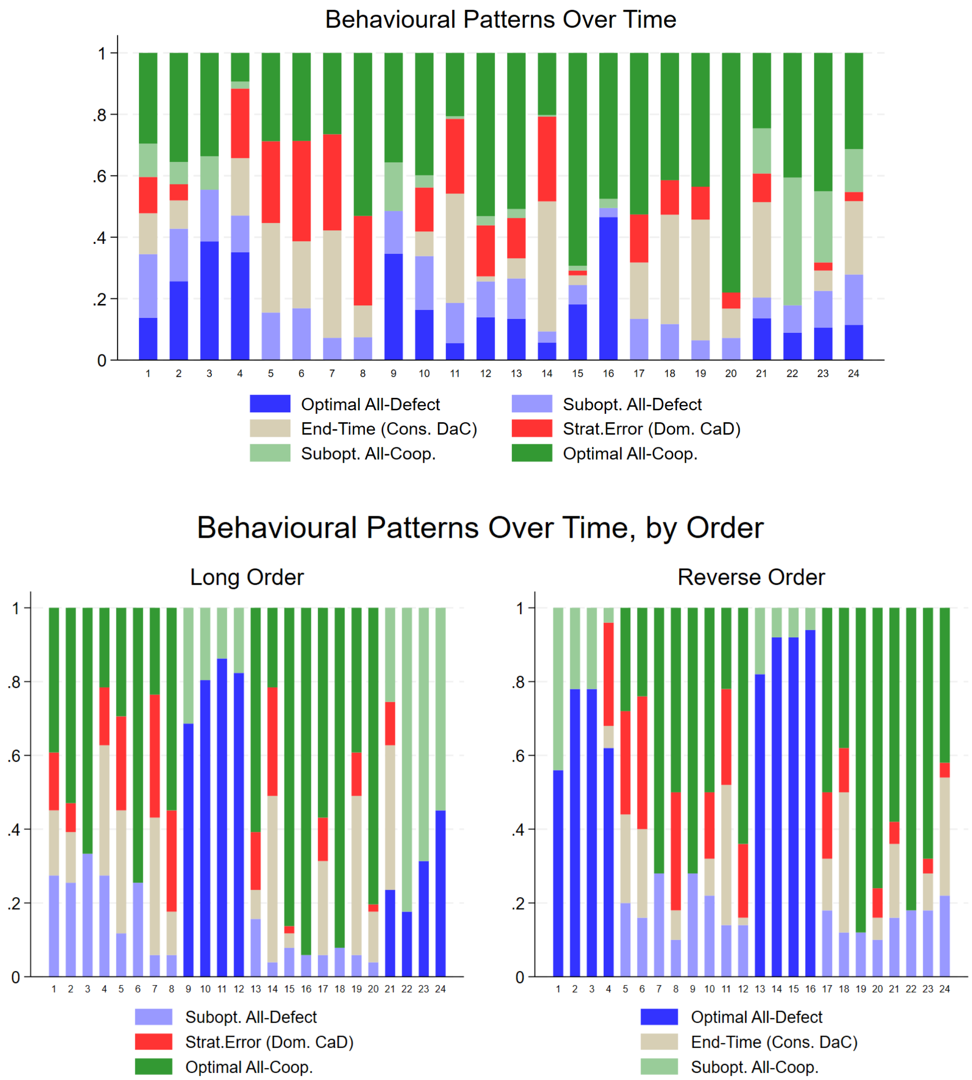

Figure C5: Subject behavior over the sequence of 24 supergames. Top panel: both supergame orders are pooled together so a supergame with a given number could involve different continuation probabilities δ and corresponding optimal actions, depending on the order treatment. Bottom panel: split by order treatment.

