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

John Duffy† Ed Hopkins‡ Tatiana Kornienko§

September 15, 2026

Abstract

Cooperation in repeated interactions is important for much socioeconomic activity. We study it in a simple dynamic setting designed to eliminate confounding factors including multiple equilibria, strategic uncertainty, and other-regarding concerns. Subjects play indefinitely repeated prisoner's dilemma games against a computer known to follow the Grim-trigger strategy, with different continuation probabilities. Cooperation increases with the continuation probability, as theory predicts. Nevertheless, many subjects make strictly dominated choices by cooperating after previously defecting; many also defect after initially cooperating, consistent with attempted "end-timing." A behavioral-inattention model organizes these findings: subjects with higher cognitive-test scores make fewer errors and earn higher payoffs.

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 Sciences. The experimental protocol was approved by the UC Irvine Institutional Review Board, HS # 2015-2082.

†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). Repeated games famously involve "the shadow of the future", trading off future benefits against current costs. Yet, little is known about how making this basic trade-off depends on cognition and how this interacts with other factors. The literature has focused on many possible other confounds, including diverse beliefs about opponents' strategies, heterogeneous risk attitudes, and social preferences.

This paper 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. 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 δ^*, here $\delta^* = 0.5$. Since opponent behavior is perfectly predictable here, we can observe whether subjects implement the optimal policy over the *entire supergame*, rather than concentrating on the first round as in much existing experimental analysis of repeated games.

We report on two related studies. Study 1 is a baseline laboratory experiment in which subjects face a set of continuation probabilities δ skewed toward lower values, and thus shorter expected supergames. In this design, it is theoretically optimal to always cooperate in one-third of the supergames and to always defect in the remaining two-thirds. Study 2 provides a laboratory robustness check in which a different group of university subjects face a set of continuation probabilities skewed toward higher values, and thus longer expected supergames, comparable to prior repeated-game experiments.[1] By contrast with Study 1, in

[1]Study 2 further differs from Study 1 in that the continuation probability δ was fixed for four consecutive

Study 2 it is theoretically optimal to always cooperate in two-thirds of the supergames and to always defect in the remaining one-third. We also report, in an appendix, an Amazon Mechanical Turk (AMT) replication of Study 1 using a more heterogeneous subject pool, with similar results.

Despite our simple setup involving a robot opponent playing a known strategy, the support for the theoretical predictions is mixed. We find that, overall, subjects do respond to the dynamic incentives by increasing their cooperation with an increase in the continuation probability δ. However, this response is more gradual rather than the step function predicted by standard theory. For example, the theory predicts zero cooperation when $\delta = 0.1 < \delta^*$ and full cooperation when $\delta = 0.67 > \delta^*$. But we find that, for these δ values, the cooperation rates across all rounds were, respectively, 9.5% (15.3%) and 66.1% (62.3%) in Study 1 (Study 2). In terms of theoretically optimal actions (i.e., those that would maximize their monetary earnings *ex ante*, given the uncertain end of a supergame) in all rounds of all supergames, the median subject made just under two thirds of such decisions in Study 1 and three fifths in Study 2. There is little reason to expect subjects to make fewer dynamic optimization mistakes in a more complex setting with human opponents, with added strategic uncertainty, multiple equilibria, and uncertainty about others' preferences.

Using subjects' choices in the first round of each supergame as a simple test of the theoretical predictions, we find that first-round cooperation rises sharply with the continuation probability δ. In the first rounds of each supergame, at the lowest value of δ tested ($\delta = 0.1$), cooperation rates are 9.5% (15.3%) of all first round choices in Study 1 (Study 2); at the highest values tested ($\delta = 0.7$ in Study 1 and $\delta = 0.85$ in Study 2), they rise to 76.3% and 77.7%, respectively. This responsiveness to δ is considerably larger than the estimates implied by Dal Bó and Fréchette (2018) for standard human subject-versus-human subject experiments, suggesting that our design succeeds in reducing strategic uncertainty. Yet, relative to the theoretical benchmark, subjects frequently chose the wrong action: they cooperated too often when $\delta < 0.5$ and too little when $\delta > 0.5$. In particular, suboptimal cooperation in the first round accounted for approximately 15% (21%) of first-round choices

supergames to promote equilibrium learning.

in Study 1 (Study 2), indicating that behavior remains highly noisy even in this simpler, single-person decision environment.

As noted, our methodology enables us to examine behavior beyond the initial round and document two unexpected types of dynamic deviation from the optimal strategy. First, after defecting in a supergame—and thereby triggering defection by the robot for all remaining rounds—52% (60.4%) of subjects in Study 1 (Study 2) nevertheless cooperated in at least one subsequent round. Because defection strictly dominates cooperation once the robot's Grim-trigger punishment has been activated, this behavior is difficult to rationalize. This "cooperate after defecting" (CaD) behavior becomes less frequent with experience but never completely disappears.

Second, 78% (75%) of subjects in Study 1 (Study 2) defect at least once after initially cooperating, a pattern we call "defect after cooperating" (DaC). A DaC sequence is never an optimal policy: before the first defection, the decision problem is unchanged, so cooperation, if optimal initially, continues to be optimal as long as no defection has occurred. Although the end of each supergame is randomly determined and cannot be known in advance, subjects appear to engage in "end-timing": gambling on when a supergame will end by defecting in the round they guess will be the last.[2] This "end-timing" strategy resembles "sniping" in auctions (see Roth and Ockenfels (2002)), and we find that it becomes more frequent with experience. Our single-person experimental design allows us to identify and interpret this behavior particularly clearly, and our findings offer an alternative interpretation of results from repeated-game experiments involving matched pairs of human participants.[3]

Our results thus show that subjects respond to payoffs but also depart frequently from the policy that maximizes expected monetary earnings. The data reveal further substantial heterogeneity – some subjects never cooperate while others always do, regardless of δ val-

[2]As is standard in indefinitely repeated games, we employ a constant termination probability of $1-\delta$, where δ is fixed for the duration of a supergame and known to subjects. Subjects may nonetheless hold the non-Bayesian belief that termination becomes more likely as the number of completed rounds increases. Alternatively, Mengel et al. (2022) find that past *realized* supergame lengths affect subjects' decisions in subsequent supergames.

[3]Romero and Rosokha (2018) and Cooper and Kagel (2023) also report declining cooperation rates in indefinitely repeated prisoner's dilemma experiments. In those settings, however, cooperation may decline because subjects believe that their opponents are about to stop cooperating.

ues. To explain this diversity of rationality and biases, we propose and test a novel model of behavioral inattention that allows for payoff responsiveness, individual inclinations, and substantial error, with rational behavior serving as a special case. We adapt the approach of Gabaix (2019) and assume that the agent chooses an action to match an unknown payoff associated with an unknown state of the world. The agent may, however, be inattentive to the payoff-generating process, including the continuation probabilities δ, the opponent's strategy, the experimental design, and other relevant information. This formulation of inattention is convenient as it directly implies a probit choice rule with three predictions. First, subjects will only noisily respond to expected payoffs. Second, the responsiveness to payoffs is increasing in cognitive ability, with an error rate that is decreasing in cognitive ability. Third, individuals with a less precise understanding of the decision problem will be more influenced by their own "default" prior payoff which is typically obtained in comparable situations outside the laboratory.

The gradual rather than discontinuous response of cooperation rates to δ is consistent with the model's first prediction. Consistent with the model's second prediction, subjects with higher cognitive reflection test (CRT) scores, which proxy for cognitive ability, tend to earn higher payoffs. They are less likely to cooperate after triggering the opponent's irreversible defection and behave more closely in line with the theoretical benchmark. There is also evidence that they are more likely to engage in end-timing. The model's third prediction may explain why, in both studies, first-round cooperation (before any response by the computer opponent) is correlated with subjects' intrinsic patience (which we interpret as a default prior for cooperation) – but only among subjects with lower cognitive test scores. By contrast, the standard theoretical benchmark predicts that behavior is independent of individual characteristics.

Experimental economists have used robot players for control purposes in a number of studies.[4] Two prior studies are most closely related to our own: Roth and Murnighan (1978) and Murnighan and Roth (1983) (the first studies to run experiments on supergames with

[4]For instance, Houser and Kurzban (2002) and Johnson et al. (2002) use robot players to remove the influence of social preferences in applications involving finitely repeated games. For surveys of experiments combining human subjects and robot players, see March (2021) and Bao et al. (2022).

an uncertain end) have subjects play a repeated PD game against a fixed strategy. However, subjects in these two studies faced some strategic uncertainty, as they were *not informed* of the strategy they faced or that their opponent was, in fact, the experimenter.[5] By contrast, in this experiment we tell subjects explicitly that they are playing against a programmed opponent who plays the *Grim trigger strategy*, and we carefully explain what this means and test subjects' understanding of this strategy. Duffy and Xie (2016) study play against robot players known to follow 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 as we do here. Also related is Andreoni and Miller (1993), who study *finitely repeated* PD games in which there is a known probability that the opponent is a robot tit-for-tat player rather than 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 that players with higher cognitive ability cooperate more, make fewer mistakes, and earn 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 defecting immediately after mutual cooperation 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.

In recent and independent work, Reverberi et al. (2021) also experimentally study repeated play by human subjects against robot opponents although with a different design

[5]In Roth and Murnighan (1978), subjects were told they faced a programmed opponent but were not informed of its strategy (p. 194); the opponent was actually an experimenter playing Tit for Tat. In Murnighan and Roth (1983), subjects were told they faced a different, unidentified person in each session, but all in fact played an experimenter using either Tit for Tat or Grim trigger (p. 289). These designs were intended to "control for differences in subjects' behavior due to differences in their opponents" (Roth and Murnighan, 1978, p. 194).

and research objectives. In their experiment, both the stage game and the robot strategy could change randomly over time. They examine how the complexity of the robot's strategy interacts with subject's cognitive ability to determine the frequency of mistakes. Normann and Sternberg (2023) and Normann et al. (2025) investigate interactions between pricing algorithms and human agents in repeated oligopolies and point out that human–algorithm interactions may be of increasing practical importance. Finally, Blonski et al. (2025) have human subjects play against two types of robot players. Unlike our design, however, none of these studies fully disclosed the algorithmic strategy to the human subjects, so that strategic uncertainty remained an inherent part of the decision problem.

Surprisingly, despite the importance of dynamic optimization in contemporary economic theory, only a few experiments, Noussair and Olson (1997) and Carbone and Duffy (2014), have explored subjects' capacity for dynamic optimization and also found deviations from optimal behavior. In our study, the problem is even simpler, as it does not involve a 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 resembles that of Charness and Levin (2009), who experimentally demonstrate the winner's curse in a single-person bidding task: in both studies, simplifying the environment identifies 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, in which subjects face conditions where cooperation and non-cooperation are each optimal. Thus, our study of repeated interactions builds on the methodology of Duffy et al. (2021) and Charness et al. (2021), who also use experimental designs with contrasting environments.

2 Theory and Hypotheses

In our experiment, subjects play an 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),

$$\begin{array}{cc} & \begin{array}{cc} \text{X} & \text{Y} \end{array} \\ \begin{array}{c} \text{X} \\ \text{Y} \end{array} & \begin{array}{cc} \textbf{75}, \textit{75} & \textbf{15}, \textit{120} \\ \textbf{120}, \textit{15} & \textbf{30}, \textit{30} \end{array} \end{array} \tag{1}$$

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 (Y) in the first round, and every round thereafter 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 $\delta^* = 0.5$.

Since the continuation probability δ is constant over time, the decision problem is *stationary*: if cooperation is optimal in round 1, it is optimal in every subsequent round. Hence, when cooperation is initially optimal, it remains optimal after every history in which both players have cooperated. [6] 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. *Optimal Play: in order to maximize expected money earnings, subjects should play Cooperate/X (Defect/Y) in every round of every supergame when* $\delta > (<)\ \delta^* = 0.5$.

This assumes risk neutrality. Risk aversion changes the calculation of δ^*. For example, applying CRRA utility to a subject's total accumulated earnings over the (randomly determined) supergame length, one can show numerically that δ^* rises with risk aversion – so defection can become optimal even at δ values 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, among the δ values appearing in our experimental design, consider the lowest for which cooperation is optimal under risk neutrality – which is $\delta = 0.67$. To make defection optimal at this value, a CRRA parameter above 1.1 is needed.[7] Lastly, risk aversion does not change the stationarity of the problem - in a given supergame, subjects should either always cooperate or always defect. Thus, we do not expect risk attitudes to affect the main predictions for play in our games.

Further, 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 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.

[6] 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 final round. This very different prediction is one of the reasons why we concentrate on uncertain length supergames here.

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

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 of Behavioral Inattention

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 to explain heterogeneity in deviations from optimality in single-person decision problems. Importantly, this model generates the novel and testable predictions that subjects with lower cognitive ability both make more mistakes and are also more influenced by irrelevant non-cognitive factors.

We assume subject i faces a single-person decision problem, entering with a prior default payoff from cooperation, d_i – a non-cognitive characteristic reflecting her experience with cooperation outside the laboratory. In the lab, faced with a supergame having 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 assume that the precision of the signal an agent receives, rather than being fixed exogenously, depends on her cognitive ability, which varies across individuals. As in Gabaix (2019) and 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,[8] 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 low cognitive ability, we have the result that the posterior is close to the individual's default payoff d_i. This shows *inattention*, in that the individual does not fully respond to the environment she faces. Again, our innovation is to link the level of inattention to cognitive ability.

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 $P_i > 0$.[9] Thus, when facing a decision problem with continuation probability δ, the subject's probability of

[8] 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.

[9] 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.

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, and is a continuous, strictly increasing function. 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. Given the probability of cooperation is increasing in π which is increasing in the continuation probability δ, the model leads to this simple hypothesis.

Hypothesis 2. *The cooperation rate is increasing in the continuation probability δ.*[10]

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 testable 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.[11] Second, the probability of cooperation C will be closer to its optimal value, resulting in higher ability (lower noise) individuals earning higher payoffs.

Hypothesis 3. *High-noise, low-cognitive-ability subjects are less responsive to the true payoff to cooperation, and hence to the continuation probability δ, and earn lower average payoffs than low-noise, high-cognitive-ability subjects.*

Further, the model predicts that for noisier individuals payoffs matter relatively less than initial priors or propensities toward cooperation.

Hypothesis 4. *High-noise, low-cognitive-ability subjects are more influenced by their default payoff to cooperation based on ex ante factors that favor cooperation.*

[10] Strictly speaking, this is not implied by the earlier Hypothesis 1. For example, standard theory does not predict any change in behavior for changes in δ that do not cross the threshold level of δ^*.

[11] 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.

3 Experimental Design

We conducted two main experimental studies, where the main task consisted of 24 indefinitely repeated prisoner's dilemma (IRPD) games, or "supergames," played against a computer program known to use the Grim trigger strategy. 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 IRPD 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.

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. 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 B.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.

After subjects were presented the list of all continuation probabilities δ, but before play

of the first 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 B.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.[12]

Subjects were told that at the end of the session, six supergames—one for each δ—would be randomly selected for payment. During play, they were always informed of the probability that the current supergame would continue. On the *same* decision screen, where they made their own choice for the round they were also reminded which action (X or Y) the computer would take for that same round under the Grim trigger strategy, given the history of the current supergame (see Appendix B.3). Thus, strategic uncertainty about the computer's play was eliminated.

3.1 Study 1 and Diverse Subject Pool Replication Study ("AMT")

For Study 1,[13] 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 both *extremely biased* types of subjects – for those who always cooperate and those who always defect. 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

[12]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.

[13]Study 1 (and its AMT replication) is a part of a project pre-registered on the AEA RCT registry, https://doi.org/10.1257/rct.6318-1.0.

play instead entails cooperating in only 8 of the 24 supergames (where $\delta = 0.67$ or 0.7) and defecting in the remaining 16. Thus, unlike most existing studies, the distribution of continuation probabilities here is intentionally weighted toward shorter expected supergame durations. Half of 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. (See Tables C1 and C2 for the sequence of randomly drawn continuation probabilities δ, realized durations and their expected durations, and Appendix C.2 for a discussion of order effects.)

Study 1 involved a gender-balanced laboratory sample consisting 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 laboratory subjects were conducted online via Zoom in the aftermath of the Covid pandemic (late 2020). An experimenter verified each participant's identity and was continuously present to monitor activity and answer questions, ensuring attention and supervision comparable to conventional laboratory settings.

Study 1 subjects were paid USD \$0.01 per point from six randomly selected supergames, one for each δ value, with a \$7 fixed show-up payment. Total earnings averaged \$17.90 for a one-hour session.[15] The order of supergames was balanced across subjects. Exactly half of the subjects (50/100, or 4 of 8 sessions) faced the "long" order first (first supergame $\delta = 0.67$), the remainder faced the reverse order (starting with $\delta = 0.33$).

We also conducted a replication of Study 1 (which we will refer to as "AMT") on a more diverse subject pool, with 149 U.S.-resident subjects from Amazon's Mechanical Turk (AMT) platform. For direct comparability with Study 1, it used the same experimental program, and followed exactly the same experimental specifications, including the randomly

[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]After completing the 24 repeated PD games, Study 1 subjects (only) were randomly paired for another two-player task, not reported here, in which they could earn an additional \$1.00–1.70 not included in the \$17.90 average.

drawn and subsequently hard coded sequences of the supergames. The recruitment details, payments, demographics, and experimental results are reported in Appendix G.

3.2 Study 2: Robustness Check

As a robustness check, we conducted Study 2, which differed from Study 1 in four key respects (see implementation details in Appendix D.1–D.3).[16] First, it involved *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. Second, the δ values were implemented in *blocks* of four supergames rather than randomly, allowing learning within each δ-block; instructions were modified accordingly. Third, to ensure full experimental control, the new sessions were conducted using standard, *in-person* procedures (in 2025) at the UC Irvine laboratory: instructions were read aloud, and subjects completed all tasks individually at workstations. Finally, we added a Holt and Laury (2002) incentivized risk-elicitation task, as risk attitudes may influence optimal play (see Section 2.1); Study 1 used only an unincentivized Likert-scale measure. Study 2 involved a new panel of 101 UC Irvine undergraduates (64% female), distinct from the Study 1 sample. Average earnings were \$26.52—higher than in Study 1 due to longer games under higher average δ and a \$3 increase in the show-up payment.

The design changes to Study 2 were made to increase the level of control, to closer align it with existing experiments on repeated games, and to give subjects additional opportunities to learn. Nevertheless, as our findings below show, the subject behavior was found to closely parallel that in Study 1, and the risk attitudes elicited in Study 2 using the incentive compatible Holt-Laury method were found to have no explanatory power. For comparability, the results of the two studies are presented below side-by-side.

4 Results

In both studies, each subject completed 24 supergames, each lasting at least one round. Two statistical aspects of our design are important. In Study 1 (involving $\delta \in \{0.1, 0.25, 0.33, 0.4, 0.67, 0.7\}$,

[16] Study 2 was preregistered on the AEA RCT registry `https://doi.org/10.1257/rct.15844`.

i.e., two thirds of δs below 0.5), the theoretically optimal strategy implies 8 cooperative and 16 defecting first-round choices, so first-round optimal choices are skewed toward defection. In Study 2 (involving $\delta \in \{0.1, 0.4, 0.67, 0.75, 0.8, 0.85\}$, i.e., two thirds of δs above 0.5), theoretically optimal play is reversed, with 16 cooperative and 8 defecting first-round choices, so first-round optimal choices are skewed toward cooperation. But, in *both* studies, the expected theoretically optimal play is skewed toward cooperation in later rounds – as, due to random termination, later rounds are more likely to occur when δs are higher – and higher δ values make cooperation optimal.

