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

John Duffy† Ed Hopkins‡ Tatiana Kornienko§

July 2023

Abstract

Cooperation in repeated interactions is important for much socio-economic activity. In this paper we put subjects in the simplest dynamic setting that can rationalize cooperative behavior while eliminating confounding factors such as multiple equilibria, strategic uncertainty, and other regarding concerns. We find that, over all supergames, only 1-2% of subjects behave perfectly consistently with rational choice predictions, and only 3-5% behave consistently with the theory at least 95% of the time. These low frequencies suggest that the rational choice framework used to explain cooperative behavior may not be empirically relevant, either in the standard subject pool or in a more representative online subject pool. We document that while the majority of subjects make dominated choices resulting in money left on the table, a substantial minority are able to earn payoffs in excess of rational choice predictions by "end-timing", or gambling on the end to a supergame.

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

JEL Codes: C72, C73, C92.

*For helpful comments and suggestions, we thank Maria Bigoni, Gabriele Camera, Tim Cason, Subhasish Chowdhury, David Cooper, Guillaume Fréchette, Eugenio Proto, Yaroslav Rosokha, Yefim Roth, and Alistair Wilson, as well as audiences at UC Irvine, the MiddExLab Seminar Series, the Virtual East Asia Experimental and Behavioral Economics Seminar, the North American ESA meetings in Tucson and Santa Barbara, the Behavioral Game Theory Conference at the University of East Anglia, the Conference on Learning, Evolution and Games in Lucca, King's College London, the BETEcon Workshop at the University of St Andrews, and the Barcelona Summer Forum. We are grateful to Patrick Julius and Yaoyao Xu for programming the experiment. Funding for this project was provided by the UC Irvine School of Social Science. The experimental protocol was approved by the UC Irvine Institutional Review Board. The reported experiment is part of a project that was pre-registered on the AEA RCT registry, https://doi.org/10.1257/rct.6318-1.0.

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

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

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

1 Introduction

Cooperation in repeated interactions is important for much socio-economic activity. However, despite an extensive experimental literature, it is still unclear what exactly determines cooperative behavior in prisoner's dilemma settings (see Dal Bó and Fréchette (2018) for a survey). One common observation is that subjects sometimes cooperate even in a one-shot prisoner's dilemma. Equally, subjects sometimes cooperate too little when a dilemma is indefinitely repeated. For example, when the continuation probability (discount factor) is sufficiently high that strategies supporting cooperation such as the Grim trigger strategy could be an equilibrium, a significant proportion of subjects always choose to defect, thereby leaving money on the table. However, it is difficult to interpret either behavior as mistaken, given many possible confounds, including diverse beliefs about opponents' strategies, heterogeneous risk attitudes, social preferences and cognitive limitations.

In an effort to simplify the analysis, this study introduces a novel experimental design that deliberately eliminates various confounding factors. The experiment involves two subject pools: a traditional group consisting of university students (referred to as "Lab" subjects), and a second group consisting of Amazon Mechanical Turk (AMT) workers (referred to as "Mturkers"). Both sets of subjects play a series of indefinitely repeated prisoner's dilemma (IRPD) games with different continuation probabilities against a robot opponent known to play the Grim trigger strategy.[1] This design reduces or eliminates multiple equilibria, strategic uncertainty, and social preferences as factors influencing cooperation. This allows one to focus attention on the cognitive task of trading off present gains against future rewards, relying on basic dynamic programming arguments. The optimal policy is simple in theory: a subject should cooperate in each round if and only if the continuation probability δ is above a critical level, here 0.5. Further, much existing experimental analysis of repeated games focuses on first round behavior because behavior in subsequent rounds depends on a subject's experience of playing against her opponent. In our setting, opponent behavior

[1] While Tit-for-Tat (TFT) seems to be a reasonable alternative to Grim Trigger strategy, it has a number of disadvantages - see Section 3 for a discussion.

is perfectly predictable, and thus we can analyse whether subjects implement the optimal policy over the *entire supergame*.