C.8 High Deltas: The Patterns of Play Within Each Supergame

We classify subjects' patterns of play within each supergame into 6 *mutually exclusive* types: optimal All-C, optimal All-D, suboptimal All-C, suboptimal All-D, strategic errors (CaD), and end-timing (DaC). As in the baseline, there is no prevalent pattern to subjects' play - see Figure C6. Out of 101 subjects, only three subjects always made perfect, theoretically optimal choices (depicted by dark green and dark blue bars joined at the solid red line close to the middle). Two (one) subjects who always cooperated (defected) are depicted by dark and light bars joined at the solid red line, green on the far left (for All-C) and blue on the far right (for All-D). Both types of non-constant play (CaD and DaC) and both optimal and suboptimal constant play (All-C and All-D) all tend to co-exist in subjects' patterns of play.

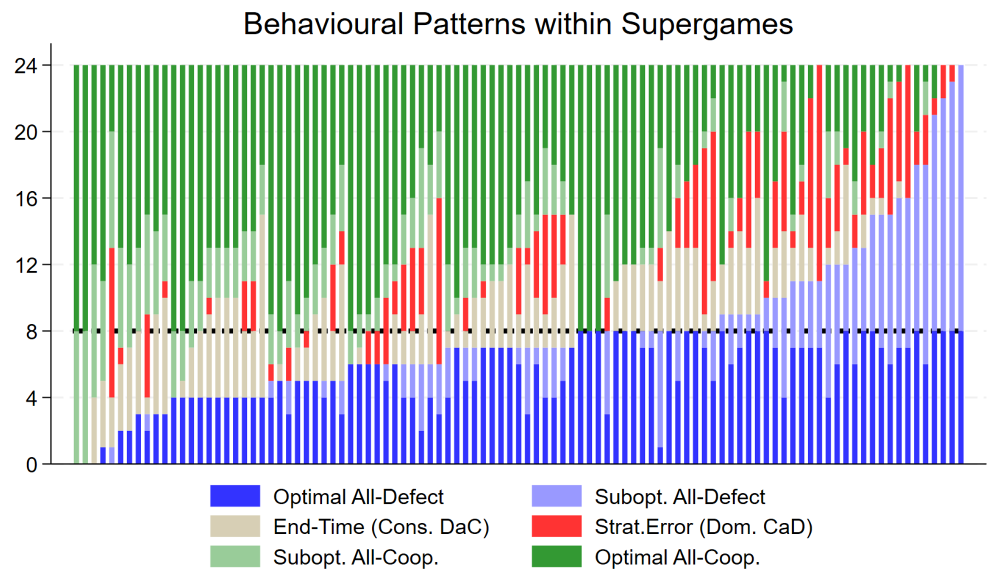

Figure C6: Subject heterogeneity in patterns of choices within supergames, out of 24 supergames, by subject, ordered by the count of supergames with (combined optimal and sub-optimal) All-Defect choices (N=101 subjects). The theoretically optimal strategy involves always defecting in 8 supergames and always cooperating in the remaining 16 supergames (depicted by the solid red horizontal line).

Finding 25. *As in the baseline, there is a notable heterogeneity in subjects' choices in the longer horizon treatment, without any representative pattern. Only three out of 101 subjects always followed the theoretically optimal strategy. Two (one) subjects are fully biased towards cooperation (defection). The rest of the subjects appear to pursue strategies that are neither theoretically optimal nor purely biased.*

C.9 High Deltas: The Effect of Cognitive Abilities

The mean (st.dev.) of the CRT7 score was 3.61 (2.47), with a median of 4 (Figure C7, left).