In Study 1, owing to random termination, 11 supergames ended after a single 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 decisions per subject (see Tables C1 and C2). 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).

In Study 2, by chance, 11 supergames ended after one round, while the remaining 13 lasted 2–15 rounds, yielding 24 first-round and 75 subsequent-round choices, for a total of 99 decisions per subject (see Table D2). There, the theoretically optimal strategy is to defect in all rounds of the 8 supergames with $\delta \in \{0.1, 0.4\}$ (11 defections) and to cooperate in all rounds of the remaining 16 supergames (88 cooperative choices).

4.1 The Headline Result

Figure 1 summarizes the full dataset. For each of the six δ values, it reports the shares of cooperate and defect actions pooled across all subjects in the relevant 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.

For both studies, three broad features stand out. First, cooperation broadly increases with the continuation probability δ. Second, cooperation is not driven to zero in the supergames with $\delta < \delta^* = 0.5$, and it is less than full for the δ values above δ^*. Third, beginning with the second supergame, behavior changes little across the remaining supergames for

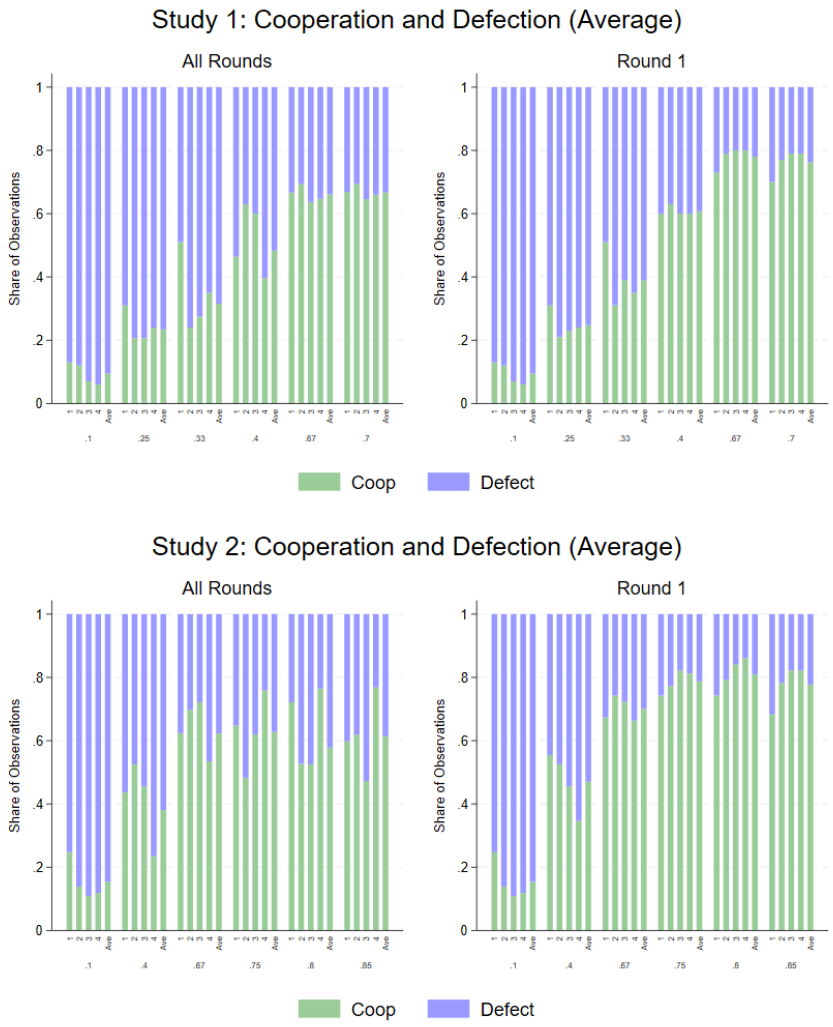


Figure 1: Cooperation and Defection rates in the four supergames in chronological order (first horizontal axis) of each of the six δ values (second horizontal axis). Left: across all rounds. Right: across first rounds only. Top: Study 1 ($N = 100$ subjects); Bottom: Study 2 ($N = 101$ subjects).

a given δ, suggesting limited learning over time – even in Study 2 where the four supergames for each δ were played consecutively, potentially facilitating learning. While these aggregate patterns are broadly consistent with the standard game-theoretic predictions (Hypothesis 1), the observed deviations from these predictions are notable.

These deviations from the theoretical predictions translate into low rates of fully theoretically optimal behavior as shown in Figure 2 (left panels) by the low population shares with high counts of optimal choices (on the far right of the horizontal axis). In Study 1 (top left panel), the mean (s.d.) number of theoretically optimal choices across *all* 48 decisions

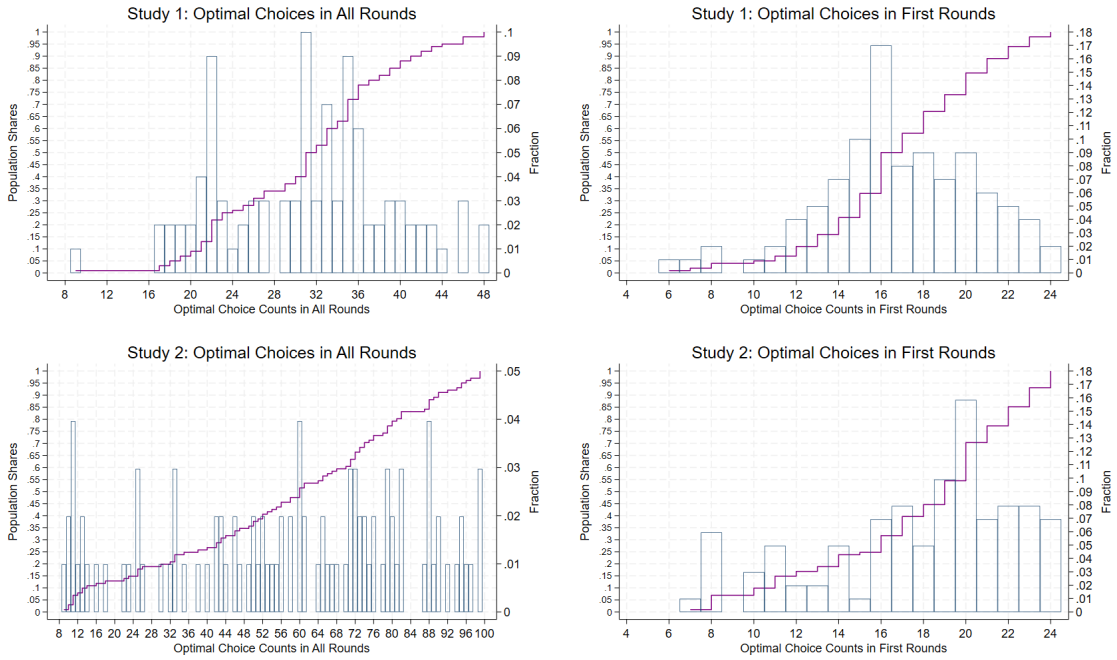


Figure 2: Frequency densities and cumulative distributions of per-subject counts of optimal choices. Left: all rounds of the 24 supergames (48 decisions per subject in Study 1; 99 in Study 2). Right: first rounds of the 24 supergames (24 decisions per subject). Top: Study 1 ($N = 100$ subjects); bottom: Study 2 ($N = 101$ subjects).

is 31.05 (8.06), or 64.7%, with a median subject making 31.5 (or 65.6%) of theoretically optimal choices. In Study 2 (bottom left panel), the corresponding mean (s.d.) is 57.20 (26.41) optimal choices out of 99 decisions, or 57.8%, and a median of 60 (or 60.6%) – lower than in Study 1 in percentage terms.

Finding 1. *In Study 1 (2), across all 48 (99) decisions in all 24 supergames, the median subject made just under two thirds (Study 1) and about three fifths (Study 2) of theoretically optimal decisions in all rounds of all supergames.*

4.2 First Round Behavior

First-round choices, made before any feedback, capture subjects' initial strategic tendencies. Figure 1 (right panels) shows average cooperation and defection rates in the *first round* of supergames 1–4 for each δ. Again, cooperation is not driven to zero in the games with $\delta < \delta^* = 0.5$, and it is less than full for δ values above δ^*. There is again not much evidence

of learning after the second supergame.

Figure 2 (right panels) displays the distributions of optimal first-round choices across 24 supergames. In Study 1 (top right), the mean (s.d.) of theoretically optimal choices across first-round decisions in 24 supergames is 16.81 (3.74), with the median subject making 16.5 (or 68.8%) of theoretically optimal choices. In Study 2 (bottom right), the corresponding mean (s.d.) is 17.8 (4.60) and the median is 19 (or 79.2%). While theoretically optimal behavior in the first rounds is more frequent in Study 2 than in Study 1, it does not translate into higher frequency of overall theoretically optimal choices (see Section 4.1).

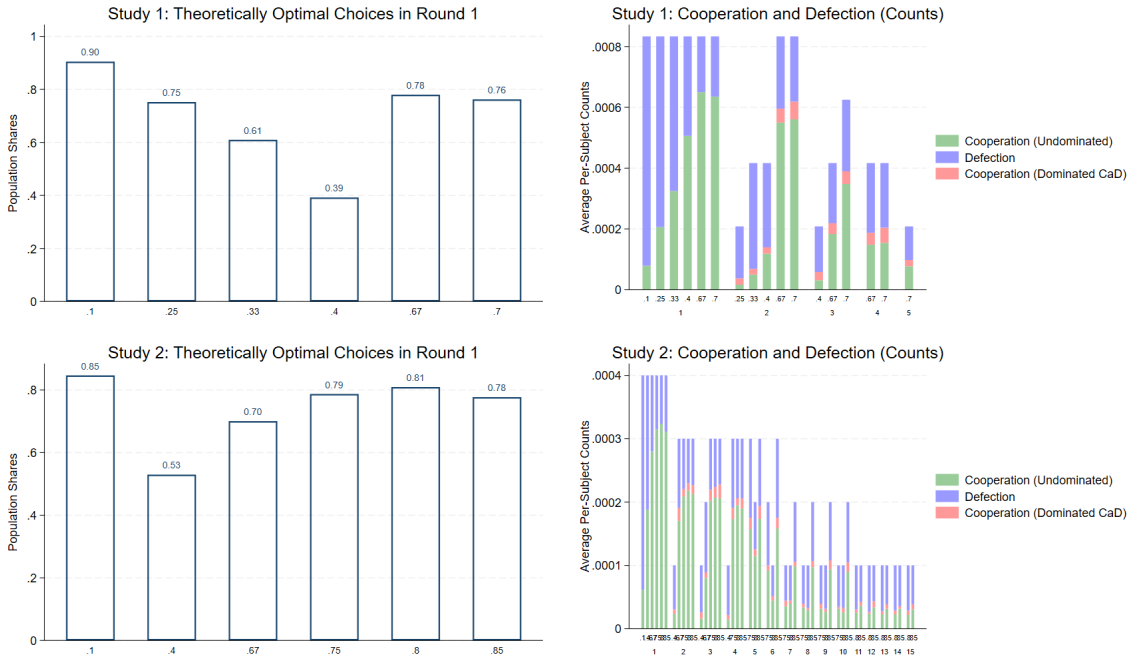


Figure 3: Patterns of optimality, cooperation and defection: Left panel: population shares of theoretically optimal choices in the first rounds of all supergames by the continuation probability δ (see also the leftmost set of bars in the right panel). Right panel: Average per-subject counts of cooperation 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 un-dominated cooperation and dominated *cooperation after defection* (CaD). The later rounds were never reached for some δ values. (see Tables C2 and D2). Top: Study 1 (N=100 subjects, 2,400 supergames); Bottom: Study 2 (N=101 subjects, 2,424 supergames).

Figure 3 (left panels) displays the overall frequency of first round optimal play by continuation probability δ. In both studies, first round optimal play is highest when δ is the lowest ($\delta = 0.1$) and most subjects optimally defect. It is also high when δ is relatively high ($\delta > \delta^* = 0.5$), and most subjects optimally cooperate. First-round optimal play is

lowest at intermediate δs, especially at $\delta = 0.4$ in both studies, where over (almost) half the subjects in Study 1 (Study 2) cooperated suboptimally, despite the δ value lying below the (unknown) threshold $\delta^* = 0.5$.

This excessive cooperation at $\delta = 0.4$ (as well as at the intermediate $\delta = 0.33$ in Study 1) cannot be attributed to risk aversion, since — as noted in Section 2.1 —- risk aversion should instead lead to excessive defection. Rather, these apparent mistakes might be better explained by the behavioral inattention theory of Section 2.2. In such binary stochastic choice models (as well as in simple Probit or Logit models), the rate of optimal behavior is lowest when the expected-payoff difference is smallest, near the critical value $\delta^* = 0.5$. This is *precisely* what we observe, at $\delta = 0.4$, which is the value closest to δ^*.

Yet, as shown in Appendix A, while the cost of first round mistakes is indeed lowest (0) when $\delta = \delta^*$, 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. But, as discussed earlier, Figure 3 (left panels) 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 values, so the costs of mistakes alone cannot fully explain these deviations from theoretical predictions.

The rates of theoretically optimal first-round choices in Figure 3 (left panels) are easily translatable into the first-round cooperation rates (corresponding to the leftmost bars in Figure 3, right panels). There, first-round cooperation rises (almost) monotonically with the continuation probability δ (confirmed later on in regressions of Table 4). First-round cooperation responds more strongly to δ here than in standard subject-versus-subject experiments.[17] In Study 1, first-round cooperation rises from 9.5% at $\delta = 0.1$ to 76.3% at $\delta = 0.7$,

[17]We compute two benchmarks from the probit estimates in (Dal Bó and Fréchette, 2018, p. 66, Table 4), which regress first-round cooperation on normalized payoffs and δ separately, by experience. At our payoffs and continuation probabilities, predicted first-round cooperation in the first supergame (the *inexperienced* benchmark) ranges from 45.4% ($\delta = 0.1$) to 55.9% ($\delta = 0.7$), a spread of 10.5 percentage points. In the 15th supergame (the *experienced* benchmark), it ranges from 15.8% to 62.1%, a spread of 46.3 points.

a spread of 66.8 points, well above both benchmarks. In Study 2, it rises from 15.4% at $\delta = 0.1$ to 70.0% at $\delta = 0.67$, a spread of 54.6 points, exceeding the experienced benchmark over a narrower δ range, and reaches 77.7% at $\delta = 0.85$.

Finding 2. *In both studies, during the first rounds of the supergames, (a) cooperation (defection) tends to increase (weakly decrease) with the continuation probability δ. However, cooperation is excessively high at the "intermediate" continuation probability $\delta = 0.4$. (b) Relative to the theoretical optimum, subjects behave suboptimally, cooperating too much at low continuation probabilities and too little at high continuation probabilities. Across the first rounds of all supergames, the median subject made just over two-thirds (Study 1) and three-quarters (Study 2) of the theoretically optimal decisions.*

4.3 Cooperation After Defection (CaD)

Our design allows us to examine behavior beyond the first rounds of the supergames. 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 decision. Thus, once a subject had defected, cooperating later in the *same* supergame – behavior we call cooperation after defection (CaD) – was dominated for any δ and therefore constituted a *strategic error*.

Figure 3 (right panels) shows the average per-subject numbers of cooperate and defect choices, by continuation probability δ, separately for the first round and subsequent rounds. Beginning in round 2, the figure distinguishes dominated cooperation after defection (CaD) from undominated cooperation. By chance, each study contained 13 supergames lasting more than one round. Among all post-first-round decisions in these supergames, CaD accounted for 7.6% of choices in Study 1 and 5.7% in Study 2.

While CaD errors account for a relatively small share of decisions, their distribution across subjects complicates the interpretation of subject-level departures from theoretically optimal behavior. Figure 4 (left panels) shows that 52% of subjects in Study 1 and 60.4%

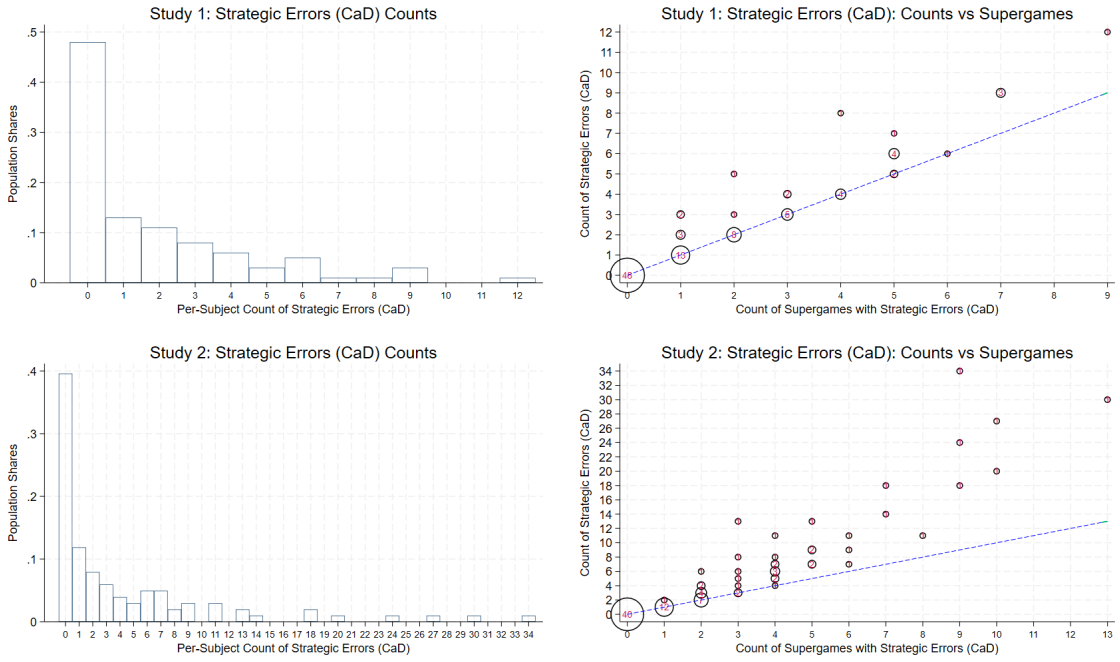


Figure 4: Strategic errors of dominated cooperation after defection (CaD) in the relevant supergames (i.e., those lasting longer than one round). Left: Distribution of per-subject counts of instances of cooperation after defection (CaD). Right: Per-subject counts of CaD instances vs. count of supergames with those instances. Bubble size is proportional to the share of subjects. Top: Study 1 (N=100 subjects, 13 relevant supergames); Bottom: Study 2 (N=101 subjects, 13 relevant supergames).

in Study 2 made at least one CaD error. Moreover, 20% and 34.7%, respectively, made at least four dominated CaD choices.

Some CaD choices could be intentional. For example, a subject might cooperate after defecting to verify the displayed information about the computer opponent's behavior. A CaD observation could also arise from a genuine "trembling-hand" error if a subject accidentally pressed the "defect" button without noticing this and subsequently returned to the cooperative action originally intended. In either case, an attentive, payoff-maximizing subject would be unlikely to make dominated CaD choices repeatedly or in multiple supergames.[18]

To assess whether CaD can be attributed solely to isolated trembling-hand errors, Figure 4 (right panels) compares each subject's total number of CaD choices (vertical axis) with the number of supergames in which the subject made at least one such choice (horizontal axis). Although the overall CaD rate is relatively low, CaD behavior is persistent for some

[18]Kloosterman (2020) also find that subjects in repeated human-versus-human prisoner's dilemma games sometimes return to cooperation after defection.

subjects: 24% of subjects in Study 1 and 33.7% in Study 2 made CaD errors in at least 3 of the 13 supergames lasting more than one round. This persistence suggests that some dominated CaD behavior may reflect inattention or limited understanding of the dynamic consequences of defection, even though subjects were required to pass a comprehension quiz before playing.