Despite this simple setup, we find that both the Lab student subjects and the more representative population of Mturkers (AMT) behave in a manner that is strikingly different from the predictions of the rational choice framework used to explain cooperative behavior. Overall, across all supergames, only 2% and 1% of subjects (in the Lab and AMT pools, respectively) behave consistently with the rational choice predictions, and only 5% and 3% behave consistently with that theory at least 95% of the time (making no more than 2 suboptimal choices out of 48). These low frequencies essentially amount to noise, and suggest that the rational choice framework used to explain cooperative behavior may not be empirically relevant. While we acknowledge that our design overlooks a crucial aspect of repeated games, namely the presence of other *human players*, it is difficult to imagine that subjects would make fewer errors when confronted with the additional complexities of strategic uncertainty, uncertainty regarding others' preferences and beliefs, and the existence of multiple equilibria that are all features of the standard setting.

Using subjects' choices in the first rounds of each supergame as a clear upper bound for the overall measure of rational play, we find that only 2% and 3% of subjects (in the Lab and AMT pools, respectively) behave consistently with the rational choice predictions at the start of each supergame. We further find that first round cooperation is strongly increasing in the continuation probability δ, from 10% and 26% when $\delta = 0.1$ to 76% and 74% when $\delta = 0.7$ (again, in the Lab and AMT pools, respectively). This responsiveness to the continuation probability δ (particularly in the Lab pool) is much greater than estimates based on Dal Bó and Fréchette (2018) in standard subject versus subject experiments, which suggests that our design is successful in reducing strategic uncertainty. However, on average, subjects cooperate too much in the first round (48% and 52% of decisions rather than the theoretically optimal value of 33%).

As noted, our methodology allows for a detailed examination of behavior in rounds beyond the initial one. Our analysis reveals significant deviations from the optimal strategy. First, 52% and 54% of subjects (in the Lab and AMT pools, respectively) cooperate at least once after already having defected in a supergame, behavior that is difficult to rationalize. Further, 24% and 30% of subjects make this type of mistake repeatedly, in at least 3 out of 17 relevant supergames. While the frequency of such behavior diminishes with experience, it never completely disappears.

Second, subjects commonly defect after having begun the play of a supergame by cooperating. Specifically, we find that cooperation is significantly decreasing with the round number – when theory suggests that, given initial cooperative play, a player should continue cooperating for the duration of the supergame. Although the supergames have an unknown random end, subjects appear to be engaging in what we call "end-timing", which is gambling on the end of a supergame, defecting in the round they guess will be the last round of the supergame.[2] This "end-timing" strategy is similar to "sniping" in auctions (see Roth and Ockenfels (2002)), and we find that such"end-timing" behavior increases with experience. We are able to identify these behaviors only because of our novel, single-person design, and our findings offer an alternative interpretation of results from other repeated game experiments involving matched pairs of human participants.[3]

We further find that test scores from a cognitive reflection test (CRT) predict earnings and are negatively associated with the error of cooperating after having defected, supporting the idea that cognitive failures and/or lack of attention are a cause of deviations from optimal behavior. However, we also find that individuals with higher cognitive test scores are more likely to "end-time", or use timing strategies. More generally, our findings indicate that cognitive factors play a crucial role in explaining the excess cooperation observed when the continuation probability δ is low, as well as the insufficient level of cooperation observed when δ is high.

[2]As is standard in indefinitely repeated games, we employ a constant termination probability of $1-\delta$, where δ is fixed and known to subjects. Subjects may nonetheless believe that the length of a supergame rises with the termination probability. Alternatively, Mengel et al. (2022) find that subjects respond to past *realized* supergame lengths.

[3]Romero and Rosokha (2018) and Cooper and Kagel (2022) also report decreasing cooperation rates in indefinitely repeated prisoner's dilemma game experiments. However in those settings, the decrease in cooperation may be caused by beliefs that it is the cooperation by opponents which may about to end.

Behind these aggregate results, there is considerable heterogeneity - some subjects never cooperate while others always do. In an effort to explain this diversity, as well as the incidence of dominated choices, we propose and test a novel model of inattentive behavior. Specifically, we hypothesize that subjects might be inattentive to the payoff-generating process, including continuation probabilities δ, opponent's strategies, and so on. First, we adapt the approach of Gabaix (2019) and assume that there is an unknown payoff associated with an unknown state of the world, which the agent seeks to match with her action. As we show, this formulation of rational inattention is convenient as it directly implies a probit choice rule. Second, inattention theory suggests that individuals with less precise information about a decision problem will therefore be more influenced by their own default prior payoff for that decision which is typically obtained in comparable situations outside the laboratory.