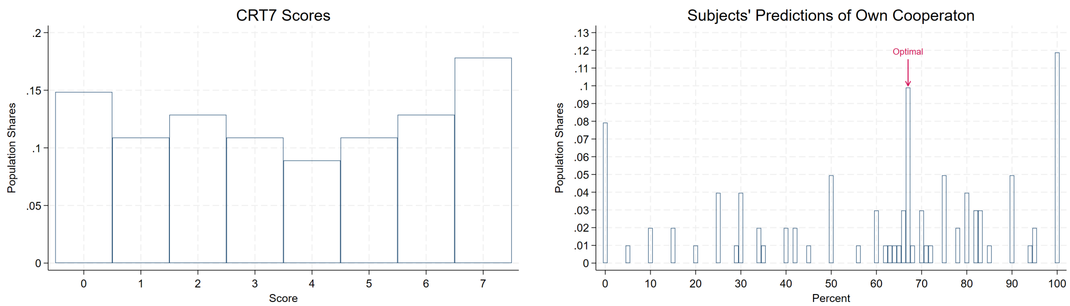

Figure C7: Frequency distributions of CRT7 scores (left) and of the "Prediction" variable (right).

As in the baseline, the CRT7 score predicts the rational aspects of subjects' behavior (see Table C4, even-numbered specifications). Unlike in the baseline, CRT7 does not predict end-timing; instead, end-timing is predicted by being male. The gender coefficient in the odd-numbered specifications may partly reflect the strong correlation between CRT7 and gender ($r = 0.4275$, $p = 0.0000$). (An analysis of gender effects are beyond the scope of this paper). The effect of the CRT7 score remains even if one controls for personality characteristics (see specifications 11-20 in the bottom panel), of which only the number of safe Holt-Laury lottery choices is marginally significant.

Again, the CRT7 score on its own has no effect on the choice to cooperate which can be seen in specifications 3-4 of Table C5, which contains average marginals (dy/dx) from mixed-effects probit regressions. The same applies to the number of safe Holt-Laury choices.

Finding 26. *Higher CRT7 subjects are more likely to make payoff-maximizing choices.*