A comparison of the right panels of Figure 4 shows that, despite its lower overall frequency, CaD behavior was more dispersed across subjects and repeated choices in Study 2. Among subjects who made at least one CaD error, 63.5% in Study 1 made exactly one such choice in every affected supergame, as indicated by bubbles on the diagonal in the top-right panel. By contrast, 63.9% of the corresponding subjects in Study 2 lie above the diagonal, indicating that they made multiple CaD choices in at least one affected supergame. This more complex pattern of subject-level behavior in Study 2 may partly reflect its longer supergames, which provided more opportunities for trembling-hand errors: subjects made 99 choices in Study 2, compared with 48 in Study 1. Table 1 illustrates the persistence of this CaD behavior for subject 20207, who made CaD errors in ten supergames. Although the overall CaD rate accounts for a relatively small share of observations, its pattern complexity underscores the difficulty of interpreting departures from optimal behavior in standard IRPD experiments, which typically involve longer supergames and more choices.

Finding 3. *In Study 1 (2), over all rounds of all supergames, a majority of subjects — 52% (60.4%) — 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 7.6% (5.7%) of relevant observations, with 24% (33.7%) of subjects making this error in at least 3 of the 13 supergames lasting more than one round.*

4.4 Overall Point Totals

Recall that the theoretically optimal policy – cooperating in every round of every supergame when $\delta > 0.5$ and defecting in every round when $\delta < 0.5$ (see Hypothesis 1) – is derived *ex ante*, before the lengths of the supergames are realized. In Study 1, given the realized random supergame terminations, following the *ex ante* optimal policy would have yielded

Supergame	δ	Duration	Subject 20207 ($CRT7 = 0$)	Subject 20503 ($CRT7 = 5$)
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 1: Study 2: Examples of subjects engaging in non-stationary play. Subject 20207: dominated cooperation after defection (CaD) vs. subject 20503: end-timing (DaC). Both subjects faced the long order.

an *ex post* total of 4,050 points across all 48 decisions.[19] As Figure 5 (top) shows, subjects' overall point totals ranged from 3,285 to 4,185, with a mean of 3,835.05 (s.d. 203.38). We decompose each subject's total into a "fixed" component of 2,745 points, equal to the lowest *ex post* total that could have been earned given the realized supergame lengths,[20] and a "variable" component equal to the subject's earnings above that minimal, fixed amount. On average, subjects earned 83.5% of the variable component of the overall point totals that would have been obtained by following the theoretically optimal policy.

As Figure 5 (top) further reveals, some subjects earned as little as 41.4% of the "variable" component generated by following the optimal policy. Furthermore, 16% of subjects could have *increased* their overall point totals to 3,600 points (65.5% of the "variable" com-

[19]In Study 1, the theoretically optimal total of 4,050 points consists of 75 points in each round of the 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 $\delta \in \{0.1, 0.25, 0.33, 0.4\}$.

[20]In Study 1, the *ex post* theoretical minimum of 2,745 points is obtained as follows. In the 11 supergames lasting only one round, the lowest payoff is 75 points, obtained by cooperating (11 decisions). In each of the remaining supergames, the lowest total is obtained by defecting in the first round and earning 120 points (13 decisions), and then cooperating in every subsequent round and earning 15 points per round (24 decisions). Thus, attaining this minimum requires repeated cooperation after defection, although not every action used to construct the minimum is necessarily a strategic error.

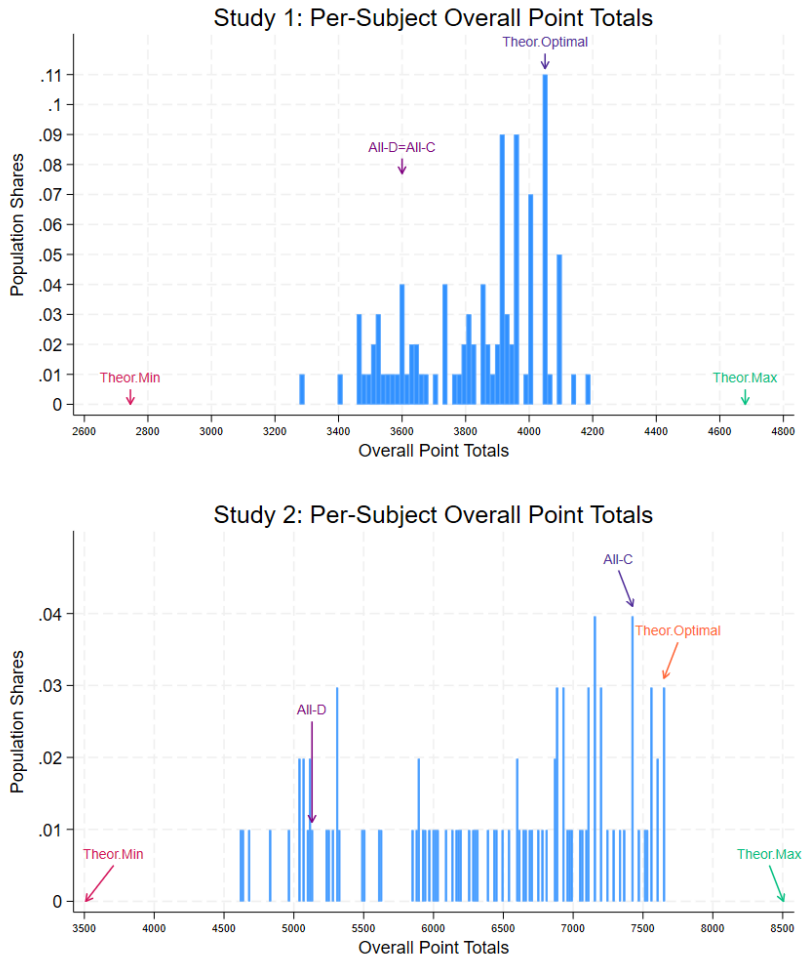


Figure 5: Distribution of overall point totals, or the sum of point earnings across all decisions. Top: Study 1 (N=100 subjects, 48 decisions); Bottom: Study 2 (N=101 subjects, 99 decisions). *Ex post* theoretical point total from following the *ex ante* optimal policy is 4,050 in Study 1 and 7,650 in Study 2; while the *ex post* theoretical minimum and (omniscient) maximum point totals are, respectively, 2,745 and 4,680 in Study 1; and 3,510 and 8,505 in Study 2.

ponent) by simply choosing either always to cooperate (All-C) or to always defect (All-D),[21] highlighting the payoff consequences of strategic errors (CaD).

As Figure 5 (top) further shows, the mode is at the theoretically optimal point total of 4,050. Yet, strikingly, the maximum point total is still higher! Here, 19% of subjects earned at least 4,050 points, the *ex post* total achievable by the *ex ante* theoretically optimal policy. However, as Figure 2 (top left) shows, only two subjects (2%) always followed this policy.

[21]Recall that, in Study 1, the expected payoffs to these two extremely biased strategies are the same by design. *Ex post*, by cooperating always, 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).

The other 17 subjects earned at least 4,050 points despite pursuing theoretically suboptimal strategies, – including those hypothesized in the next Section 4.5.

In Study 2, given the realized random supergame terminations, following the *ex ante* optimal policy would have yielded an *ex post* total of 7,650 points across all 99 decisions.[22] Figure 5 (bottom) shows that the earned point totals ranged from 4,620 to 7,650 – i.e., no subject earned more than the theoretically optimal-policy total. Subjects earned an average of 6,428.8 points (s.d. 867.7), or 70.5% of the "variable" component available above the *ex post* theoretical minimum of 3,510 points for Study 2. Moreover, 11.9% of subjects earned less than the All-D benchmark of 5,130 points, while 85.6% earned less than the All-C benchmark of 7,425 points. No subject earned more than the theoretically optimal total.

Finding 4. *In Study 1 (2), subjects' overall point totals are 83.5% (70.5%) of what following the* ex ante *optimal policy could achieve above the* ex post *theoretical minimum. Overall, 16% of subjects in Study 1 earned less than the points guaranteed by either All-C or All-D, and 11.9% in Study 2 earned less than the All-D benchmark. Notably, 19% of subjects in Study 1 reached the* ex post *optimal total points, but only two of these (2%) did so by following the optimal policy throughout; in Study 2, no subject exceeded the optimal total.*

4.5 "End-Timing" (DaC)

Against a Grim-trigger robot, a subject who knew exactly when each supergame would end could earn the highest feasible payoff by cooperating until the final round and then defecting. Subjects did not know the final round and therefore could not implement this "end-timing" strategy perfectly.[23] Nevertheless, some subjects appear to have attempted it, perhaps based on subjective beliefs about when the supergame would end.[24] This is a form of "gambler's

[22] In Study 2, the theoretically optimal ex post point total 7,650 is the sum of 75 points in each round of the supergames with a $\delta \in \{0.67, 0.75, 0.8, 0.85\}$ (88 decisions) plus 120 points in the first rounds (8 decisions) and 30 points in the subsequent rounds (3 decisions) of the supergames with a $\delta \in \{0.1, 0.4\}$.

[23] In both studies, the maximum realized totals (4,185 in Study 1 and 7,650 in Study 2) were well below the omniscient end-timing maxima (4,680 and 8,505, respectively), consistent with subjects having no advance information about termination.

[24] Suppose a subject believes that the continuation probability is initially δ but that the experimenter will end the supergame with probability one in round T. When $\delta > \delta^* = \frac{1}{2}$, the optimal strategy under this belief is to cooperate until round T and defect in that round.

fallacy": the belief that surviving earlier rounds raises the probability of termination in the current round, although the true termination probability remains constant (Cowan, 1969).

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]

Figure 6: 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. Top: Study 1 (N=100 subjects, 2,400 supergames); Bottom: Study 2 (N=101 subjects, 2,424 supergames).

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

Indeed, Figure 6 (left panels) shows that 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 6 (right panels) shows, the share of supergames where subjects always defected (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. Importantly, 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.

Note that the shorter supergames of Study 1 complicate interpretation there: subjects might have *intended* to time their defection, but the supergame ended earlier than expected. In contrast, Study 2 is biased toward higher continuation probabilities, so supergames last longer on average and display greater variance in duration, allowing us to more clearly detect when subjects used end-timing strategies. For example, subject 20503 in Table 1 appears to have "gambled" by defecting in what they believed was the final round for $\delta \in \{0.67, 0.75, 0.8\}$ – though, given the realized draws, this strategy proved *ex post* unsuccessful. As discussed earlier, such a strategy is less risky when continuation probabilities δ are relatively low (as in Study 1) than when they are high.

Furthermore, as the mixed-effects probit regressions in Table 4 (specifications 1 and 3) reveal, subjects' tendency to choose cooperation increases with δ, but *decreases* significantly with the round number, consistent with the use of the end-timing strategy. In contrast, if behavior were theoretically optimal, there should be no significant effect of the round number on the probability of cooperating – over and above the trembling hand effect.

Figure 7 (left panels) shows that only 22% (24.75%) of subjects in Study 1 (Study 2) never engaged in end-timing, i.e., never switched from cooperation to defection within the

[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 C2. 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.

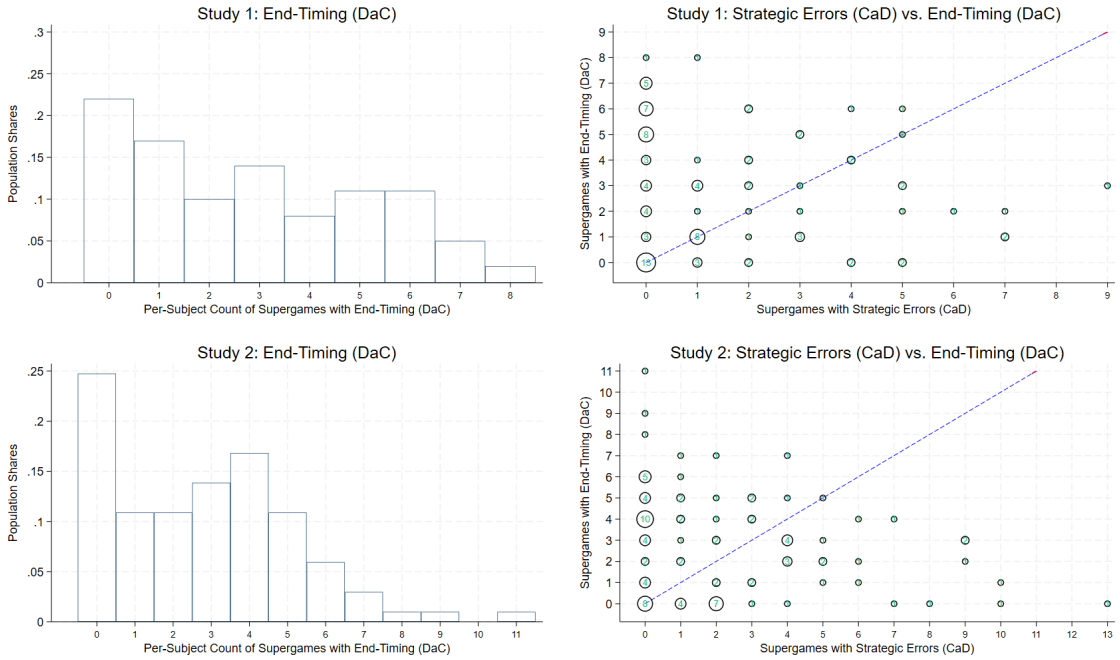


Figure 7: Left: Distribution of per-subject counts of supergames with "end-timing" (DaC) among the relevant supergames. Right: Per-subject counts of supergames with strategic errors (CaD) vs. those with end-timing (DaC). Bubble size is proportional to the share of subjects. Top: Study 1 (N=100 subjects, 13 relevant supergames); Bottom: Study 2 (N=101 subjects, 13 relevant supergames).

same supergame (DaC) – corresponding to the column atop 0 on the horizontal axis. As Figure 7 (right panels) shows, some apparent end-timing (DaC, on the vertical axis) may be unintentional, reflecting "mistakes" by subjects who also commit frequent strategic errors (CaD), shown by high values on the horizontal axis. Yet, some subjects who rarely commit strategic errors (CaD) (those near zero on the horizontal axis) appear to "end-time".

Finding 5. *Most subjects (78% in Study 1, 75% in Study 2) used, at least once, a non-optimal "end-timing" strategy – defecting after initially cooperating (DaC), – some possibly attempting to time their first defection to the unknown final round of a supergame. Following this risky strategy enabled a few subjects in Study 1, by chance, to earn more than the theoretically optimal payoff.*

4.6 Learning Over Time

In Study 1, as Table 2 (top) shows, subjects respond to their experience, changing their play in the second half of the experiment (last 12 supergames) relative to the first (first 12). While the incidence of dominated CaD errors declines over time, it does not disappear entirely in the second half of session, and the slight improvement in overall optimal play per round is only marginally significant. There is a clear decrease in supergames with dominated errors (CaD) but an increase in end-timing activity (DaC), without any significant increase in overall optimal behavior or in overall point totals.

Do subjects in Study 2 learn to play the theoretically optimal strategy when supergames with the same δ value are played consecutively in blocks? Table 2 (middle) shows that behavior changes from the first to the second half of the session, although the ordering of the δ-blocks complicates interpretation (see Figures C2 and D3). Subjects make fewer dominated CaD errors and cooperate more often, shifting away from both optimal and suboptimal All-D play toward All-C play. Because Study 2 favors cooperation, this shift increases the overall frequency of optimal choices, although CaD errors do not disappear.

Table 2 (bottom) considers learning *within these δ-blocks*, comparing the first two supergames (SG) with the last two at the same δ. Subjects make significantly fewer optimal decisions and use end-timing (DaC) significantly more often in the second half of a block. The shift appears in All-C play: optimal All-C play falls, while suboptimal All-C play rises, indicating that subjects increasingly cooperate where defection is optimal. The frequency of dominated CaD errors and overall point totals remain statistically unchanged within δ-blocks.

Finding 6. *In both studies, subjects commit fewer dominated CaD errors over time, though such errors never disappear. In Study 1, this does not lead to a significant gain in overall optimal play, as end-timing (DaC) becomes more frequent. In Study 2, overall optimal play does improve across the session, but within δ-blocks it* declines *as end-timing rises.*

Study 1: Learning: Within Session		1st Half (SG 1-12)		2nd Half (SG 13-24)		t-stat	df	pvalue
		Mean	StDev	Mean	StDev			
Per Round	Cooperate	0.52	0.50	0.48	0.50	-2.63	4798	0.01***
	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
	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
Study 2: 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*
Study 2: 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 2: The effect of learning: Means and standard deviations, and t-tests of differences between the two halves. Top panel (Study 1, N=100 subjects) and Middle panel (Study 2, N=101 subjects): the first (12) vs. second (12) supergames of each session. Bottom panel (Study 2, N=101 subjects): Within-δ-block comparisons (4 supergames per block), first vs. second half (2 supergames each). Df stands for degrees of freedom and pvalue denotes the two-sided probability $\Pr(|T| \geq |t|)$ under the null hypothesis. (Significance * 0.10 ** 0.05 *** 0.01 **** 0.001.)

4.7 Patterns of Play Within Each Supergame

In existing IRPD experiments, it is common to classify subjects' play using more than a dozen strategies, of which Grim and Tit-for-Tat (TFT) are the best known (Dal Bó and Fréchette, 2018, Section 2.5). Here, however, the computer opponent never defects first and never forgives, so these and more complex strategies cannot be identified, even though they could be optimal in repeated interactions among humans.

Instead, our two design innovations let us separate optimal from suboptimal stationary play (All-C and All-D, by conditioning on δ) and document the prevalence of two non-stationary suboptimal patterns that do not fit existing classifications: end-timing (DaC) and cooperation after defection (CaD). Here, subjects' within-supergame behavior falls into just four mutually exclusive patterns – All-C, All-D, CaD, and DaC – or six types once All-C and All-D are further split by whether they are optimal or not at the prevailing δ.[27]

As Figure 8 shows, there is no dominant pattern among subjects' play in either study. The two out of 100 subjects in Study 1 (top) and three out of 101 in Study 2 (bottom), who always made theoretically optimal choices, are represented by dark green and dark blue bars meeting at the solid line. In each study, two subjects always cooperated and one always defected. These appear as dark and light bars of the same color joined at the solid line: green at the far left for the two All-C players, blue at the far right for the All-D player. Both types of non-stationary 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.

Finding 7. *There is notable heterogeneity in subjects' choices to cooperate or defect, without any representative pattern. Only two (three) out of 100 (101) subjects in Study 1 (Study 2) always followed the theoretically optimal strategy. In each study, two subjects were fully biased toward cooperation and one was fully biased toward defection. The remaining subjects utilized suboptimal stationary strategies as well as suboptimal non-stationary strategies of end-timing (DaC) and of cooperation after defection (CaD).*

[27] 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.