Indeed, we find that choices to cooperate by subjects with lower cognitive test scores (particularly for the Lab sample) are correlated with their intrinsic level of patience (which we elicited and interpret as a default prior for cooperation) even though the only fundamental factor relevant to making cooperate/defect decisions is the continuation probability δ. In contrast, subjects with higher test scores respond more strongly to the continuation probability δ, and there choices are not influenced by their own patience level. Furthermore, only the choices of subjects with high cognitive test scores correlate with their interpretation of the structure of the game, as represented by another elicited measure, their prediction of the aggregate frequency with which they will choose to cooperate in the first round of all supergames, after they were informed about the set of continuation probabilities, but before they begin play of the games.

Experimental economists have used robot players for control purposes in a number of studies.[4] Two prior studies, by Roth and Murnighan (1978) and Murnighan and Roth (1983), are most closely related in having a population of subjects play the repeated PD game against a fixed strategy, as well as being the first studies to run experiments on supergames

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

with an uncertain end. However, subjects in those studies were *not informed* of the strategy they faced or that their opponent was in fact the experimenter. Thus, subjects in those two studies, who participated in sessions along with other subjects, faced some strategic uncertainty.[5] By contrast, in this experiment we instruct subjects that they are playing against a programmed opponent who *plays the Grim trigger strategy*, and precisely what this means. In addition, these two prior studies did not allow subjects to play multiple supergames with the same continuation probability. Dal Bó and Fréchette (2018) argue that such repetition is an important feature, in that more recent experiments have found significant learning effects with experience. Learning may be less important in our setting where there is no strategic uncertainty, but nonetheless we think it is important to give subjects an opportunity to learn by doing. In the most similar paper, Duffy and Xie (2016) consider play against robot players known to play the Grim trigger strategy but in an n-player Prisoner's Dilemma game under random matching, where they vary n and the stage game payoffs but *not* δ. Also related is Andreoni and Miller (1993) who study play of *finitely repeated* PD games when there is a known probability that the opponent could be a robot, tit-for-tat player (and not another human subject). They find that cooperation increases with the probability of facing such a robot player.

Of course, there are many experiments involving play of the indefinitely repeated prisoner's dilemma game, where subjects play other subjects. For example, Proto et al. (2019) (see also Proto et al. (2022)) also find that individual personality differences between subjects affect play, with higher cognitive ability players being more cooperative, making fewer mistakes and earning higher payoffs. Individual differences across subjects also help to explain differences in the frequency of cooperative behavior as documented, e.g., by Davis et al. (2016), Kölle et al. (2020) and Gill and Rosokha (2022).

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

However, in all of these studies involving human vs. human subject pairings, there is no unique optimum policy as there is in our study. Consequently, any errors must be inferred. For example, Proto et al. (2019) assume that playing defect directly after both players chose to cooperate is an error in implementation. However, our experiment reveals that such behavior may represent an attempt to guess the final round of a supergame. Further, as noted, with our design we can also identify the dominated behavior of cooperating after one has previously defected.

Recently and independently, Normann and Sternberg (2021) and Reverberi et al. (2021) also carry out experiments that involve human subjects playing repeated games against robot strategies. Their designs and research questions are very different from ours, however. In Reverberi et al. (2021), subjects played repeated games against robot opponents but the game and the strategy faced could change randomly over time. They then look at the interaction between the complexity of the strategy subjects face and their cognitive ability in determining the frequency of mistakes. Normann and Sternberg (2021) investigates three and four firm oligopolies in which one firm's prices are determined by an algorithm. The question is whether such algorithmic pricing facilitates collusion. This latter study indicates that interactions between humans and algorithms or robots may have increasing practical importance.

Surprisingly, only a few experiments have been conducted to assess subjects' capacity for dynamic optimization, despite its fundamental importance in contemporary economic theory. These experiments, such as Noussair and Olson (1997) and Carbone and Duffy (2014), also reveal deviations from optimal behavior. However, in our study, the problem at hand is even simpler than in these other studies as it does not involve conditioning on some changing, continuous state variable such as capital or wealth. Further, the trade-off between immediate gains and future cooperation in our experiment is arguably much more intuitive.