	Overall Point Totals (OLS)		Dominated(CaD) (Tobit, ll=0)		Theor.Optimal (Tobit, ul=24)		End-Time(DaC) (Tobit, ll=0)		Th.Opt.+End-T(DaC) (Tobit, ul=24)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
CRT7	67.61*	114.82***	-0.51***	-0.71****	0.27	0.51**	0.01	0.16	0.33	0.71***
	(39.37)	(35.15)	(0.19)	(0.19)	(0.21)	(0.20)	(0.13)	(0.13)	(0.23)	(0.22)
Female	-495.87***		2.25**		-2.61***		-1.79***		-4.35****	
	(170.22)		(0.91)		(0.95)		(0.65)		(1.06)	
Age	47.45*		-0.10		0.20		-0.02		0.18	
	(27.77)		(0.19)		(0.17)		(0.12)		(0.19)	
Order Long	42.08	36.78	-0.74	-0.69	-1.92**	-1.95**	1.19**	1.18*	-1.05	-1.09
	(159.09)	(165.03)	(0.82)	(0.85)	(0.87)	(0.91)	(0.58)	(0.60)	(0.95)	(1.03)
Constant	5508.00****	5995.26****	3.91	3.95****	11.86***	13.43****	3.21	1.21*	15.77****	15.22****
	(627.11)	(169.44)	(4.12)	(0.80)	(3.89)	(0.83)	(2.71)	(0.63)	(3.95)	(0.94)
F	8.03	5.81	5.83	7.69	7.15	4.14	3.11	2.56	9.16	5.21
p	0.00	0.00	0.00	0.00	0.00	0.02	0.02	0.08	0.00	0.01
	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)
CRT7	52.54	91.26**	-0.49**	-0.67***	0.19	0.36*	-0.01	0.12	0.24	0.54**
	(46.06)	(40.93)	(0.21)	(0.21)	(0.22)	(0.20)	(0.13)	(0.14)	(0.25)	(0.24)
Female	-466.60**		2.44**		-2.16**		-1.87***		-3.90***	
	(197.44)		(1.00)		(1.03)		(0.68)		(1.16)	
Age	45.77		-0.09		0.20		-0.04		0.17	
	(32.05)		(0.19)		(0.19)		(0.11)		(0.19)	
Lottery	80.98	86.74	-0.40*	-0.43*	0.43*	0.45*	0.15	0.17	0.54*	0.58*
	(50.51)	(53.21)	(0.24)	(0.24)	(0.25)	(0.26)	(0.18)	(0.19)	(0.27)	(0.31)
Risk	-22.18	-2.64	0.25	0.16	0.03	0.12	-0.09	-0.01	0.03	0.20
	(45.13)	(45.85)	(0.26)	(0.27)	(0.22)	(0.23)	(0.15)	(0.16)	(0.26)	(0.29)
Patience	40.62	34.96	-0.14	-0.10	0.20	0.18	0.09	0.06	0.27	0.22
	(47.84)	(48.98)	(0.25)	(0.26)	(0.23)	(0.24)	(0.17)	(0.17)	(0.24)	(0.28)
Punishment	-17.43	-3.75	0.15	0.08	0.06	0.13	-0.05	0.01	0.05	0.16
	(42.13)	(42.15)	(0.20)	(0.21)	(0.19)	(0.20)	(0.15)	(0.15)	(0.22)	(0.22)
Altruism	-31.52	-53.25	0.04	0.13	-0.21	-0.31	-0.18	-0.27	-0.44	-0.61*
	(49.87)	(50.91)	(0.26)	(0.27)	(0.24)	(0.25)	(0.17)	(0.18)	(0.29)	(0.32)
Reciprocity	58.88	93.73	0.09	-0.07	0.30	0.46	0.07	0.18	0.29	0.55
	(75.11)	(78.16)	(0.40)	(0.40)	(0.42)	(0.45)	(0.23)	(0.25)	(0.47)	(0.52)
Retribution	55.86	51.20	-0.10	-0.08	0.38*	0.36	-0.30*	-0.33*	0.14	0.11
	(40.36)	(41.43)	(0.22)	(0.23)	(0.21)	(0.22)	(0.17)	(0.17)	(0.26)	(0.28)
Trust	6.84	6.24	-0.15	-0.15	0.00	-0.00	0.11	0.10	0.05	0.03
	(35.27)	(36.25)	(0.18)	(0.19)	(0.19)	(0.19)	(0.11)	(0.11)	(0.18)	(0.20)
orderlong	27.94	37.90	-0.57	-0.57	-1.97**	-1.92**	1.15**	1.21**	-1.04	-0.96
	(158.63)	(161.76)	(0.78)	(0.81)	(0.84)	(0.86)	(0.55)	(0.58)	(0.92)	(0.97)
Constant	4449.01****	4640.18****	5.00	6.86**	4.70	5.38	3.99	0.66	9.92**	7.53**
	(930.57)	(664.56)	(5.08)	(3.20)	(5.28)	(3.44)	(3.72)	(2.76)	(4.91)	(3.65)
F	3.37	2.74	2.59	2.33	4.30	3.34	2.03	1.75	4.94	3.15
p	0.00	0.01	0.01	0.02	0.00	0.00	0.03	0.08	0.00	0.00

Table C4: Individual differences in rationality, N=101 subjects. (Signif.: *0.10 **0.05 ***0.01 ***†0.001).