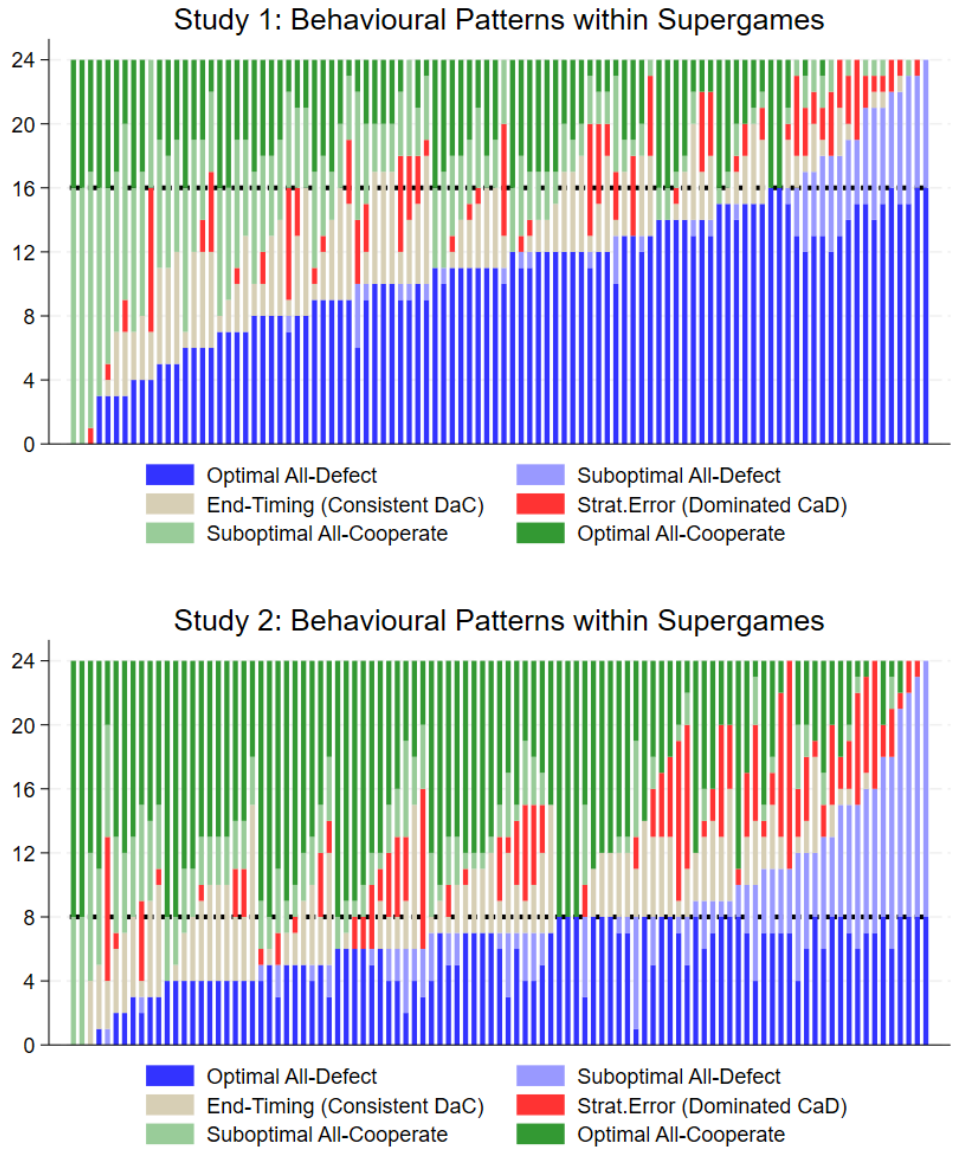


Figure 8: Subject heterogeneity in patterns of choice across the 24 supergames. Each column represents one subject; subjects are ordered by the count of supergames with All-Defect play (optimal and suboptimal combined). The theoretically optimal strategy, depicted by the solid horizontal line, involves always defecting in 16 (8) supergames and always cooperating in the remaining 8 (16) in Study 1 (Study 2). Top: Study 1 ($N = 100$ subjects); bottom: Study 2 ($N = 101$ subjects).

4.8 The Effect of Cognitive Abilities

Our hypothesis is that 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 (s.d.) CRT7 score was 3.78 (2.26) in

Study 1 and 3.61 (2.47) in Study 2, with a median of 4 in both studies (see Figure F5).

Study 1	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)	
CRT7	29.36***	23.78**	-0.56****	-0.44***	0.34*	0.25	0.14	0.17	0.50***	0.43**
	(8.66)	(9.63)	(0.16)	(0.15)	(0.19)	(0.18)	(0.14)	(0.14)	(0.18)	(0.19)
Order Long	55.68	62.97	-0.16	-0.21	0.76	0.88	0.39	0.33	0.95	1.03
	(38.35)	(37.93)	(0.69)	(0.65)	(0.79)	(0.79)	(0.60)	(0.59)	(0.76)	(0.77)
Female		-45.51		0.89		0.25		-0.86		-0.37
		(50.85)		(0.77)		(0.99)		(0.69)		(1.01)
Age		-4.98		0.12		0.25		-0.24*		0.08
		(9.01)		(0.16)		(0.18)		(0.14)		(0.18)
Patience		14.43		-0.28*		0.17		-0.11		0.13
		(10.52)		(0.16)		(0.17)		(0.15)		(0.20)
Retribution		-25.93**		0.16		-0.10		-0.23		-0.30
		(12.11)		(0.24)		(0.27)		(0.17)		(0.25)
Prediction		136.33**		-1.34		-0.87		0.63		-0.35
		(64.95)		(1.14)		(1.37)		(1.02)		(1.44)
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
F	7.07	3.64	6.28	2.01	2.34	0.81	0.92	1.37	4.83	1.51
p	0.00	0.00	0.00	0.03	0.10	0.64	0.40	0.19	0.01	0.14

Study 2	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)	
CRT7	114.82***	28.76	-0.71****	-0.43**	0.51**	0.07	0.16	0.00	0.71***	0.14
	(35.15)	(41.25)	(0.19)	(0.20)	(0.20)	(0.20)	(0.13)	(0.13)	(0.22)	(0.25)
Order Long	36.78	-43.06	-0.69	-0.36	-1.95**	-2.32***	1.18*	1.20**	-1.09	-1.33
	(165.03)	(150.15)	(0.85)	(0.77)	(0.91)	(0.81)	(0.60)	(0.55)	(1.03)	(0.89)
Female		-408.63**		2.26**		-1.87**		-1.90***		-3.66***
		(170.43)		(0.96)		(0.91)		(0.69)		(1.09)
Age		54.97*		-0.12		0.25		-0.04		0.21
		(28.32)		(0.18)		(0.17)		(0.11)		(0.19)
Lottery		62.31		-0.35		0.34		0.17		0.46*
		(47.62)		(0.23)		(0.23)		(0.18)		(0.26)
Patience		53.39		-0.18		0.27		0.08		0.32
		(46.85)		(0.25)		(0.23)		(0.17)		(0.25)
Retribution		74.74**		-0.15		0.47**		-0.31*		0.21
		(35.76)		(0.22)		(0.19)		(0.17)		(0.24)
Prediction		847.28***		-2.08		4.21***		-0.61		3.36*
		(299.31)		(1.29)		(1.49)		(1.00)		(1.70)
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
F	5.81	4.74	7.69	2.58	4.14	4.85	2.56	1.96	5.21	4.60
p	0.00	0.00	0.00	0.00	0.02	0.00	0.08	0.03	0.01	0.00

Table 3: Individual differences in rationality. Only the relationships which were significant in at least one laboratory study are presented. "Controls" stands for further demographic and individual differences controls. (Significance: * 0.10 ** 0.05 *** 0.01 **** 0.001). Top: Study 1 ($N = 100$ subjects); Bottom: Study 2 ($N = 101$ subjects).

In Study 1, as Table 3 (top) shows, consistent with Hypothesis 3 of Section 2.2, the CRT7 score is positively correlated with overall point totals (specifications 1-2) and negatively with the count of supergames involving strategic errors (dominated CaD) (specifications 3-4). The count of theoretically optimal supergames (i.e., either All-D for $\delta < 0.5$ or All-C for $\delta > 0.5$) is marginally positively correlated with CRT7 (specifications 5), but the overall fit, as indicated by the F statistic, is poor. The count of apparent end-timing behavior (DaC) is not correlated with the CRT7 score, or with any other characteristic (specifications 7-8). Yet, as discussed

in Section 4.5, our classification assigns each supergame to exactly one pattern, even though optimal play and end-timing can be observationally equivalent. When $\delta < 0.5$, immediately defecting is optimal – but it is also indistinguishable from pursuing an end-timing strategy. And when $\delta > 0.5$, always cooperating is optimal – but it is also indistinguishable from end-timing when a supergame ended earlier than a subject expected it.[28] To address this ambiguity, specifications 9 and 10 report the combined count of supergames classified as either theoretically optimal or end-timing. This *is* positively correlated with the CRT7 score, in contrast to the above negative correlation with dominated CaD errors.

In Study 2, as Table 3 (bottom) shows, consistent with Hypothesis 3, the CRT7 score is positively correlated with overall point totals (specification 1) and negatively with the count of supergames involving strategic errors (dominated CaD) (specifications 3-4). As in Study 1, the combined count of theoretically optimal play and end-timing is positively correlated with CRT7 (specification 9), although this disappears once controls are added (specification 10). CRT7 is uncorrelated with end-timing alone (specifications 7–8). The effect of the demographic controls is likely due to high correlation between CRT and gender ($\rho = -0.43, pvalue = 0.0000$).[29] The apparent order effect for end-timing behavior (DaC) in specifications 7-8 disappears in mixed-effects probit regressions of Tables 4 and 5.

Figure 9 visualizes how the patterns of behavior differ among the subjects according to how their CRT7 score compare to the median (equal to 4 in both studies). Low-CRT subjects have a greater number of strategic errors (colored red) and a lower number of optimal choices.

Finding 8. *In both studies, subjects with higher CRT7 scores earn higher payoffs and make fewer errors.*

[28]For example, as one can see in Table 1, the behavior of subject 20503 in Study 2 is consistent with pursuing an end-timing strategy in all 24 supergames, was, instead, classified here as behaving optimally in 10 supergames. Such behavior is more difficult to identify in the shorter supergames of Study 1.

[29]Study 1 was gender balanced – 48 males (with mean (s.d.) CRT of 4.33 (2.19) and median of 5) and 52 females (with mean (s.d.) CRT of 3.26 (2.22) and median of 3), and correlation between CRT and gender is $\rho = -0.24, pvalue = 0.0178$. In contrast, Study 2 involved 36 males (with mean (s.d.) CRT of 5.03 (1.92) and median of 5) and 65 females (with mean (s.d.) CRT of 2.83 (2.41) and median of 2).

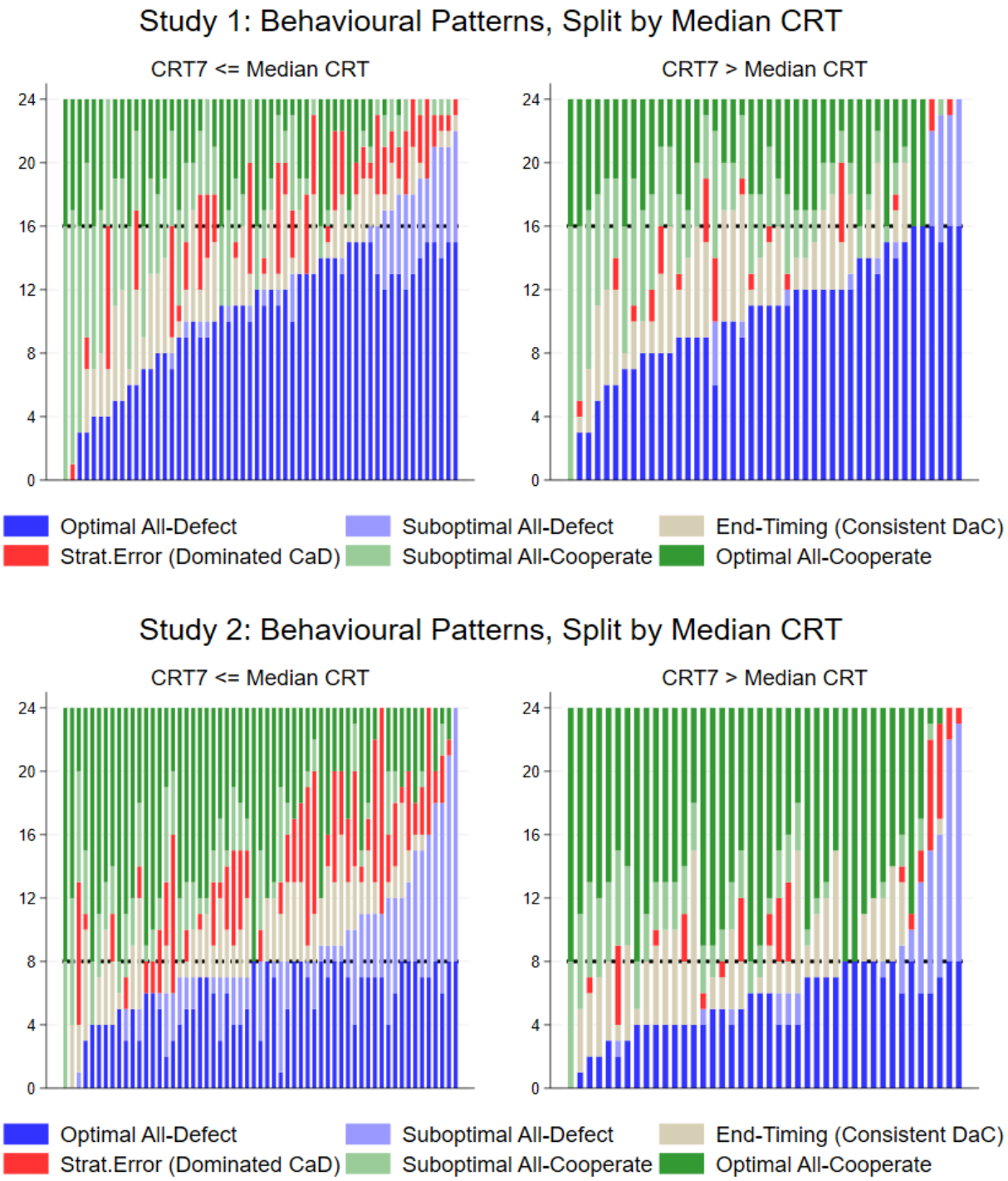


Figure 9: Cognitive Ability. Each subject's patterns of play within supergames (out of 24 supergames), split by median $CRT7$. Patterns are ordered by the count of supergames with combined optimal and sub-optimal All-Defect choices. The theoretically optimal strategy is depicted by solid horizontal line and it involves in Study 1 (2) always defecting in 16 (8) supergames and always cooperating in the remaining 8 (16). Each column represents the behavior of one subject. Top: Study 1 (N=100 subjects); Bottom: Study 2 (N=101 subjects).

4.9 Behavioral Inattention

Earlier in Section 2.2, we show that inattention theory implies Hypothesis 2 that cooperation rates should be increasing in the continuation probability δ. A further prediction, Hypothesis 3, is that, when deciding whether to cooperate or defect, individuals with higher cognitive ability (higher attention) respond more to actual payoffs, but lower cognitive ability individuals respond more to default values (Hypothesis 4).

Cooperate	Study 1		Study 2	
	(1)	(2)	(3)	(4)
Delta	4.18****	2.39****	3.18****	2.48****
	(0.31)	(0.52)	(0.20)	(0.35)
Round	-0.24****	-0.02	-0.04****	-0.01
	(0.05)	(0.10)	(0.01)	(0.01)
Supergame	-0.25***	-0.47***	0.64****	0.66***
	(0.09)	(0.15)	(0.12)	(0.21)
Prior Defection	-0.76****	-0.44**	-1.88****	-1.44****
	(0.11)	(0.21)	(0.12)	(0.18)
Order Long	0.19	0.18	0.18	0.18
	(0.17)	(0.18)	(0.13)	(0.13)
CRT7		-0.12**		-0.03
		(0.06)		(0.06)
Delta × CRT7		0.52****		0.20**
		(0.13)		(0.09)
Round × CRT7		-0.07***		-0.01**
		(0.02)		(0.00)
Supergame × CRT7		0.06		-0.00
		(0.04)		(0.05)
Prior Defect × CRT7		-0.09*		-0.14***
		(0.05)		(0.05)
chi2	225.03	270.51	408.50	450.90
p	0.00	0.00	0.00	0.00
N	4800.00	4800.00	9999.00	9999.00

Table 4: Behavioral Inattention: Choices to cooperate: mixed-effects probit regressions, coefficients, robust standard errors in parentheses. "Supergame" is the supergame number in the sequence of supergames (scaled down by 24),"Prior Defection" is a dummy variable for whether the subject defected in prior rounds of a given supergame and "Order Long" is a dummy variable for whether the first supergame in the sequence had $\delta = 0.67$. (Significance * 0.10 ** 0.05 *** 0.01 **** 0.001.)

To investigate these hypotheses, we report in Table 4 the results of probit regressions of the probability of cooperation on the continuation probability δ, our cognitive proxy, the CRT7 score, and interaction terms. One can see that in both studies cooperation is very strongly increasing in δ, supporting Hypothesis 2. Cooperation is also decreasing in round number, reflecting perhaps end-timing. The negative coefficient on Prior Defection reflects that indeed it is rational not to cooperate in a supergame where one has already defected. Notably, cooperation decreases with the supergame number in Study 1, which is biased toward defection by design, but increases with the supergame number in Study 2, which is

biased toward cooperation by design—a pattern consistent with learning in both studies.

Further, the interaction of cognitive score CRT7 with the continuation probability δ is significantly positive. Also, the interaction with "Prior Defect" is significantly negative, suggesting that higher cognitive ability subjects are less likely to commit a CaD error (which would decrease their payoff). These observations support Hypothesis 3 that higher cognitive ability subjects are more sensitive to the objective payoffs in the experiment, and are more attentive to the payoff consequences of their own prior actions. Interestingly, (and something that our inattention model cannot explain), the negative interaction with the round number suggests that higher-CRT7 subjects' cooperation declines more steeply within a supergame, a pattern consistent with more end-timing behavior.

Finding 9. *In both studies, subject behavior is broadly consistent with a simple model of inattention. (a) Consistent with Hypothesis 2, subjects' cooperative choices are strongly increasing in the continuation probability δ. (b) Consistent with Hypothesis 3, the sensitivity to the continuation probability δ increases with the subjects' cognitive scores.*

A less obvious prediction of the inattention model, Hypothesis 4, is that, when deciding whether to cooperate or defect, individuals with lower cognitive ability (lower attention) tend to rely more on default values – but not the individuals with higher cognitive ability (higher attention). To test this hypothesis, we split the sample by the median CRT7 score (equal to 4). Table 5 reports the results from mixed-effects probit regressions of the choice to cooperate or defect across all 48 (99) decisions in Study 1 (Study 2).

As the middle specification for each study shows, subjects with lower proxies for cognitive ability (CRT7$\leq$4) have a single factor correlated with their decision to cooperate in *both* studies. Specifically, 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)[30]. In contrast, subjects with higher proxies for cognitive ability (CRT7>4) show no such reliance on the Falk et al. (2018) measures; as shown below, their choices instead track the "Prediction" variable.