Our methodology shares similarities with the approach taken by Charness and Levin (2009) who experimentally demonstrate the existence of the winner's curse phenomenon in a single-person bidding problem. In both their study and ours, simplifying the environment enables the identification of individual cognitive failures that would be more challenging to detect in the original strategic settings. In our case, the individual failure pertains to the inability to adhere to a consistent strategy within a stationary environment, resulting in suboptimal behavior even without strategic uncertainty. One key distinction (aside from the different game under investigation) is our utilization of a within-subject design, wherein subjects encounter situations where cooperation is optimal and situations where it is not. In this way, our study of repeated interactions draws upon the methodology employed by Duffy et al. (2021) and Charness et al. (2021), who also employed experimental designs involving contrasting environments.

2 Theory and Hypotheses

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

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

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

2.1 Insights from the Theory of Repeated Games

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

To see this, note first that since the computer is programmed to play the Grim trigger strategy, it begins each supergame by choosing to cooperate. It continues to cooperate so long as all previous play by the human opponent has been to cooperate as well, but after any defection by the human opponent, the computer program switches to defect for all remaining periods of the supergame. Thus, any human player should understand, given that the continuation probability is fixed at δ, that the return to playing cooperate (X) forever is

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

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

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

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

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

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

Three important factors present in the standard two player repeated prisoner's dilemma are removed in our experimental design. First, our design eliminates the problem of *multiple equilibria*. When the continuation probability δ is sufficiently high for cooperation to be supported, there is typically an infinite number of equilibria which presents subjects with difficult coordination problems. Given that the opponent in our design is playing the Grim trigger strategy, the set of equilibria is reduced to just a singleton – either always cooperate, or always defect, depending on the continuation probability δ.

Second, our design minimizes *strategic uncertainty*. This type of uncertainty is always present in the standard design because subjects do not know which strategy their opponent is following. Indeed, a simplification used by Dal Bó and Fréchette (2018), following Blonski et al. (2011), is to suppose that strategy choices are limited to the Grim trigger strategy and the strategy of always defecting. They show that there is a $\delta^{RD} > \delta^*$ such that, if and only if $\delta > \delta^{RD}$, it is risk dominant to choose the Grim trigger strategy and hence to start out cooperating. Or, in other words, although it is an equilibrium to cooperate as long as $\delta > \delta^*$, strategic uncertainty can make it difficult to cooperate unless $\delta > \delta^{RD}$, a higher hurdle.

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

2.2 A Simple Cognitive Model

In this section, we present a simple cognitive model that can be used to explain deviations from optimality in single-person decision problems. The optimal strategy is to cooperate if and only if the current δ exceeds a threshold of $\delta^* = \frac{1}{2}$. This cognitive model tries to place some structure on deviations from this ideal so as to better understand how they may vary with subjects' cognitive abilities. The model is based on ideas about attention in Gabaix (2019) and cognitive uncertainty in Enke and Graeber (2019).[6] While our model doesn't account for all observed deviations, it is particularly useful for explaining subjects' choices in the initial round of a supergame.

The model works in the following way. A decision maker is faced with a single-person decision problem. She has a default payoff which is derived from previous experience. But she also attempts to determine the optimal action by introspection. This is modelled by assuming that she receives a signal, the informativeness of which depends on her cognitive ability which varies across individuals. A further assumption is that she is in effect aware of her cognitive limitations and will place greater weight on her default payoff the higher is her cognitive uncertainty.

We can represent this mathematically, adapting Gabaix (2019)'s simple Gaussian framework, by saying that a subject i has an initial default payoff d_i, which we can think of as a typical payoff to cooperation in situations outside the laboratory. Now in the laboratory, faced with a supergame with a particular continuation probability δ, she must try to update the payoff to cooperation to match the particular circumstances faced. Let the true relative

[6] We show here that 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.

return to cooperation (that is, the payoff to cooperation minus the payoff to defection) in a repeated game with continuation probability δ be $\pi(\delta)$ so that $\pi(\cdot)$ is a strictly increasing function such that, given our chosen parameters, $\pi(\frac{1}{2}) = 0$. Then assume a subject i, faced with a specific decision problem where the continuation probability is δ, subjectively estimates $\pi(\delta)$ as being normally distributed with expectation d_i and variance σ^2, i.e., $\pi(\delta) \sim N(d_i, \sigma^2)$. That is, the subject's initial evaluation of the return to cooperation is influenced by her outside default relative payoff d_i, and this default varies across subjects (though we assume for simplicity that σ^2 is constant and common to all subjects).