Cooperate (Marginals, dy/dx)	All		Split by Median CRT	
			$CRT7 \leq Med$	$CRT7 > Med$
delta=0.4	0.24****	0.23****	0.21****	0.22****
	(0.03)	(0.03)	(0.04)	(0.04)
delta=0.67	0.37****	0.35****	0.31****	0.34****
	(0.03)	(0.03)	(0.04)	(0.04)
delta=0.75	0.46****	0.44****	0.42****	0.41****
	(0.03)	(0.03)	(0.04)	(0.03)
delta=0.8	0.46****	0.43****	0.41****	0.41****
	(0.03)	(0.03)	(0.05)	(0.03)
delta=0.85	0.49****	0.47****	0.44****	0.44****
	(0.04)	(0.03)	(0.05)	(0.04)
Round 2	0.11****	0.10****	0.11****	0.09**
	(0.02)	(0.02)	(0.03)	(0.03)
Round 3	0.13****	0.12****	0.17****	0.04
	(0.03)	(0.02)	(0.03)	(0.04)
Round 4	0.08***	0.08***	0.12****	0.00
	(0.03)	(0.03)	(0.03)	(0.04)
Round 5	0.07**	0.06**	0.11***	-0.01
	(0.03)	(0.03)	(0.04)	(0.04)
Round 6	0.02	0.02	0.07*	-0.05
	(0.03)	(0.03)	(0.04)	(0.04)
Round 7	0.02	0.01	0.06	-0.05
	(0.03)	(0.03)	(0.04)	(0.04)
Round 8	-0.01	-0.01	0.06	-0.08**
	(0.03)	(0.03)	(0.04)	(0.04)
Round 9	0.03	0.02	0.07*	-0.03
	(0.03)	(0.03)	(0.04)	(0.04)
Round 10	0.02	0.02	0.06	-0.04
	(0.03)	(0.03)	(0.04)	(0.03)
Round 11	-0.01	-0.01	0.04	-0.06
	(0.03)	(0.03)	(0.05)	(0.04)
Round 12	-0.02	-0.02	0.07	-0.13***
	(0.04)	(0.04)	(0.04)	(0.05)
Round 13	-0.02	-0.02	0.03	-0.07
	(0.04)	(0.04)	(0.05)	(0.05)
Round 14	-0.03	-0.03	0.02	-0.07
	(0.04)	(0.03)	(0.05)	(0.05)
Round 15	-0.00	-0.00	0.05	-0.07
	(0.03)	(0.03)	(0.04)	(0.05)
Supergame	0.01****	0.01****	0.01****	0.00****
	(0.00)	(0.00)	(0.00)	(0.00)
orderlong	0.04	0.02	0.02	-0.03
	(0.03)	(0.02)	(0.04)	(0.03)
Prior Defect	-0.46****	-0.43****	-0.46****	-0.38****
	(0.01)	(0.01)	(0.02)	(0.02)
CRT7		-0.00		
		(0.01)		
Female		0.01	0.04	-0.00
		(0.03)	(0.05)	(0.02)
Age		0.01**	0.01	0.01
		(0.00)	(0.01)	(0.00)
Prediction		0.02****	0.01**	0.03****
		(0.01)	(0.01)	(0.01)
Lottery		0.00	0.00	0.00
		(0.01)	(0.01)	(0.01)
Risk		0.00	0.00	-0.00
		(0.00)	(0.01)	(0.00)
Patience		0.01	0.02**	-0.00
		(0.01)	(0.01)	(0.01)
Punishment		-0.00	-0.00	-0.00
		(0.01)	(0.01)	(0.01)
Altruism		-0.00	-0.01	0.00
		(0.01)	(0.01)	(0.01)
Reciprocity		0.00	0.00	0.01
		(0.01)	(0.01)	(0.01)
Retribution		0.01**	0.01	0.01*
		(0.00)	(0.01)	(0.00)
Trust		0.00	-0.01	-0.00
		(0.00)	(0.01)	(0.00)
chi2				
p				
N	9999	9999	5841	4158

Table C5: Choices to cooperate: mixed-effects probit regressions, marginals (dy/dx), robust errors in parentheses. (See Table C6 for the corresponding odds.) “Supergame” is the supergame number in the sequence of supergames (scaled down by 24), “Order Long” is a dummy variable for whether the first supergame in the sequence had $\delta = 0.67$, “Prior Defection” is a dummy variable for whether the subject defected in prior rounds of a given supergame, “Prediction” is the subjects’ predictions of the share of their own cooperative choices in Round 1 across all 24 supergames (scaled down by 100), “Lottery” is the number of safe choices in Holt-Laury elicitation. Median CRT7 is 4. (N =101 subjects.) Chi2 and corresponding p-values are from the odds regressions (see Table C6). (Significance * 0.10 ** 0.05 *** 0.01 ***† 0.001.)