[30]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 B.4)

Cooperate	Study 1			Study 2		
	All	$CRT7 \leq 4$	$CRT7 > 4$	All	$CRT7 \leq 4$	$CRT7 > 4$
Delta	4.18****	3.38****	5.55****	3.18****	2.80****	3.83****
	(0.31)	(0.37)	(0.55)	(0.20)	(0.25)	(0.35)
Round	-0.25****	-0.15**	-0.42****	-0.04****	-0.02*	-0.07****
	(0.05)	(0.06)	(0.10)	(0.01)	(0.01)	(0.02)
Supergame	-0.25***	-0.32***	-0.16	0.64****	0.61****	0.74****
	(0.09)	(0.12)	(0.16)	(0.12)	(0.16)	(0.21)
Order Long	0.23	0.26	0.22	0.09	0.08	-0.17
	(0.15)	(0.21)	(0.22)	(0.12)	(0.16)	(0.20)
Prior Defect	-0.75****	-0.58****	-1.12****	-1.88****	-1.68****	-2.32****
	(0.11)	(0.14)	(0.21)	(0.12)	(0.14)	(0.21)
CRT7	-0.00			-0.01		
	(0.04)			(0.03)		
Female	-0.33*	-0.32	-0.35	0.02	0.16	-0.04
	(0.17)	(0.22)	(0.27)	(0.14)	(0.25)	(0.16)
Age	-0.03	-0.03	-0.05	0.05**	0.06	0.04
	(0.02)	(0.04)	(0.04)	(0.02)	(0.04)	(0.03)
Prediction	0.97***	0.72	1.09***	0.94***	0.57**	2.02****
	(0.35)	(0.53)	(0.36)	(0.29)	(0.28)	(0.59)
Lottery				0.02	0.01	0.01
				(0.05)	(0.05)	(0.07)
Risk	-0.01	-0.03	0.02	-0.00	0.00	-0.03
	(0.05)	(0.08)	(0.09)	(0.02)	(0.04)	(0.02)
Patience	0.06	0.14***	-0.09	0.05	0.11**	0.00
	(0.04)	(0.04)	(0.08)	(0.04)	(0.04)	(0.05)
Punishment	0.01	0.03	-0.03	-0.00	-0.01	-0.00
	(0.05)	(0.07)	(0.06)	(0.03)	(0.04)	(0.04)
Altruism	-0.06	-0.09	-0.09	-0.01	-0.04	0.01
	(0.05)	(0.06)	(0.08)	(0.04)	(0.03)	(0.05)
Reciprocity	0.07	0.04	0.25	0.02	0.00	0.08
	(0.07)	(0.08)	(0.18)	(0.04)	(0.04)	(0.08)
Retribution	-0.04	0.01	-0.10*	0.06**	0.05	0.05*
	(0.04)	(0.06)	(0.05)	(0.03)	(0.05)	(0.03)
Trust	0.01	-0.03	0.09*	0.00	-0.03	-0.00
	(0.03)	(0.04)	(0.05)	(0.02)	(0.04)	(0.03)
chi2	336.28	183.23	207.61	529.89	326.10	243.31
p	0.00	0.00	0.00	0.00	0.00	0.00
N	4800.00	2688.00	2112.00	9999.00	5841.00	4158.00

Table 5: Behavioral inattention: mixed-effects probit regressions of the choice to cooperate. Reported estimates are probit coefficients, with robust standard errors in parentheses. "Supergame" is the supergame number in the sequence, (scaled down by 24). "Order Long" is an indicator for whether the first supergame had $\delta = 0.67$. "Prior Defect" is an indicator for whether the subject defected in an earlier round of the same supergame. "Prediction" is the subject's predicted share of cooperative choices in round 1 across all 24 supergames, scaled by 100. The seven variables from "Risk" through "Trust" are self-reported preference measures from Falk et al. (2018). Significance levels: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$, and $^{****}p < 0.001$.

Furthermore, in both studies, the higher CRT7 group's cooperative choices are strongly correlated with the "Prediction" variable, elicited before any choices were made, possibly reflecting these subjects' understanding of the task – the structural response that Hypothesis 4 predicts in place of reliance on defaults.[31] In Study 2 only, there is some weaker relationship

[31]Before making any choices, subjects were asked what percentage of their Round 1 choices across all 24 supergames would be cooperative (see Section 3). For the theoretically optimal strategy, this "Prediction" variable is 33.33% (66.67%) in Study 1 (Study 2). The observed mean (s.d.) in Study 1 was 60.96 (29.57) with a median of 62. In Study 2 it was 59.70 (30.28) with a median of 67 – very close to the optimal choice of $\frac{2}{3}$. The Prediction variable is significantly correlated with subjects' actual first-round choices – $r = 0.3517$,

for lower CRT7 subjects as well – but this 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).

Furthermore, the relative size of the coefficients on the payoff-relevant decision variables is consistent with the above-mentioned findings – namely that the higher CRT7 subject group is less likely to cooperate after defection within the same supergame (i.e., make fewer CaD errors), and more frequently follows the end-timing strategy, as indicated by larger negative and significant coefficients on "Prior Defect" and "Round" variables. Notably, there is no robust, significant effect of either our proxy "Risk"[32] for risk propensity (elicited in both studies) or the incentivized Holt–Laury measure "Lottery" (elicited in Study 2 only).

Finding 10. *In both studies, behavior is broadly consistent with Hypothesis 4. For lower-CRT7 subjects (higher noise), the only elicited measure consistently correlated with cooperation is* Patience *– an* ex ante *disposition elicited before play and independent of the payoffs at stake, and thus a default in the sense of the model. For higher-CRT7 subjects (lower noise), Patience carries no such weight; their choices instead correlate with* Prediction*, a proxy for understanding the game's structure.*

5 Conclusion

We report an experiment that tests fundamental aspects of the standard game-theoretic model of repeated interactions. 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 or greatly reducing confounds such as strategic uncertainty, multiple equilbria, or social preferences. Subjects completed many supergames with varying continuation probabilities, allowing classification of within-supergame play into distinct behavioral patterns and linking deviations from the

$p = 0.0003$ in Study 1 and $r = 0.3879$, $p = 0.0001$ in Study 2.

[32]The Risk measure is also from Falk et al. (2018): "In general, how willing are you to take risks?" (see Appendix B.4).

theoretical optimum to individual characteristics, especially cognitive ability.

Our simplified individual-choice design pinpoints subjects' mistakes and shows how individual characteristics explain them. Not only did subjects make mistakes by playing the wrong action in the first round of supergames, but they also often seemed confused about the stationarity of the situation they were facing, making dynamic mistakes. A majority made at least one strategic error of cooperating after defection (CaD), and most subjects (78% in Study 1, 75% in Study 2) followed an end-timing strategy – defecting after initially cooperating (DaC) – at least once. The latter strategy can yield higher payoffs than the theoretical optimum yet is riskier for high continuation probabilities. These patterns correlate with cognitive ability, and differences between high- and low-ability subjects align with a simple model of inattention introduced here. 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!" *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.

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., N. Rulié, O. Stypa, and T. Werner (2025): "Delegate pricing decisions to an algorithm? experimental evidence," *arXiv preprint arXiv:2510.27636*.

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 Cost of First-Round Mistakes

Not all round-1 mistakes are equally costly. For each supergame, we calculate the *ex ante* expected payoff loss associated with the subject's observed round-1 action relative to the expected-payoff-maximizing action for the given continuation probability δ. To isolate the marginal payoff consequence of an incorrect round-1 action, we compare the value of the observed action with that of the benchmark action, assuming optimal play thereafter in each case and accounting for the Grim-trigger state induced by the round-1 choice. Let $a_1 \in \{C, D\}$ denote the subject's observed round-1 choice, and let $a_1^*(\delta)$ denote the benchmark choice. We focus on round 1 because it is observed for every subject in every supergame.

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 on either side of the threshold. If $\delta > \delta^*$, the benchmark prescribes cooperation. A round-1 defection activates Grim-trigger punishment for all subsequent rounds, 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 round; the optimal response is then 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, so $\mathcal{L}^{EA}(\delta^*) = 0$.

Table A1 reports the implied losses (in points) for the δ values used in Study and Study 2. 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 round, if players understand their mistake and correct it.

Table A1: 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
Study 1	36.0	22.5	15.3	9.0	46.36	60.0			
Study 2	36.0			9.0	46.36		90.0	135.0	210.0

‘

B Study 1: Experimental Instructions & Questions

This Appendix contains the instructions, quiz questions, belief elicitation, and post-experimental questions in the order in which subjects encountered these in Study 1.

B.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

B.2 Belief Elicitation ("Prediction")

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 (see Figure B1). Prior to making this choice, they were told that they would play 4 supergames for each of the 6 different delta values. After submitting this "Prediction" belief, the experiment proceeded on to the first indefinitely repeated PD game.

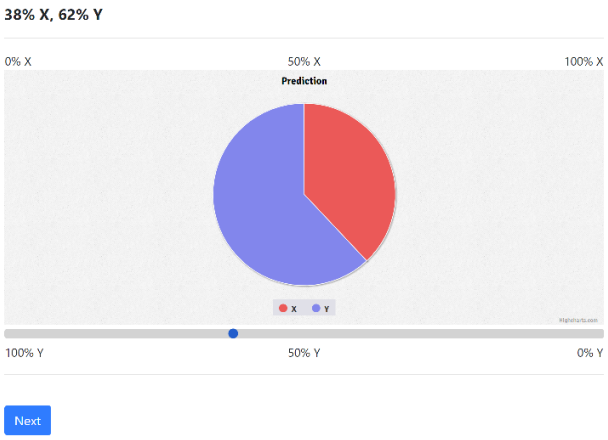
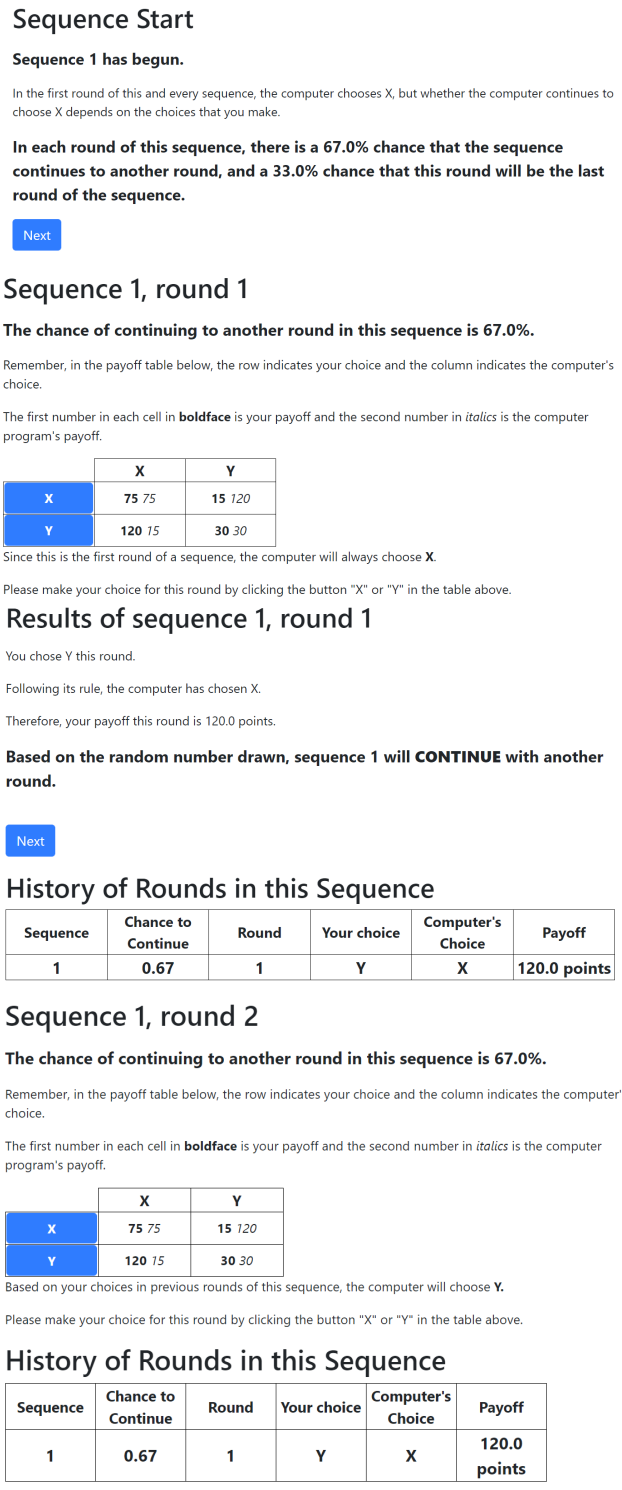


Figure B1:

B.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 B1-B2 provide an illustration of the screens that subjects faced in the first round and in continuation rounds of the first supergame of the "Order Long" treatment, which had a continuation probability of 0.67 and lasted for 4 rounds.



Sequence Start

Sequence 1 has begun.

In the first round of this and every sequence, the computer chooses X, but whether the computer continues to choose X depends on the choices that you make.

In each round of this sequence, there is a 67.0% chance that the sequence continues to another round, and a 33.0% chance that this round will be the last round of the sequence.

Next

Sequence 1, round 1

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*

Since this is the first round of a sequence, the computer will always choose **X**.

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

Results of sequence 1, round 1

You chose Y this round.

Following its rule, the computer has chosen X.

Therefore, your payoff this round is 120.0 points.

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

Next

History of Rounds in this Sequence

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

Sequence 1, round 2

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

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

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

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

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

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

History of Rounds in this Sequence

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

Figure B1: (1) Start screen for a new sequence. (2) Main decision screen for a round in the sequence. (3) Results screen for a round in the sequence. (4) Decision screen for a continuation round in the sequence, noting what the robot player will do, based on the history of play.

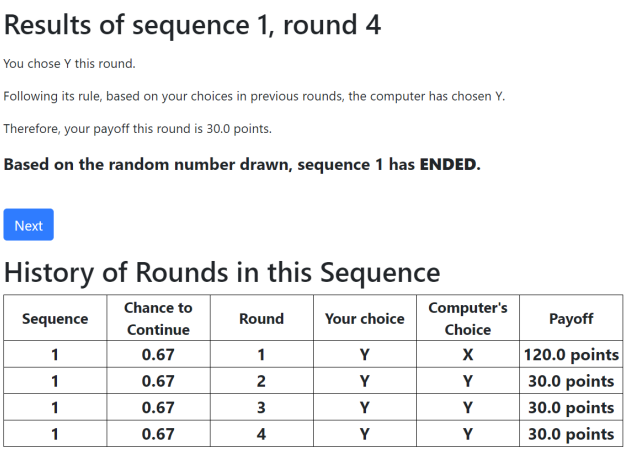


Figure B2: Screen for the final round of a sequence noting that based on the random drawn, the sequence has ended.

B.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.[33]

Questionnaire

We now ask for your willingness to act in a certain way in 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.

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

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

B.5 CRT questions

Subjects were asked to provide numerical answers to the following cognitive reflection test (CRT) questions.[34]

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?

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

C Study 1: Further Details

C.1 Study 1: Continuation Probabilities and Realizations

Tables C1 and C2 report on the continuation probabilities δ for each of the 24 supergames along with the actual number of rounds played for the two treatment orders of Study 1.

	OrderShort		Order Long	
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 C1: Study 1: 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	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 C2: Study 1: The distribution of the supergames, split by continuation probabilities δ. That is, out of 24 supergames, 11 lasted only 1 round, 7 only two rounds, and so on. The average theoretical and realized supergame durations are 1.99 rounds and 2 rounds, respectively.

C.2 Study 1: 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 Study 1, while the mean (s.d.) 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 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, being, respectively, 16.2 (3.49) and 17.42 (3.91), but this difference is only significant on the Kolmogorov-Smirnov test (one-sided $p = 0.034$), and only marginally on t-test ($t = 1.65, p = 0.051$). The overall optimal choice counts are, again, higher in the long order treatment, 32.54 (8.16) in long order, and 29.56 (7.76) in reverse, but this is marginally significant only on t-test ($t = 1.87, p = 0.032$), but not on Kolmogorov-Smirnov test (one-sided $p = 0.135$). Controlling for subjects' individual differences, the order effect is not discernible in mixed effects panel regressions in Tables 4 and 5.

Finding 11. *In the Study 1 sample, there is no consistent order effect.*

C.3 Study 1: More on Cooperation and Optimality

Figure C1 presents the marginal frequency distributions of choices to cooperate (bottom left panels) and of optimal choices (top right panels). The bottom right panels further provide two-dimensional distributions of cooperative and optimal choices, where the possible choice combinations are restricted to the polygons delineated by the dashed lines. The shape of the polygon in the bottom right panel for the overall all round choices (in the bottom set of panels) are due to the possibility of dominated CaD choices.

As the histograms in Figure C1 (top set of panels), subjects tend to excessively cooperate in the first round of each supergame. The mean (s.d.) count of cooperative choices per subject is 11.53 (5.73), far above the theoretical prediction of 8, with 16.81 (3.74) theoretically optimal choices per subject – prominently short of the theoretical prediction of 24.

As for all 48 choices in all 24 supergames, the mean (s.d.) of the overall count of cooperative choices is only 24.09 (11.24), significantly lower than the theoretical prediction of 26 (one-sided $t = 1.670, p = 0.046$). These cooperative choices are often sub-optimal – as depicted in Figure C1 (bottom set of panels) by the two-dimensional distribution in the bottom right panel. Indeed, with a mean (s.d.) of 31.05 (8.06), the overall optimal choice counts are significantly short of the theoretical prediction of 48. That is, the early excessive cooperation in the first rounds is followed by the subsequent defection. The mean (s.d.) of cooperation count in the subsequent rounds is only 12.56 (6.29), significantly lower than the theoretical prediction of 18 ($p = 0.000$) – despite the excessive cooperation of 1.82 counts per subject on average due to strategic (CaD) errors.

As Figure C1 shows, there is a significant heterogeneity in subjects' behavior, without any clear "representative" pattern. The initial heterogeneity of play in the first rounds (top set of panels) is further amplified by the heterogeneous strategies in the subsequent rounds. Only

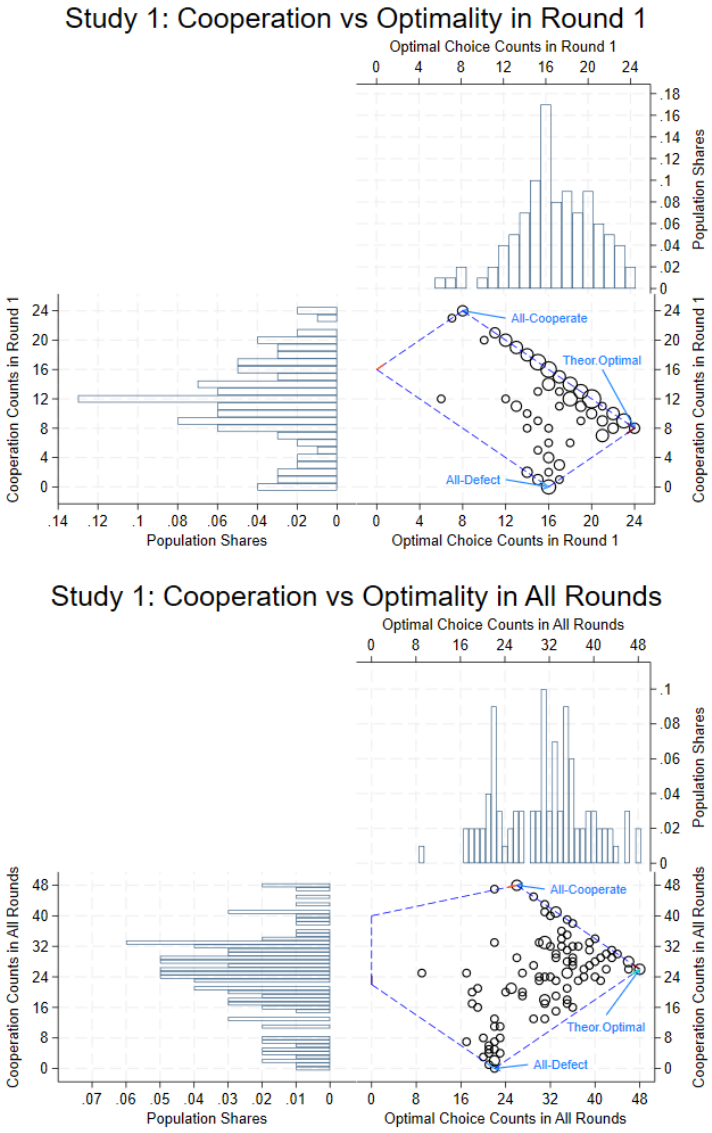


Figure C1: Study 1: Two-dimensional distributions of per-subject counts of cooperation and of optimal choices across all 24 supergames, together for the frequency distributions of cooperative (bottom left) and optimal (top right) choices. Top panel: Choices in the first rounds of each supergame. Bottom panel: choices in all rounds. Bubble size is proportional to the share of subjects (N=100 subjects).

two (out of 100) subjects made perfect theoretically optimal choices, in the far right corners of the bottom right polygon. The presence of strategic CaD errors complicates the interpretation of subject behavior, as most are located away from the boundaries. Many of those observations represent the overall early excessive cooperation in the first rounds followed by the subsequent defection within a supergame, possibly due "end-timing" strategies.