However, by further cognitive introspection, she can gain a potentially more accurate estimate of $\pi(\delta)$. We model this by assuming the subject i receives a noisy payoff signal s_i which is equal to the true value $\pi(\delta)$ plus noise ε_i, where $\varepsilon_i \sim N(0, \sigma_i^2)$, so that

$$s_i(\delta) = \pi(\delta) + \varepsilon_i. \tag{4}$$

The noisiness of the payoff signal varies across individuals with σ_i^2 being the variance of the noise ε_i for individual i. The hypothesis is that the higher is the subject's cognitive ability, the lower is the variance σ_i^2, and the more precise is the signal.

Following Gabaix (2019) and by standard Bayesian updating, the subject's posterior estimate of $\pi(\delta)$ is a weighted average of the signal and the prior estimate,

$$P_i(\pi(\delta)|s_i) = \lambda_i s_i(\delta) + (1 - \lambda_i) d_i \tag{5}$$

where the weight λ_i is determined by the relative variances,

$$\lambda_i = \frac{\sigma^2}{\sigma^2 + \sigma_i^2}.$$

Note that as the 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 value of the payoff π. However, for a subject with a high σ_i^2, the posterior estimate is in fact very close to the individual's default payoff d_i.

Remember 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 $P_i > 0$.[7] Thus, when the subject faces a decision problem with an arbitrary continuation probability δ, the subject's probability of cooperation C_i is the probability that the posterior estimate P_i is positive which is,

$$C_i(\delta) = \Pr(\lambda_i s_i + (1-\lambda_i)d_i > 0) = \Phi\left(\pi(\delta) + \frac{\sigma_i^2}{\sigma^2} d_i\right) \tag{6}$$

where Φ is the normal CDF of ε_i with variance σ_i^2. So, the individual's actions are given by a probit in which the probability of cooperation is influenced both by the true payoff and the individual specific default payoff.

Further, one can see that those individuals with higher cognitive ability, and thus with lower cognitive noise σ_i^2, will place less weight on the default payoff d_i and more weight on the true payoff π. Because $\pi(\delta)$ is increasing in the continuation probability δ, one can draw a similar conclusion: those with higher cognitive ability should be more sensitive to δ in their choice of cooperation.[8]

Hypothesis 2. *Cognitive Ability and Cooperation:*

1. *for high cognitive ability (low noise) subjects the probability of cooperation will be less influenced by their default payoff value to cooperation than for low ability (high noise) subjects;*

2. *the probability of cooperation for high cognitive ability (low noise) subjects will be more influenced by the true payoff to cooperation (and, thus, the current continuation probability δ) that they face than for low ability (high noise) subjects.*

A related, natural hypothesis is that higher ability (lower noise) individuals earn higher

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

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

payoffs. This is implied by the model in that the probability of cooperation C will be closer to its optimal value, the larger the relative weight on the true payoffs (holding the default payoff d_i constant). As we have seen, this is the case when σ_i^2 is lower, which is associated with higher cognitive ability.

Hypothesis 3. *Cognitive Ability and Payoffs: subjects with higher cognitive ability (lower noise) will earn higher average payoffs than those with lower cognitive ability (higher noise).*

3 Experimental Design

The main experimental task consisted of the play of 24 indefinitely repeated prisoner's dilemma games or "supergames" against a computer program known by subjects to play the Grim trigger strategy. The payoff matrix for the prisoner's dilemma stage game was held constant across all treatment conditions and is shown in (1). Subjects were instructed that the rows referred to their action and the columns referred to the computer opponent's actions and that the first number in each cell (in bold) was their payoff in points and the second number in each cell (in italics) was the computerized opponent's payoff in points.[9]

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

The set of 6 continuation probabilities, $\delta \in \{0.1, 0.25, 0.33, 0.4, 0.67, 0.7\}$, were selected using several criteria. First, given specification 1, the expected total theoretical payoff averaged across this set of continuation probabilities is the same for subjects who are biased towards either always cooperating or always defecting. Second, the expected payoff from always following the theoretically optimal strategy relative to either of the fully biased strate-

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

gies is substantial and results in a clear difference. Finally, since the threshold probability for sustaining cooperation in the stage game (1), $\delta^* = 0.5$, we did not want the simple heuristic of cooperating in 50% of the supergames to correspond to the optimal policy. Instead, optimal play would involve cooperating in just 8 of the 24 supergames (those with $\delta = 0.67$ or 0.7) and always defecting in the other 16 supergames. Thus, in contrast to most existing studies, our set of continuation probabilities δ is biased towards producing supergames with a shorter duration.