C.10 High Deltas: Behavioral Inattention

As in the baseline treatment, we will check here the predictions of the simple inattention theory presented in Section 2.2. In deciding whether to cooperate or defect, individuals with lower cognitive ability (lower attention) will tend to be influenced more by their default values. In contrast, those with higher cognitive ability (higher attention) will tend to be influenced by the structure of the game.

To explore these hypotheses, we again split our samples in two, according to the median CRT7 score (equal to 4). Table C5 presents average marginals (dy/dx) from mixed-effects probit regressions of the choice to cooperate or defect in all 99 rounds of the prisoner's dilemma game (for odds ratios, see Table C6 in the Appendix). The differences between the two subject types are broadly similar to what was found in the baseline treatment. Subjects with relatively high proxies for cognitive ability (CRT7>4, 42 subjects) tend to cooperate (weakly) more than lower CRT7 subjects (CRT7≤4, 59 subjects) for $\delta \geq 0.4$ (see specifications 3 and 4). As the longer-horizon treatment involved longer supergames by design and thus is biased towards cooperation, higher CRT7 subjects do not exhibit any systematic sensitivity to the low round numbers, but they tend to defect in the higher rounds (here, in rounds 8 and 12), consistent with following the end-timing strategy. As in the baseline treatment, cooperation of lower CRT7 subjects is strongly negatively correlated with their self-reported measure of Patience, which also was marginally higher for the higher CRT7 group (one-sided $t = 1.54$, $p = 0.0629$). Note that neither measure of attitudes towards risk is correlated with the decision to cooperate for either subject type.

Comparing Tables C5 and 3, one may spot some qualitative differences in the behavior which, at first glance, are less consistent with the predictions of the rational inattention model. Yet the interpretation of subject choices in longer supergames is more complex than in the shorter ones of the baseline treatment, as there is more room for multiple mistakes within a longer supergame. Looking at subjects' patterns of play (available upon request), lower CRT7 subjects tend to switch from defection to cooperation and back to defection, and so on. This is possibly why lower CRT7 subjects appear as if they tend to cooperate less in response to their prior defection, and to cooperate more in lower-numbered rounds (relative to the baseline of the supergames with $\delta = 0.1$, which all lasted one round only).

Furthermore, while the behavior of higher CRT7 subjects is more consistent with their own Prediction variable, there is some weaker relationship for lower CRT7 subjects as well. However, this latter relationship appears to be driven mostly by subjects with the median CRT7 score of 4—as once the median CRT7 subjects are, instead, ascribed to the higher CRT7 type, the differences between the two types become notably more pronounced (results available upon request). Interestingly, the mean (st. dev.) of the "Prediction" variable was 59.70 (30.28) with a median of 67—bang on the optimal choice of $\frac{2}{3}$ (66.67%)—see Figure C7, right panel.

Finding 27. *Subject behavior is broadly consistent with a simple model of inattention. Co-operative choices made by subjects with a lower proxy for cognitive ability (lower attention)* *correlate with an elicited proxy for their degree of patience. In contrast, cooperative choices made by those with a higher proxy for cognitive ability (higher attention) are more affected by the structure of the game, and do not correlate with their individual characteristics.*

The above regression results are summarized visually in Figure C8 which shows that the two groups of subjects, divided by CRT7 score, exhibit different patterns of play. Subjects in the lower CRT7 group (the left panel of Figure C8 make relatively more frequent strategic errors (CaD) and more suboptimal consistent defection behavior (Suboptimal All-D). By contrast, subjects in the higher CRT7 group (the right panel of Figure C8) are closer to the theoretically optimal policy and engage in end-timing behavior (DaC) more often.