Finding 12. *In Study 1, compared with the theoretical predictions, on average, subjects cooperate too much at the start of supergames with* $\delta < 0.5$ *and stop cooperating too early in supergames with* $\delta > 0.5$*, with only 64.69% of all choices being theoretically optimal.*

C.4 Study 1: Learning

Figure C2 presents the patterns of intra-supergame play across all 24 sequences, split by the sequence order. By comparing behavior in the first few supergames to that in the last few, one can see an increase in the "end-timing" play (supported by the tests in Table 2). While the incidence of dominated CaD errors decline over time, they do not disappear entirely.

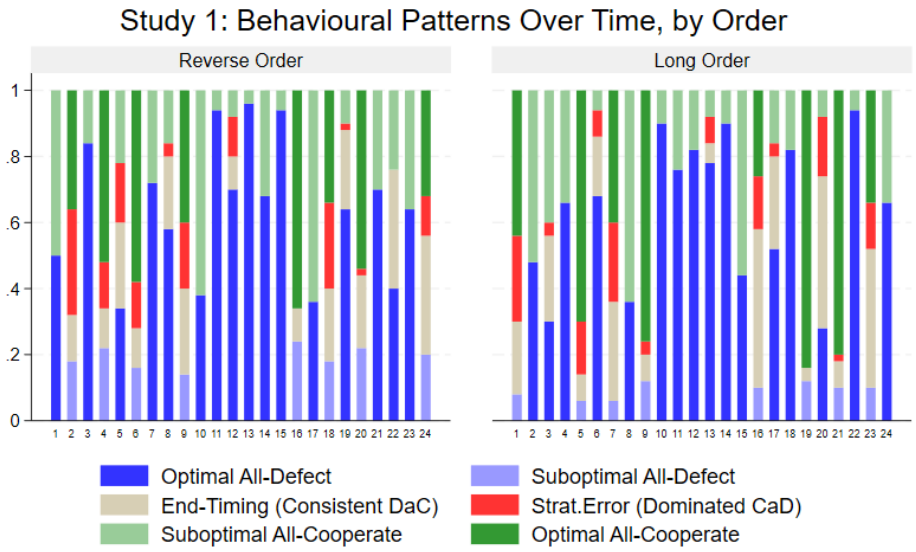


Figure C2: Study 1: Subject behavior over the sequence of 24 supergames, split by order treatment.

D Study 2: Further Details

D.1 Study 2: Experimental Instructions and Questions

The Repeated PD Game Instructions and Comprehension Quiz are essentially the same as for the original design Appendix B. 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)."

D.2 Study 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.

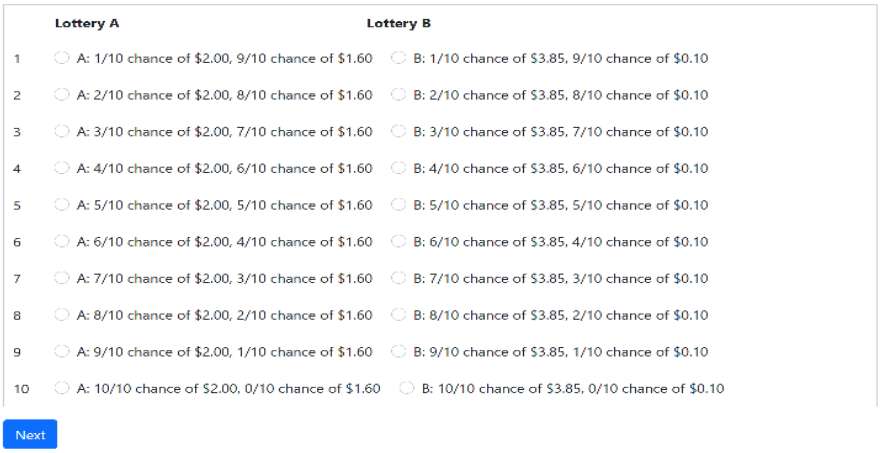


Figure D1: Study 2: Screenshot of Holt Laury Lottery Choice Task

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

D.3 Study 2: Continuation Probabilities and Realizations

Tables D1 and D2 report on the continuation probabilities δ for each of the 24 supergames along with the actual number of rounds played for the two treatment orders of Study 2.

	OrderShort		Order Long	
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 D1: Study 2: 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).

D.4 Study 2: 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 (s.d.) age was 20.53 (2.02) and 20.54 (2.30); the mean (s.d.) CRT7 was 3.63 (2.38) and 3.60 (2.59).

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 $\rightarrow$	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 D2: Study 2: 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.

The mean (s.d.) 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 mean (s.d.) overall counts (out of 99) is 59.63 (22.57) in the long order, higher 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 mean (s.d.) counts are 18.46 (4.79) in the reverse order, higher than 17.16 (4.36) in the long order, but the difference is insignificant ($t = 1.43$, Kolmogorov-Smirnov one-sided $p = 0.017$). Controlling for subjects' individual differences, the order effect is not discernible in mixed effects panel regressions in Tables 4 and 5.

Finding 13. *In Study 2, there is no consistent order effect.*

D.5 Study 2: More on Cooperation and Optimality

Analogously to Figure C1 for Study 1, Figure D2 (top set of panels) shows that subjects tend to cooperate insufficiently in the first rounds, with the mean (s.d.) 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 (s.d.) count of theoretically optimal first-round choices per subject is 17.80 (4.60), well short of the theoretical prediction of 24.

The mean (s.d.) of the overall count of cooperative choices for all 99 rounds is only 56.78 (24.26), again, short of the theoretical prediction of 88 - see Figure D2 (bottom set of panels). Again, cooperative choices are often sub-optimal due to CaD errors, and, with a mean (s.d.) of 57.22 (26.41), the overall optimal choice counts are prominently short of the theoretical prediction of 99.

Finding 14. *Study 2: On average, subjects cooperate less than theory predicted at the start of the supergames with $\delta > 0.5$, with 57.80% of all choices being theoretically optimal.*

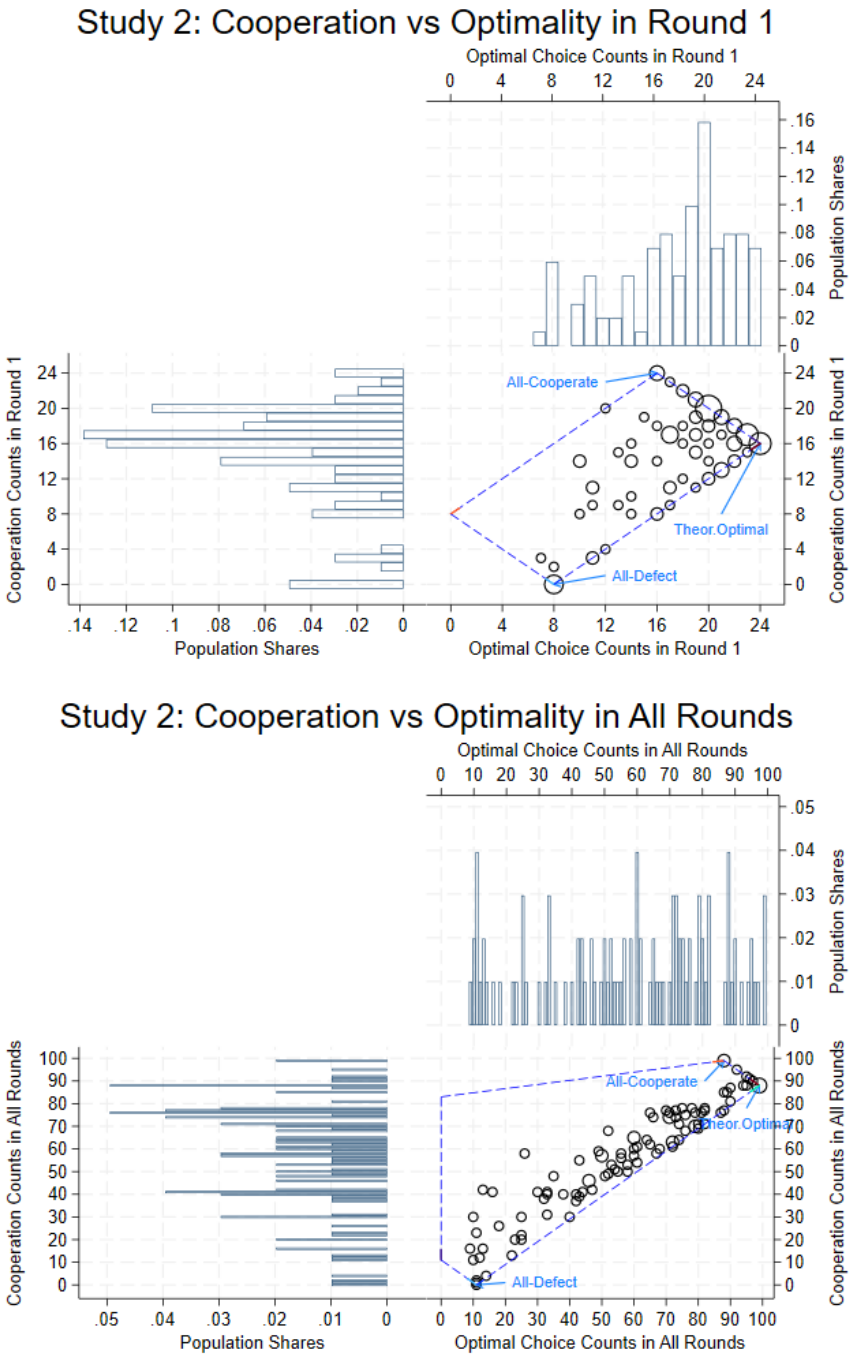


Figure D2: Study 2: Two-dimensional distributions of per-subject counts of cooperation and of optimal choices across all 24 supergames, together for the frequency distributions of cooperative (bottom left) and optimal (top right) choices. Top panel: Choices in the first rounds of each supergame. Bottom panel: choices in all 99 rounds. Bubble size is proportional to the share of subjects (N=101 subjects).

D.6 Study 2: Learning

In Figure D3, 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 2). Further, while the frequency of dominated CaD errors declines over time, such behavior does not disappear entirely. Figure D3 (bottom panel) further presents the patterns of intra-supergame play across all 24 sequences, split by the treatment order (long first or the reverse).

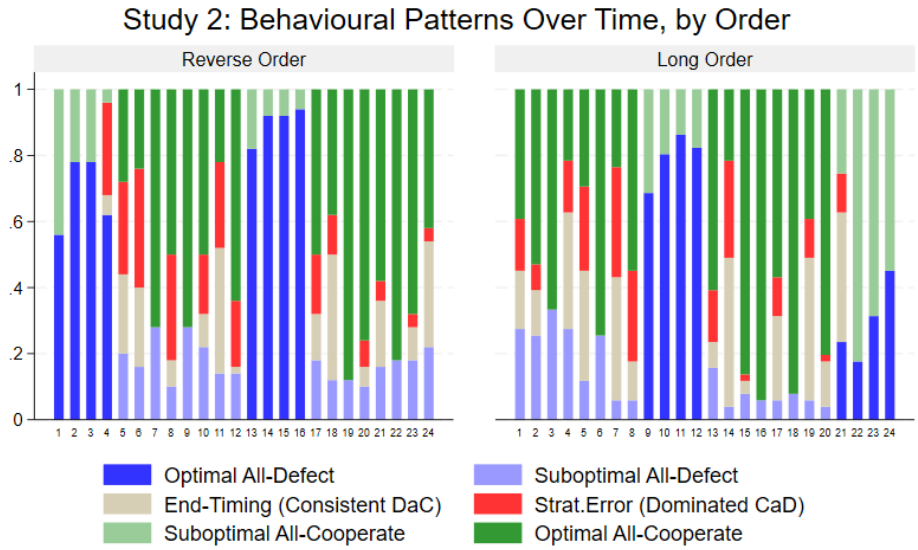


Figure D3: Study 2: Subject behavior over the sequence of 24 supergames, split by order treatment.

E Random (Study 1) vs. Blocks of Deltas (Study 2)

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 laboratory treatments of our experiment, three values $\delta \in \{0.1, 0.4, 0.67\}$ appear both in Study 1 in *random order* and in Study 2 *in blocks of 4 supergames*. As Figure E4 shows, there are differences between the random-order presentation of supergames (Study 1) and the blocked presentation (Study 2), 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 E3: for each $\delta \in \{0.1, 0.4, 0.67\}$, the rate of optimal choices under blocks is, respectively, lower, higher, and essentially unchanged relative to Study 1.

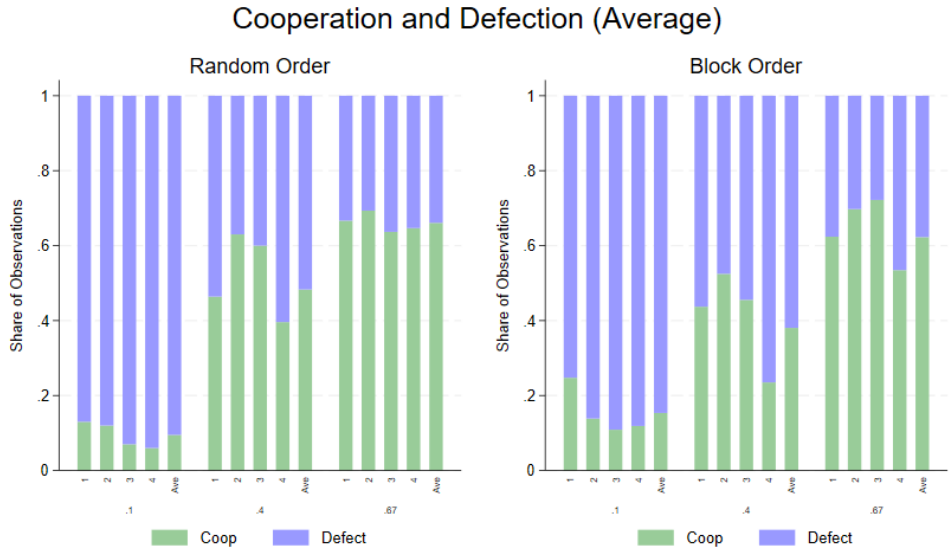


Figure E4: Study 1 vs Study 2: Cooperation and Defection rates across the supergames with $\delta \in \{0.1, 0.4, 0.67\}$ which appear both in Study 1 and Study 2. Left: Study 1 (N=100), right: Study 2 (N=101).

Interpreting error rates and payoffs is complicated by random realizations of supergame lengths (see Tables C2 and D2). For $\delta \in \{0.1, 0.4, 0.67\}$, the total number of decisions was $\{4, 7, 12\}$ in Study 1 and $\{4, 7, 9\}$ in Study 2. As a result, subjects earned higher *total* points for $\delta = 0.67$ in Study 1 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 Study 1 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 Study 1, compared to three

(respectively, one) in Study 2. Indeed, among supergames that lasted more than one round (see the final subsection of Table E3 for $\delta \in \{0.4, 0.67\}$), the ranking of normalized point totals at $\delta = 0.4$ is reversed. Although there are no significant differences between studies in optimal choices or dominated CaD behavior, end-timing is marginally more frequent in Study 1, where it happened, by chance, to be profitable overall. At $\delta = 0.67$, there are no significant differences between studies in any variable of interest among these longer supergames.

Random vs. Blocks		Random (Study 1)		Blocks (Study 2)		t-stat	df	pvalue
		Mean	StDev	Mean	StDev			
δ=0.1 (N=804)	Cooperate	0.095	0.294	0.153	0.361	-2.52	802	0.012**
Per Round	Optimal	0.905	0.294	0.847	0.361	2.52	802	0.012**
δ=0.1 (N=804)	Optimal (All-D+All-C)	0.905	0.294	0.847	0.361	2.52	802	0.012**
Per Supergame	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)	Cooperate	0.483	0.500	0.380	0.486	3.89	1405	0.000****
Per Round	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)	Optimal (All-D+All-C)	0.357	0.480	0.488	0.500	-3.76	802	0.000****
Per Supergame	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)	Optimal (All-D+All-C)	0.330	0.471	0.426	0.497	-1.63	299	0.10
Per Supergame	CaD per Round	0.040	0.126	0.059	0.128	-1.25	299	0.21
Lasting	DaC (End-Time)	0.335	0.473	0.228	0.421	1.93	299	0.06*
>1 round	Point Total per Round	75.41	12.40	58.96	10.37	11.46	299	0.000****
δ=0.67 (N=2109)	Cooperate	0.661	0.474	0.623	0.485	1.81	2107	0.07*
Per Round	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)	Optimal (All-D+All-C)	0.520	0.500	0.537	0.499	-0.49	802	0.63
Per Supergame	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)	Optimal (All-D+All-C)	0.520	0.500	0.469	0.500	1.34	701	0.18
Per Supergame	CaD per Round	0.043	0.127	0.036	0.121	0.65	701	0.51
Lasting	DaC (End-Time)	0.212	0.410	0.215	0.411	-0.07	701	0.95
>1 round	Point Total per Round	72.62	10.38	73.51	10.26	-1.14	701	0.25

Table E3: Study 1 vs Study 2: The effect of the order of supergames with the same δ value: Means, standard deviations, and t-tests comparing Study 1 (random order) and Study 2 (blocks), for each $\delta \in \{0.1, 0.4, 0.67\}$ which appeared both in Study 1 and Study 2. (Significance * 0.10 ** 0.05 *** 0.01 **** 0.001.)

Finding 15. *Study 1 vs Study 2: For each $\delta \in \{0.1, 0.4, 0.67\}$ that appears both in Study 1 and in Study 2, we find no systematic effect of presenting supergames in blocks with a common δ on either the frequency of optimal play or realized payoffs.*

F Behavioral Inattention: More Details

Figure F5 presents the distributions of the key variable for the inattention model, $CRT7$, and the Prediction variable, for each of our three studies.