We ran the current experiment with Grim Trigger as the only programmed strategy. While Tit-for-Tat (TFT) seems a reasonable alternative to Grim, it has the following difficulties. First, as it gives a weaker punishment than Grim, cooperation is only a best response for higher continuation probabilities (c.f., the critical continuation probability $\delta^* = 0.75$ for the current parameters). We would therefore have to run supergames with a higher expected length. Second and more importantly, TFT 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.

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

Subjects were first asked to provide demographic information. Next, they answered 7 questions on economic preferences (e.g., "Patience", "Risk", and so on) on an 11-point Likert scale (taken from Falk et al. (2018)), followed by 7 cognitive reflection test (CRT7) questions (based on Toplak et al. (2014) and Ackerman (2014) – see Appendix C for these questions).

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

After subjects were presented the list of all continuation probabilities δ, they were 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. 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 experimental instructions, and indeed, we use this prediction data later on in the evaluation of our model of inattentive behavior.[10]

For half the subjects, the sequence of randomly drawn continuation probabilities δ and the realised durations (in parentheses) for each of the 24 supergames (4 supergames of each of the 6 δ values) were as follows:[11] 0.67 (4), 0.33 (1), 0.4 (2), 0.25 (1), 0.7 (3), 0.33 (2), 0.7 (5), 0.4 (1), 0.67 (2), 0.1 (1), 0.25 (1), 0.1 (1), 0.25 (2), 0.1 (1), 0.4 (1), 0.67 (4), 0.33 (2), 0.25 (1), 0.7 (2), 0.4 (3), 0.67 (2), 0.1 (1), 0.7 (4), 0.33 (1), resulting in a total of 48 decisions (see Table A2 in Appendix A). For the other half of subjects, the order of these supergames was reversed.[12]

Subjects were instructed that at the end of the session, six supergames would be chosen from all 24 played, one from each of the six different values for δ.

[10]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.

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

[12]See Appendix B.1 for a discussion of order effects.

3.1 Subject Pools and Experimental Earnings

We recruited two gender-balanced subject samples, university students (Lab) and Mturkers (AMT), and we used the same program for both subject groups.

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

The Lab subjects were instructed that their total point earnings from the six randomly chosen supergames (one for each δ value) would be multiplied by USD \$0.01 and this amount would comprise their monetary earnings from the repeated PD games. Subjects were guaranteed \$7 for showing up and completing the study. The student subjects' total earnings averaged \$17.90 for a 1 hour experiment.[13]

We also recruited 300 subjects on Amazon Mechanical Turk (AMT) who resided in the United States (henceforth the AMT subjects). Only 149 AMT subjects completed the experiment. While this might seem like a low completion rate, almost all dropouts occurred at the instruction or quiz stage of our study; the study completion rate for subjects who got past the comprehension quiz was nearly 100%. Selection bias is only an issue if dropout decisions vary with variations in experimental treatment conditions, and in line with the Arechar et al. (2018) study of AMT subjects, we find that dropout choices were exogenous to treatment conditions (here, whether subjects faced the long order or the reverse order of supergame lengths - with 74 (75) out of 148 (152) subjects completing the long (reverse) order).

[13] Following the 24 repeated PD games, the Lab subjects (and only the Lab subjects) were randomly paired to participate in another two-player task which we do not report on in this paper, and where the subjects could earn an additional \$1.00-1.70 payment.

There was an equal number of AMT subjects who reported to be males and females; one subject preferred not to report their gender (coded as 0.5). The age of the AMT subjects is dramatically different from that of the Lab subjects, with a mean age of 39.8 years and a range of 21-75. (See Table B3 for a formal comparison of the two subject pools.) The AMT subjects also had more dispersed levels of educational attainment including 2 subjects who reported not having a high school diploma.