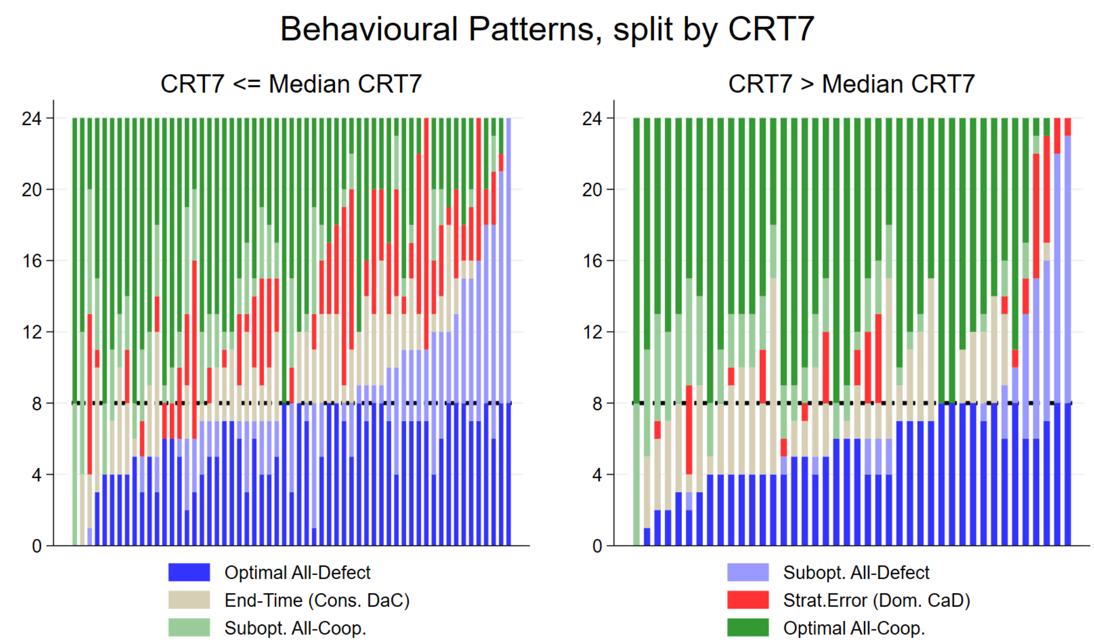

Figure C8: Inattention: subjects' patterns of choices within supergames (out of 24 supergames), split by median $CRT7$. Patterns are presented for each subject, ordered by the count of supergames with (combined optimal and sub-optimal) All-Defect choices (N=101 subjects). The theoretically optimal strategy involves always defecting in 8 supergames and always cooperating in the remaining 16 supergames (represented by the horizontal line).

C.11 Behavioral Inattention: Further Results

Figure C7 presents the distributions of the two key variables for the inattention model, $CRT7$ and the Prediction variable. Table C6 presents the odds corresponding to the regressions of Table C5.