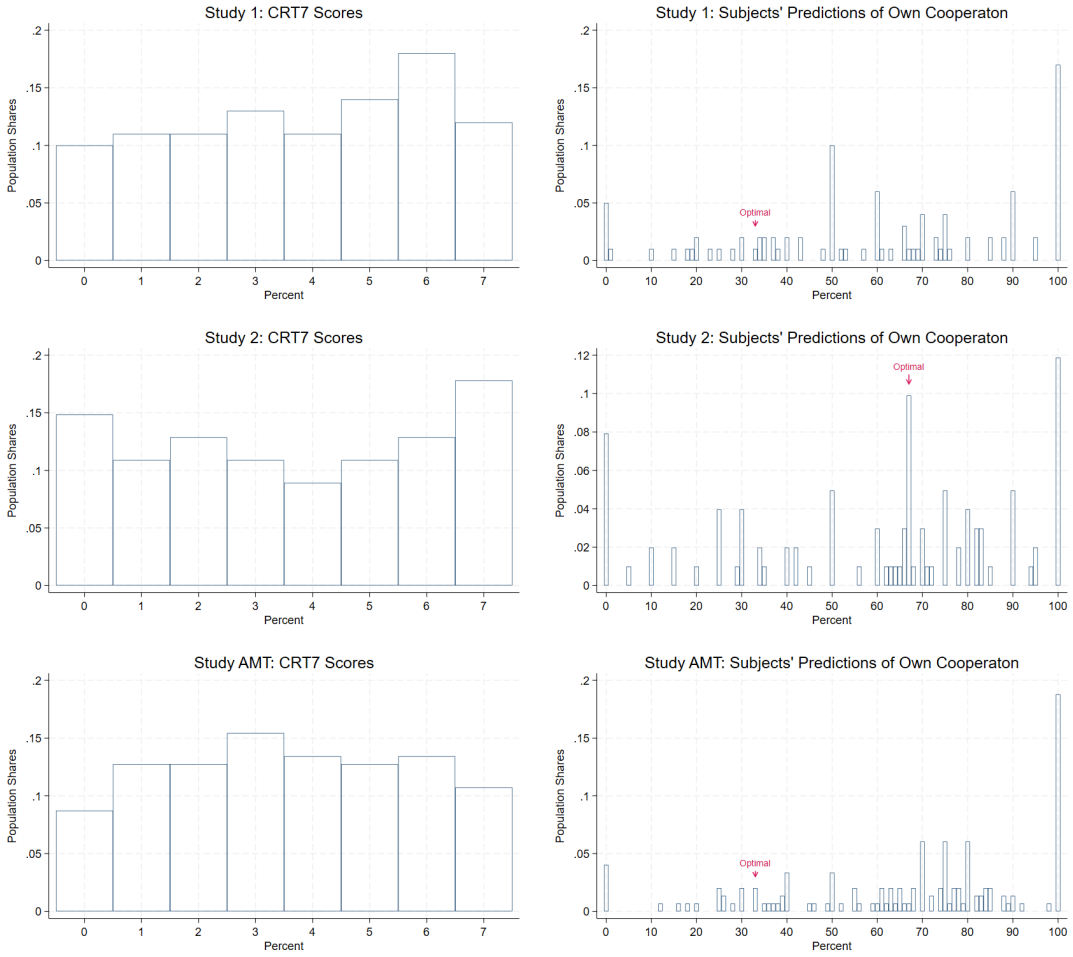


Figure F5: Frequency distributions of CRT7 scores (left) and of the "Prediction" variable (right). Top: Study 1; Middle: Study 2; Bottom: AMT.

G AMT Replication of Study 1

We conducted a replication of Study 1 using the same experimental program with 149 U.S.-resident Amazon Mechanical Turk (AMT) subjects; recruitment details and demographic comparisons with the Study 1 laboratory sample are reported in Table H7.

In Study 1 and its AMT replication, subjects were paid USD \$0.01 per point from six randomly selected supergames, one for each δ value, so both groups faced identical variable incentives. Fixed show-up payments differed – \$7 for Study 1 subjects and \$1 for AMT subjects – reflecting standard norms for laboratory versus online recruitment (see Rand (2012)). Total earnings averaged \$17.90 for a one-hour laboratory session[35] and \$10.37 for a one-hour AMT session. The order of supergames was balanced across subjects: nearly half of subjects (74/149) faced the "long" order (first supergame $\delta = 0.67$), and the remainder faced the reverse order.

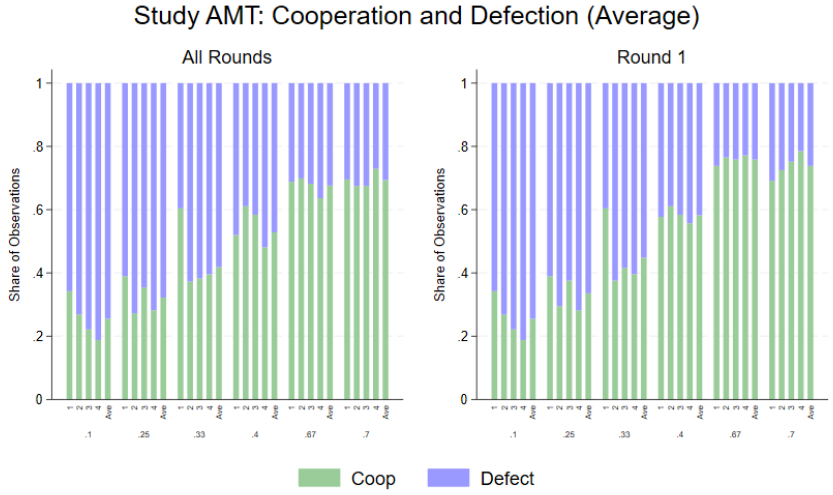


Figure G6: AMT: Cooperation and Defection rates in the four supergames in chronological order (first horizontal axis) of each of the six δ values (second horizontal axis). Left: across all rounds. Right: across first rounds only, (N=100 subjects.)

Figure G6 is the AMT replication of Figure 1 (top panels), confirming the earlier established aggregate patterns of deviations from standard game-theoretic predictions (Hypothesis 1). As expected, AMT subjects were further away from the theoretical optimum than Study 1 laboratory subjects (Kolmogorov–Smirnov one-sided test $D = 0.1799$, $p = 0.021$), especially in the supergames with $\delta < \delta^*$. (See also Figure G7 for more details.)

Finding 16. *The AMT sample displays similar deviations from the theoretical predictions to the laboratory Study 1 sample.*

Figure G6 (right) presents the average cooperation and defection rates across the *first round only* of supergames 1–4 for each δ. Again, AMT subjects are farther from the optimal

[35] After completing the 24 repeated PD games, Study 1 subjects (only) were randomly paired for another two-player task, not reported here, in which they could earn an additional \$1.00–1.70.

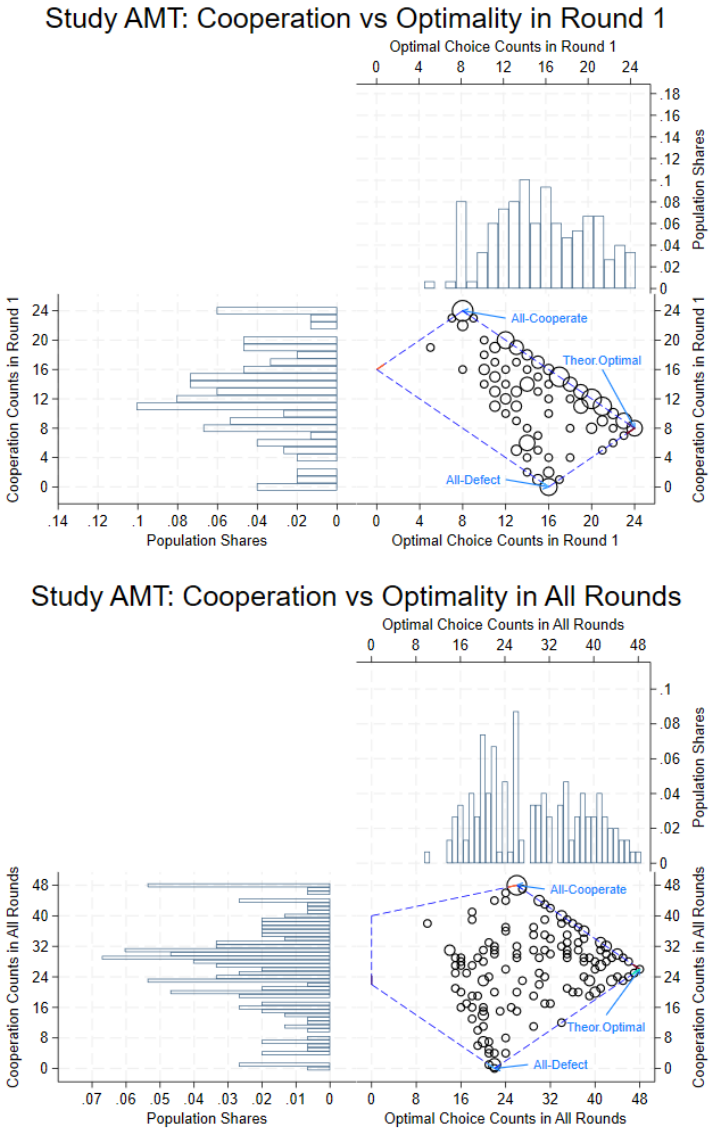


Figure G7: AMT: Two-dimensional distributions of per-subject counts of cooperation and of optimal choices across all 24 supergames, together for the frequency distributions of cooperative (bottom left) and optimal (top right) choices. Top panel: Choices in the first rounds of each supergame. Bottom panel: choices in all 48 rounds. Bubble size is proportional to the share of subjects (N=149 subjects).

policy than Study 1 subjects. As Figure G8 (right) shows, only 5 of 149 subjects (3.36%) in the AMT sample (compared to 2 of 100 subjects (2%) in Study 1) behave exactly in line with equilibrium predictions in the first rounds. Fewer than 17% of subjects in either sample made at most three mistakes (i.e., at least 21 optimal first-round choices out of 24, or 87.5%). As Figure G9 (left) shows, first-round cooperation rates range from 25.5% when $\delta = 0.1$ to 73.83% when $\delta = 0.7$.

Finding 17. *In the AMT sample, in the first rounds of supergames, subjects cooperated in excess of the theoretical optimum – on average by 55.9%. AMT subjects were significantly*

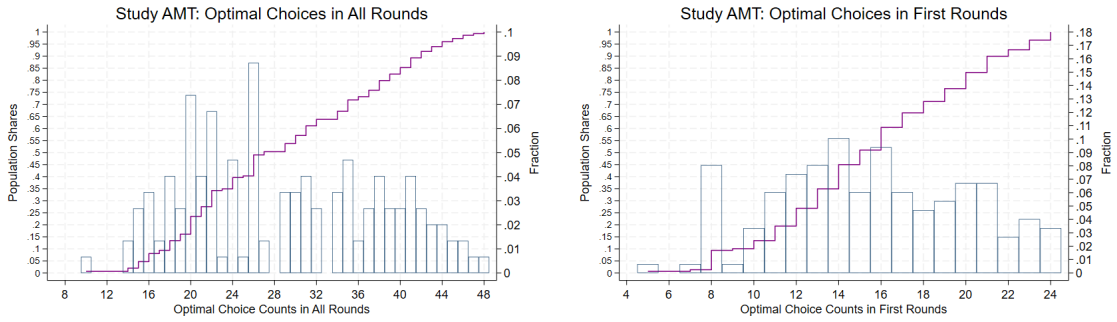


Figure G8: AMT: Frequency densities and cumulative distributions of per-subject counts of optimal choices (N=149). Left: across all supergames (48 decisions per subject). Right: in first rounds of each of 24 supergames (24 decisions per subject).

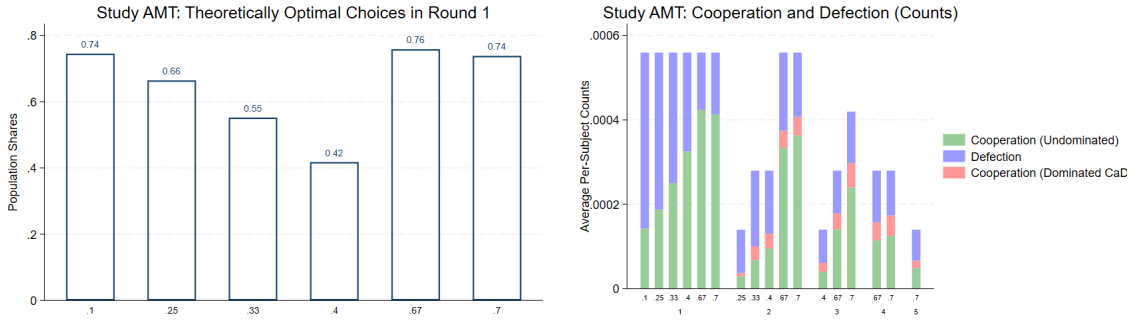


Figure G9: AMT: Patterns of optimality, cooperation and defection (N=149 subjects): Left panel: population shares of theoretically optimal choices in the first rounds of all supergames by the continuation probability δ (see also the leftmost set of bars in the right panel). Right panel: Average per-subject 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 C2).

less sensitive to changes in δ than Study 1 laboratory subjects.

Comparing the Figure G10 to Figure 4, the AMT sample displays greater rate of dominated errors of cooperation after defection (CaD). In the 13 supergames lasting more than one round, the share of CaD errors for AMT subjects is 11.47% (versus 7.57% in Study 1).

Finding 18. *In the AMT study, over all rounds of all supergames:*
(a) Cooperation (defection) increases (decreases) with the continuation probability δ.
(b) A majority of subjects (54.36%) made at least one strategic error by cooperating after previously defecting (CaD) within the same supergame, triggering a "grim" response.

Figure G11 reports on subjects' point totals, which is the theoretical measure of the payoff consequences resulting from subjects' behavior. As this figure shows, empirically overall point totals range from 3,195 to 4,185 for the AMT sample with a mean (s.d.) 3,766.21 (240.58), and some AMT subjects earned as little as 33.7% of the "variable" component

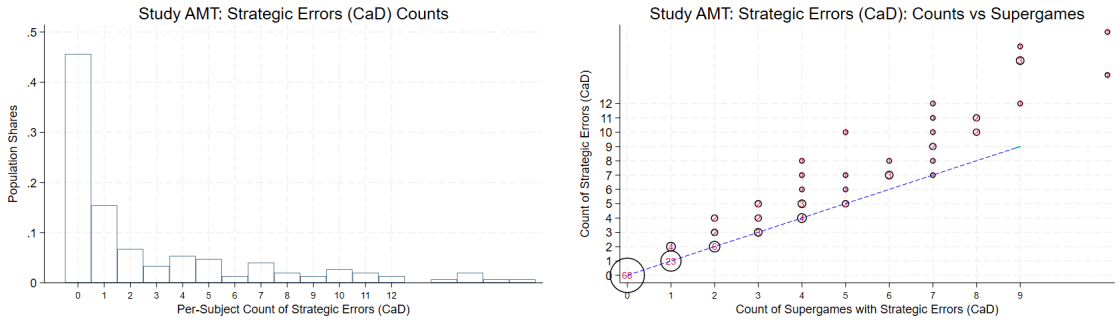


Figure G10: AMT: Strategic errors of dominated cooperation after defection (CaD) in 13 relevant supergames (i.e., those lasting longer than one round). Left: Distribution of per-subject counts of instances of cooperation after defection (CaD). Right: Per-subject counts of CaD instances vs. count of supergames with those instances. Bubble size is proportional to the share of subjects. (N=149 subjects).

achievable by following the optimal policy. Overall, 16.11% of the AMT subjects were able to achieve at least the theoretically optimal point value of 4,050, despite only one subject in the AMT sample actually behaving in a way that was fully theoretically optimal (in all 48 choices). No subject was able to get substantially close to the omniscient theoretical maximum, suggesting no transfer of information across AMT online participants.

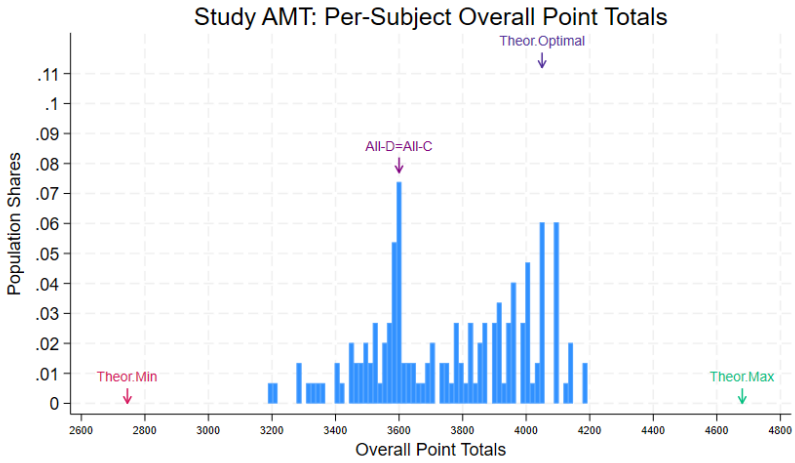


Figure G11: AMT: Distribution of overall point totals, or the sum of point earnings across all 48 decisions. *Ex post* theoretical point total from following the *ex ante* optimal policy is 4,050, while the ex post theoretical minimum and (omniscient) maximum point totals are 2,745 and 4,680, respectively.

Finding 19. *AMT subjects achieved 78.3% of what could be achieved relative to the* ex post *theoretical minimum by following the* ex ante *optimal policy.*

Figure G12 (left) shows that, just like Study 1 laboratory subjects, AMT subjects switched to defection having previously consistently cooperating (DaC), potentially engaging in "end-timing". Figure G13 (left) shows that only 30.9% of AMT subjects never engaged

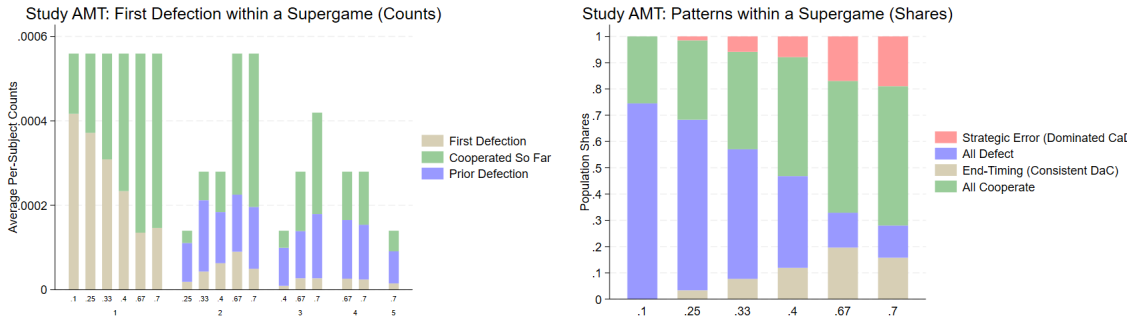


Figure G12: AMT: Patterns of cooperation and defection. Left: Average per-subject counts of first defection *within* a supergame by δ and round number. Right: The population shares of the behavioral patterns in a supergame, by δ value. By construction, the four strategies, CaD, DaC, All-D and All-C are mutually exclusive. (N=149 subjects.)

in end-timing, As Figure G13 (right) shows, some of this behavior could be unintentional "mistakes" by subjects who make frequent strategic errors (CaD), just like in Study 1, a few subjects who rarely commit strategic errors (CaD) (those near zero on the horizontal axis) appear to engage in end-timing behavior.