Like the Lab sample, the AMT subjects were instructed that at the end of the study, their total point earnings from the six randomly chosen supergames (one for each continuation probability δ value) would be multiplied by USD \$0.01 and that this amount would comprise their monetary earnings from the repeated PD games. Thus, the Lab and AMT subjects faced the *same* variable earnings potential in the repeated PD games of the study, which facilitates comparisons. However, the fixed show-up payments to the Lab and the AMT subjects were \$7 and \$1, respectively. This difference in show-up payments reflects different payment norms in the recruitment of the conventional pool of laboratory subjects versus online workforces such as the AMT sample – see, e.g., Rand (2012). The AMT subjects' total earnings averaged \$10.37 for a 1 hour experiment.

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

4 Results

In this section we report on the main results of our experiment, which we present as a number of different findings. For each finding, we analyze and report on data from the Lab and AMT samples separately.

The design for both subject populations required completion of 24 supergames each involving the play of at least 1 round. Given the randomization, 7 of these 24 supergames ended after round 1. The remaining 17 supergames lasted 2-5 rounds. Thus, each subject made 24 first round choices and (by chance) 24 non-first round choices for a total of 48 choices. (see Table A2). Given the parameters of our design, the theoretically optimal strategy involves choosing to cooperate in all rounds of the 8 supergames where $\delta = \{0.67, 0.7\}$ (and thus, a total of 26 choices to cooperate), and to defect in all rounds of the other 16 supergames.

4.1 The Headline Result

We will start with our main result. As Figure 1 shows, only a small fraction of subjects behave according to the standard game-theoretic predictions (see Hypothesis 1). Specifically, only 2 subjects out of 100 (or 2%) in the Lab student sample and only 1 subject out of 149 (or 0.7%) in the AMT sample behaved perfectly as predicted. Furthermore, in both subject pools, no more than 5% of subjects made no more than 3 mistakes out of 48 total choices (i.e., at least 45 optimal choices out of 48, or 93.75% of choices). And about 10% in each sample (10% in Lab, 10.7% in AMT) made at least 42 optimal choices out of 48 (or 87.5% of all choices). (See Appendix B.2 for corresponding results on the first rounds of each supergame, as well as for the juxtaposition with the per-subject counts of cooperative choices.)

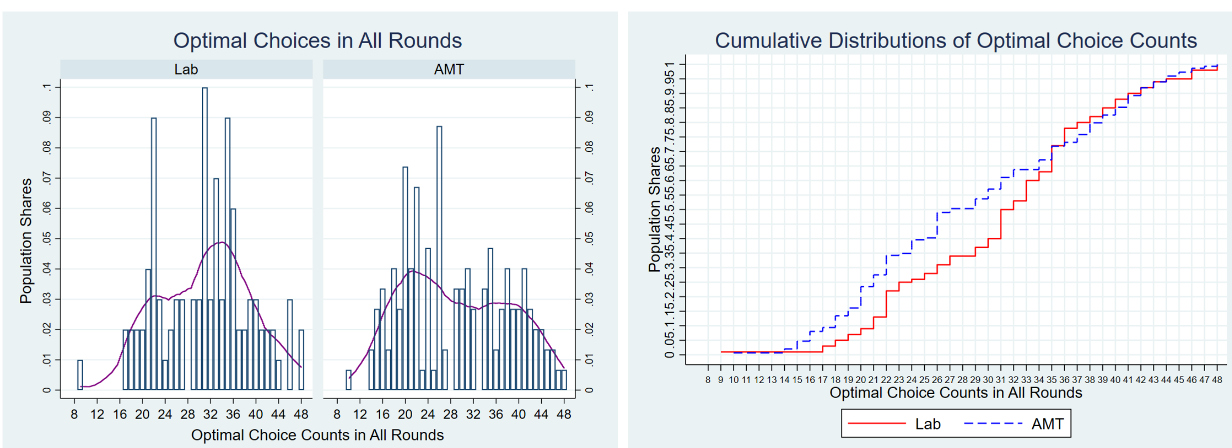

Figure 1: Left: Frequency distributions and kernel densities of per-subject counts of optimal choices across all 24 supergames (48 decisions total): Lab (N=100) vs. AMT (N=149) subjects. Right: Both cumulative distributions on the same graph.

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

Interestingly, while the Lab subjects behaved significantly more theoretically optimally than the AMT subjects (using a Kolmogorov-Smirnov one-sided test $D = 0.1799, p = 0.021$), the differences between samples are less pronounced among subjects with greater tendencies to make theoretically optimal choices as revealed in Figure 1.