Cooperate	All		Split by Median CRT	
(Marginals, dy/dx)			$CRT7 \leq Med$	$CRT7 > Med$
delta=0.4	1.08****	1.08****	0.86****	1.44****
	(0.13)	(0.13)	(0.16)	(0.20)
delta=0.67	1.63****	1.63****	1.28****	2.21****
	(0.16)	(0.16)	(0.20)	(0.29)
delta=0.75	2.04****	2.04****	1.73****	2.62****
	(0.16)	(0.16)	(0.21)	(0.25)
delta=0.8	2.04****	2.04****	1.69****	2.64****
	(0.17)	(0.17)	(0.22)	(0.30)
delta=0.85	2.20****	2.20****	1.84****	2.85****
	(0.18)	(0.18)	(0.22)	(0.34)
Round 2	0.48****	0.48****	0.48****	0.55***
	(0.10)	(0.10)	(0.12)	(0.19)
Round 3	0.58****	0.58****	0.72****	0.28
	(0.11)	(0.11)	(0.13)	(0.22)
Round 4	0.36***	0.36***	0.52****	0.02
	(0.12)	(0.12)	(0.14)	(0.26)
Round 5	0.29**	0.29**	0.46***	-0.07
	(0.12)	(0.12)	(0.14)	(0.26)
Round 6	0.10	0.09	0.30*	-0.34
	(0.14)	(0.14)	(0.16)	(0.26)
Round 7	0.07	0.06	0.25	-0.33
	(0.14)	(0.14)	(0.18)	(0.25)
Round 8	-0.02	-0.03	0.23	-0.54*
	(0.14)	(0.14)	(0.17)	(0.28)
Round 9	0.12	0.11	0.30*	-0.22
	(0.13)	(0.13)	(0.16)	(0.25)
Round 10	0.09	0.09	0.27	-0.24
	(0.13)	(0.13)	(0.17)	(0.23)
Round 11	-0.04	-0.04	0.16	-0.38
	(0.15)	(0.15)	(0.19)	(0.25)
Round 12	-0.11	-0.11	0.28	-0.85**
	(0.17)	(0.17)	(0.18)	(0.34)
Round 13	-0.11	-0.12	0.11	-0.48
	(0.17)	(0.17)	(0.21)	(0.33)
Round 14	-0.12	-0.12	0.09	-0.43
	(0.16)	(0.16)	(0.19)	(0.31)
Round 15	-0.02	-0.02	0.23	-0.43
	(0.15)	(0.15)	(0.17)	(0.31)
Supergame	0.02****	0.02****	0.02****	0.03****
	(0.00)	(0.00)	(0.01)	(0.01)
orderlong	0.18	0.10	0.09	-0.16
	(0.12)	(0.11)	(0.15)	(0.20)
Prior Defect	-2.04****	-2.04****	-1.90****	-2.43****
	(0.10)	(0.10)	(0.13)	(0.18)
CRT7		-0.01		
		(0.03)		
Female		0.04	0.17	-0.02
		(0.13)	(0.22)	(0.15)
Age		0.04**	0.05	0.04
		(0.02)	(0.03)	(0.03)
Prediction		0.09****	0.05**	0.19****
		(0.03)	(0.03)	(0.06)
Lottery		0.02	0.00	0.01
		(0.04)	(0.04)	(0.06)
Risk		0.00	0.00	-0.03
		(0.02)	(0.04)	(0.02)
Patience		0.05	0.10**	-0.00
		(0.03)	(0.04)	(0.05)
Punishment		-0.00	-0.00	-0.00
		(0.03)	(0.04)	(0.03)
Altruism		-0.01	-0.04	0.01
		(0.04)	(0.03)	(0.05)
Reciprocity		0.02	0.00	0.08
		(0.04)	(0.04)	(0.08)
Retribution		0.06**	0.05	0.05*
		(0.02)	(0.04)	(0.03)
Trust		0.00	-0.03	-0.00
		(0.02)	(0.04)	(0.03)
Constant	-1.53****	-3.69****	-3.32****	-4.70****
	(0.14)	(0.64)	(0.99)	(1.13)
chi2	1014.88	1174.30	831.35	11034.10
p	0.00	0.00	0.00	0.00
N	9999	9999	5841	4158

Table C6: Choices to cooperate: mixed-effects probit regressions, odds, robust errors in parentheses. (See Table C5 for the corresponding marginals.) "Supergame" is the supergame number in the sequence of supergames (scaled down by 24), "Order Long" is a dummy variable for whether the first supergame in the sequence had $\delta = 0.67$, "Prior Defection" is a dummy variable for whether the subject defected in prior rounds of a given supergame, "Prediction" is the subjects' predictions of the share of their own cooperative choices in Round 1 across all 24 supergames (scaled down by 100), "Lottery" is the number of safe choices in Holt-Laury elicitation. Median CRT7 is 4. (N =101 subjects.) (Significance * 0.10 ** 0.05 *** 0.01 **** 0.001.)