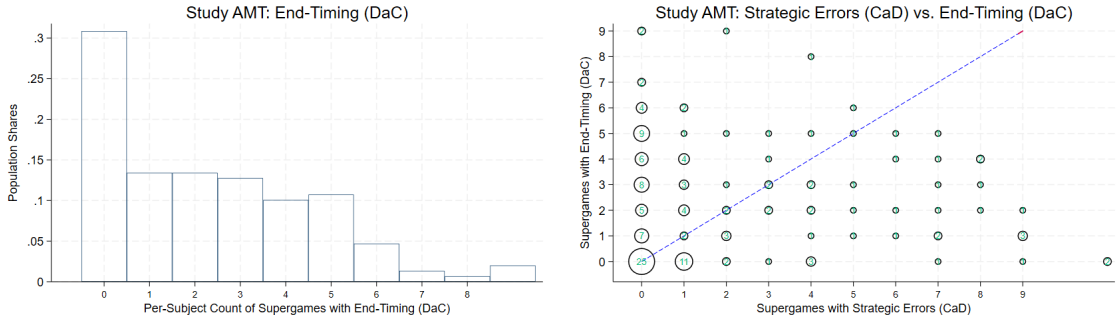


Figure G13: AMT: Left: Distribution of per-subject counts of supergames with "end-timing" (DaC) among 13 relevant supergames. Right: Per-subject counts of supergames with strategic errors (CaD) vs. those with end-timing (DaC). Bubble size is proportional to the share of subjects (N=149 subjects).

Finding 20. *Some AMT 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.*

As Table G4 shows, dominated CaD errors do not disappear entirely, comprising 7% for AMT subjects in the second half of play, and show a clear increase in end-timing activity (DaC). Overall, the improvement in optimal play per round, though significant, remains modest — rising from 59% in the first half to 62% in the second for AMT sample. Over time, only AMT subjects move toward theoretically optimal behavior within a supergame and earn slightly more points. As Figure G14 shows, again, while the incidence of dominated

CaD errors decline over time, they do not disappear entirely; and there is an increase in the "end-timing" play (supported by the tests in Table G4).

Learning		1st Half		2nd Half		t-stat	df	pvalue
		Mean	StDev	Mean	StDev			
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 G4: AMT: The effect of learning, first half (first 12 sequences) vs second half (last 12 sequences): Means, standard deviations, and t-tests of differences between the two halves. *df* stands for degrees of freedom (Satterthwaite's degrees of freedom for unequal variances of Age and Quiz Errors); *pvalue* stands for $Pr(|T| > |t|) = 0$. (Significance * 0.10 ** 0.05 *** 0.01 **** 0.001.)

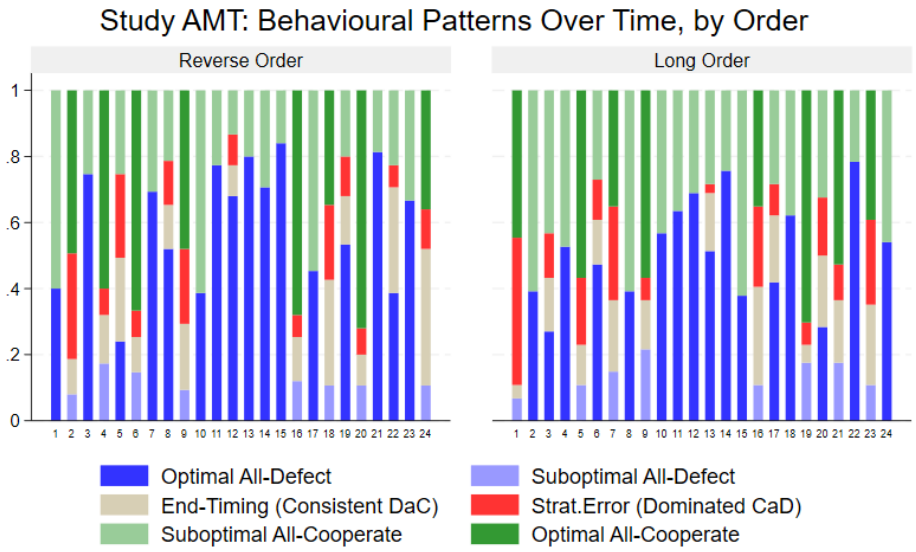


Figure G14: AMT: Subject behavior over the sequence of 24 supergames, split by order treatment.

Finding 21. *Over time, AMT 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.*

As Figure G15 shows, eight out of 149 AMT subjects always cooperated (compared to two out of 100 Study 1 laboratory subjects), a single subject always defected, and a single subject always made theoretically optimal choices.

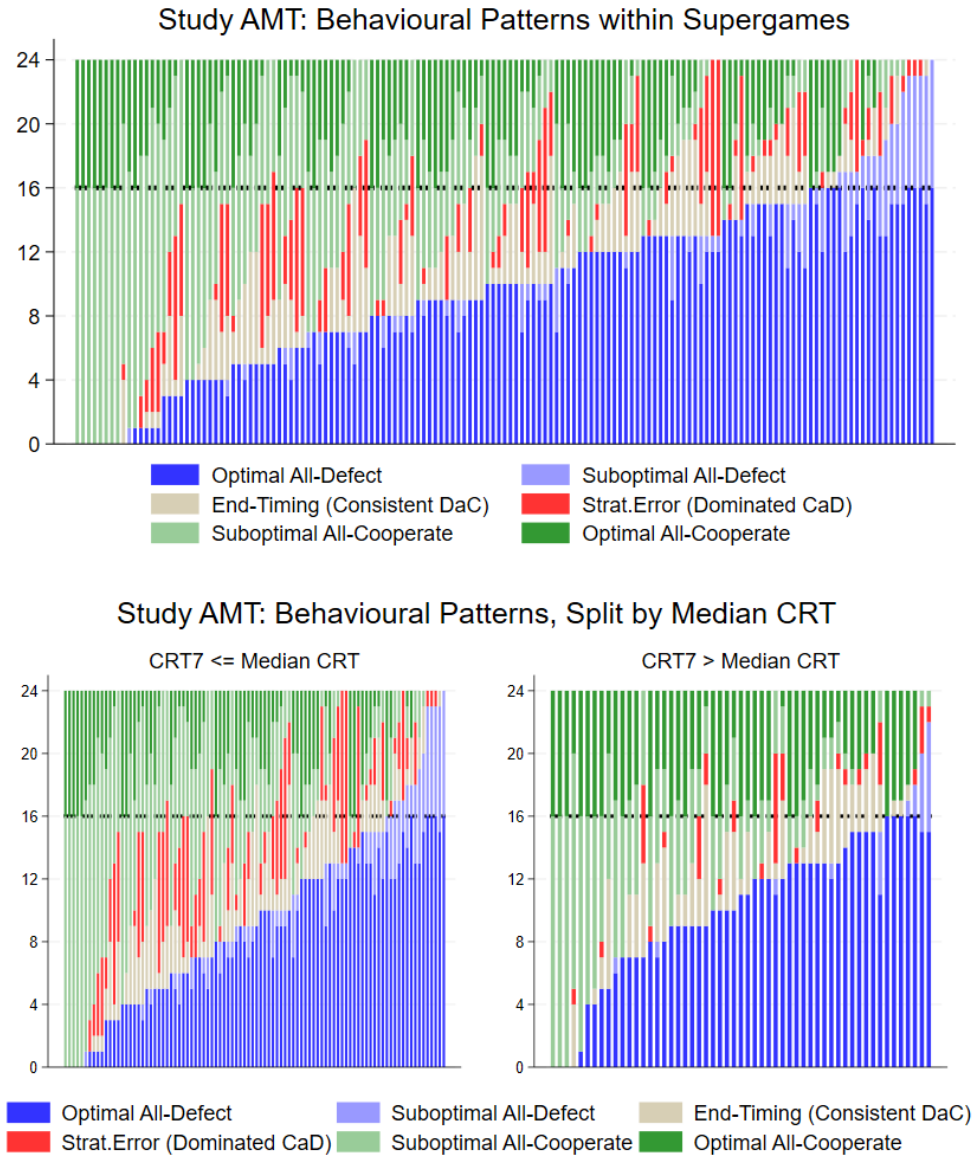


Figure G15: AMT: Subject heterogeneity in patterns of choices within 24 supergames, by subject, ordered by the count of supergames with (combined optimal and suboptimal) All-Defect choices (N=149 subjects). The theoretically optimal strategy involves always defecting in 16 supergames and always cooperating in the remaining 8 (solid red horizontal line). Top: Whole sample. Bottom: Inattention: each subjects' patterns of play within supergames (out of 24 supergames), split by median $CRT7$.

Finding 22. *Behavioural heterogeneity in AMT sample is similar to Study 1 sample.*

The mean (s.d.) CRT7 score was 3.58 (2.17) for AMT subjects, with a median of 4 as in other samples (see Figure F5, bottom left panel). There is no significant difference in CRT7 scores between the AMT and Study 1 laboratory subject pools (two-sided t-test = 0.72, $p = 0.48$). In contrast to Study 1 laboratory sample, where the CRT7 score is negatively correlated with being female ($r = -0.2365, p = 0.0178$), this is not the case for the AMT

subjects ($r = -0.0625, p = 0.4497$). As Table G5 shows, the CRT7 score predicts the rational aspects of subjects' behavior well. For AMT subjects only, the count of theoretically optimal choices (specification 6) is positively correlated with CRT7.

AMT	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)	
CRT7	62.79****	54.34****	-0.94****	-0.55****	1.01****	0.83****	0.02	0.11	1.08****	0.95****
	(6.87)	(7.53)	(0.17)	(0.16)	(0.16)	(0.16)	(0.12)	(0.12)	(0.16)	(0.17)
Order Long	-41.21	-23.87	-0.31	-0.48	-1.16	-0.81	-0.22	-0.34	-1.42*	-1.12
	(32.72)	(35.68)	(0.76)	(0.64)	(0.73)	(0.73)	(0.53)	(0.60)	(0.74)	(0.76)
Female		-84.97**		2.70****		-0.95		0.54		-0.68
		(35.87)		(0.61)		(0.70)		(0.53)		(0.75)
Age		-0.81		0.02		0.05		-0.05*		0.02
		(1.56)		(0.03)		(0.04)		(0.03)		(0.04)
Risk		-9.33		0.34**		-0.02		0.03		-0.05
		(8.51)		(0.16)		(0.18)		(0.12)		(0.18)
Patience		13.76		-0.38*		0.07		0.03		0.11
		(10.59)		(0.21)		(0.20)		(0.18)		(0.22)
Punishment		3.12		0.53****		0.07		0.08		0.11
		(8.46)		(0.14)		(0.15)		(0.13)		(0.17)
Altruism		-1.23		-0.10		-0.01		-0.02		-0.03
		(9.36)		(0.18)		(0.17)		(0.16)		(0.18)
Reciprocity		11.11		-0.19		0.57***		-0.10		0.51**
		(10.51)		(0.24)		(0.20)		(0.17)		(0.20)
Retribution		-4.51		-0.30**		-0.08		-0.06		-0.09
		(8.80)		(0.15)		(0.15)		(0.13)		(0.17)
Trust		-10.13		0.48***		-0.39**		0.09		-0.32**
		(7.51)		(0.16)		(0.16)		(0.12)		(0.16)
Prediction		7.09		-1.61		-1.77		-1.57*		-2.91**
		(54.30)		(1.13)		(1.42)		(0.93)		(1.39)
F	46.24	9.51	17.50	7.45	26.83	7.89	0.13	0.75	29.41	8.26
p	0.00	0.00	0.00	0.00	0.00	0.00	0.88	0.70	0.00	0.00

Table G5: AMT: Individual differences in rationality (N=149 subjects). All control variables except for "Altruism" were significant in at least one specification, and thus are reported. Coefficients on Constant are not reported. One subject excluded from regression with controls due to self-reported non-binary gender. (Significance: * 0.10 ** 0.05 *** 0.01 **** 0.001).

Finding 23. *In AMT sample, 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.*

As Table G6 shows, low and high CRT subject groups in AMT sample behave differently (with the regression results summarized visually in Figure G15). In the lower CRT7 group, those AMT subjects with higher self-reported *Patience* tend to cooperate marginally more frequently. There was no difference in Patience between the two CRT7 groups of the AMT subjects (two-sided $t = 0.2802$, $p = 0.7798$). In the lower CRT7 group, being female was significantly negatively correlated with cooperation for the AMT subjects only (but not for Study 1 laboratory subjects), and risk-prone AMT subjects cooperated marginally less frequently (but no effect for Study 1 subjects). This also contrasts with the theoretical prediction in Section 2.1 that risk-averse (risk-prone) subjects would cooperate less (more).

In the higher CRT7 group, Round 1 cooperation was significantly negatively correlated with Punishment for AMT (but not Study 1) subjects, and Retribution had no effect (yet it was significantly negative for Study 1 subjects). But it was strongly correlated with the

"Prediction" ($r = 0.2350$, $p = 0.0039$ for AMT, less than for Lab: $r = 0.3517$, $p = 0.0003$.) The mean (s.d.) of this "Prediction" was 66.64 (27.04) with a median of 72 for AMT subjects, with no significant difference between the two pools (two-sided t-test = 1.56, $p = 0.12$, see Figure F5, bottom right panel).

AMT	All			$CRT7 \leq 4$	$CRT7 > 4$
Delta	2.86****	-0.13	2.84****	1.81****	5.51****
	(0.24)	(0.28)	(0.24)	(0.25)	(0.54)
Round	-0.04	0.13**	-0.05	0.01	-0.28****
	(0.04)	(0.06)	(0.04)	(0.04)	(0.08)
Supergame	-0.23***	0.07	-0.23***	-0.15	-0.46***
	(0.09)	(0.17)	(0.09)	(0.11)	(0.15)
Order Long	0.12	0.17	0.13	0.22	0.17
	(0.14)	(0.16)	(0.16)	(0.21)	(0.31)
Prior Defect	-0.78****	-0.26	-0.76****	-0.59****	-1.03****
	(0.11)	(0.19)	(0.11)	(0.14)	(0.18)
CRT7		-0.21****	0.03		
		(0.05)	(0.03)		
Delta × CRT7		0.95****			
		(0.09)			
Round × CRT7		-0.07****			
		(0.02)			
Supergame × CRT7		-0.10**			
		(0.04)			
Prior Defect × CRT7		-0.13***			
		(0.05)			
Female			-0.31**	-0.39**	-0.25
			(0.15)	(0.19)	(0.25)
Age			-0.01	-0.01	0.01
			(0.01)	(0.01)	(0.01)
Prediction			0.72***	0.27	1.22**
			(0.24)	(0.30)	(0.48)
Risk			-0.05	-0.09*	-0.02
			(0.04)	(0.05)	(0.08)
Patience			0.03	0.11*	-0.05
			(0.05)	(0.07)	(0.10)
Punishment			-0.02	0.02	-0.11*
			(0.04)	(0.04)	(0.06)
Altruism			-0.02	-0.04	-0.01
			(0.03)	(0.04)	(0.07)
Reciprocity			-0.04	-0.06	0.01
			(0.03)	(0.05)	(0.06)
Retribution			0.05	0.04	0.06
			(0.03)	(0.05)	(0.06)
Trust			0.02	0.00	0.02
			(0.03)	(0.04)	(0.05)
chi2	172.89	217.56	194.45	85.45	170.42
p	0.00	0.00	0.00	0.00	0.00
N	7152.00	7152.00	7104.00	4512.00	2592.00

Table G6: AMT: Choices to cooperate: mixed-effects probit regressions, coefficients, with robust standard errors in parentheses. "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 Defect" is a dummy variable for whether the subject defected in prior rounds of a given supergame, "Prediction" is the subject's prediction of the share of own cooperative choices in Round 1 across all 24 supergames (scaled down by 100). The last 6 variables ("Risk"-"Trust") are self-reported preferences measures of Falk et al. (2018). (Significance * 0.10 ** 0.05 *** 0.01 **** 0.001.)

Finding 24. *As in the Study 1 laboratory sample, AMT subject behavior is broadly consistent with a simple model of inattention.*

H Comparison of Study 1 and AMT Subjects

AMT replication participants are farther from the optimal policy than Study 1 participants, especially in the supergames with $\delta < \delta^*$. As one would expect, Study 1 laboratory 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 H16.

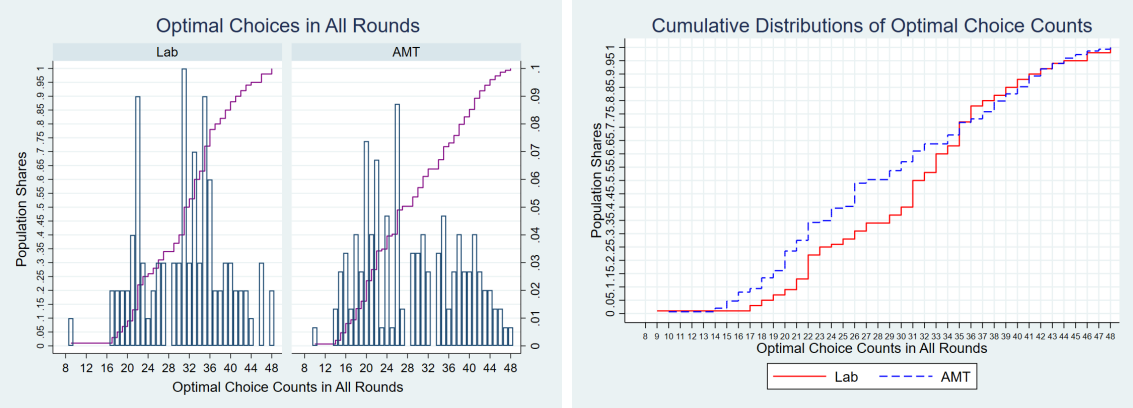


Figure H16: Study 1 (N=100) vs. AMT (N=149) subjects. Left: Frequency distributions and kernel densities of per-subject counts of optimal choices across all supergames (48 decisions per subject). Right: Cumulative distributions of optimal choice counts for both the Study 1 and AMT samples in the same graph.

The regression results of Tables 4, 5, and G6 are summarized visually in corresponding Figures 9 and G15, showing that the two subject groups differ in their patterns of play. Subjects in the lower CRT7 group 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 are closer to the theoretically optimal policy and engage in end-timing behavior (DaC) more often.

Figure G15 further reveals differences in patterns of behavior between the two subject pools (see also Table H7). 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 Study 1 subjects are less prone to pro-social behavior 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 H7.

Finding 25. *The behavior of subjects with relatively high cognitive costs (CRT7 scores weakly below the median) differs markedly for Study 1 and AMT subject pools. In contrast, the behavior of subjects with relatively low cognitive costs (CRT7 scores above the median) across the two pools differs little, despite differences in their non-cognitive characteristics.*

Study 1 vs. AMT	Study 1 (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*

Study 1 vs. AMT: $CRT7 > 4$	Study 1 (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

Study 1 vs. AMT: $CRT7 \leq 4$	Study 1 (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 H7: Study 1 vs AMT: Key variables for each subject pool: Means, standard deviations, 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 (Satterthwaite's degrees of freedom for unequal variances), *p*value stands for $Pr(|T| > |t|) = 0$. (Significance * 0.10 ** 0.05 *** 0.01 **** 0.001.)