In what follows we will attempt to decompose the observed deviations from the standard game-theoretic predictions.

4.2 Response to the Continuation Probability δ

As the left panel of Figure 2 reveals, cooperation is strongly increasing with the continuation probability δ. This finding is confirmed by mixed-effects probit regression results reported in Table 3, specifications (1)-(2). First round cooperation rates are as low as 9.5% for the Lab subjects and 25.5% for the AMT subjects when $\delta = 0.1$, and as high as 76.25% for the Lab subjects and 73.83% for the AMT subjects when $\delta = 0.7$. For the Lab subjects, this responsiveness is much greater than is observed in standard subject versus subject experiments.[14] In contrast, the AMT subjects are less responsive, particularly for $\delta = 0.1$.

Finding 2. *For every round of a supergame, the rate of cooperation (defection) tends to increase (decrease) with the continuation probability δ. The Lab subjects are more responsive to the continuation probability δ than are the AMT subjects.*

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

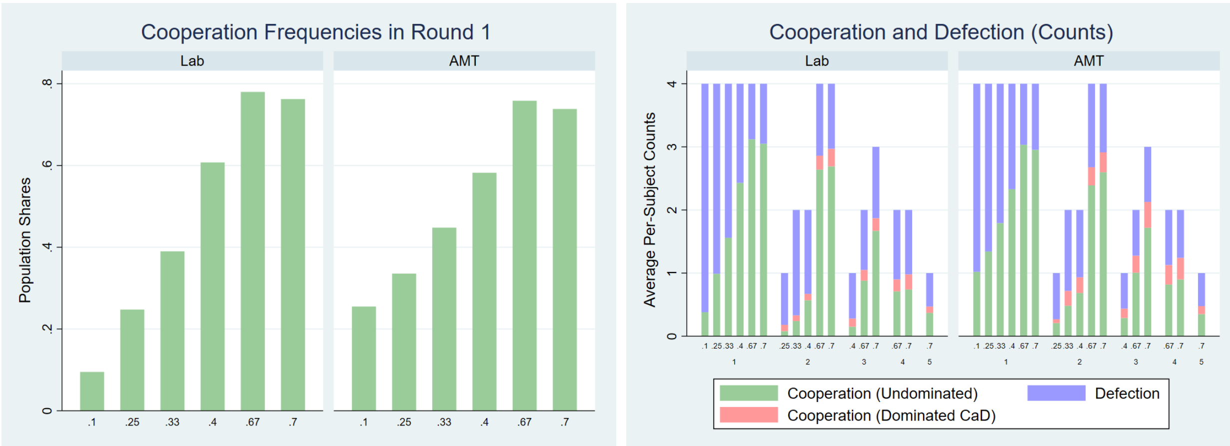

Figure 2: Patterns of cooperation and defection: Left panel: population shares of cooperative choices (or cooperation rates) in the first rounds of all supergames, split by continuation probability δ (see also the leftmost set of bars in the right panel). Right panel: Average per-subject counts of cooperation versus defection, split by continuation probability δ (the first row of the horizontal axis scale) and by the round number of a sequence (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 Table A2). See Figure B2 for the corresponding optimal versus suboptimal choices. (Lab: 100 subjects, 2,400 supergames, AMT: 149 subjects, 3,576 supergames).

4.3 Strategic Error of Cooperating after Defection (CaD)

Since the robot opponent was programmed to play the Grim trigger strategy, a defection at any time in a given supergame would trigger subsequent defection by the automated opponent in all remaining rounds. Thus, given that subjects were informed of the robot's strategy, choosing to cooperate after defecting earlier within the *same* supergame (CaD) is *dominated* for any δ, and is a *strategic error*. In the right panel of Figure 2 such suboptimal cooperation is distinguished from un-dominated/non-erroneous cooperation in the cooperation counts. Among the relevant observations (i.e., in the 17 supergames lasting more than one round), the share of CaD errors is 7.57% for the Lab subjects and 11.47% for the AMT subjects.

As Figure 3 (left) shows, slightly less than half of subjects (48% for Lab and 45.64% for AMT) never made the strategic error of CaD, and 20% of the Lab subjects and 28.86% of the AMT subjects made at least 4 dominated CaD choices. Some such choices could be

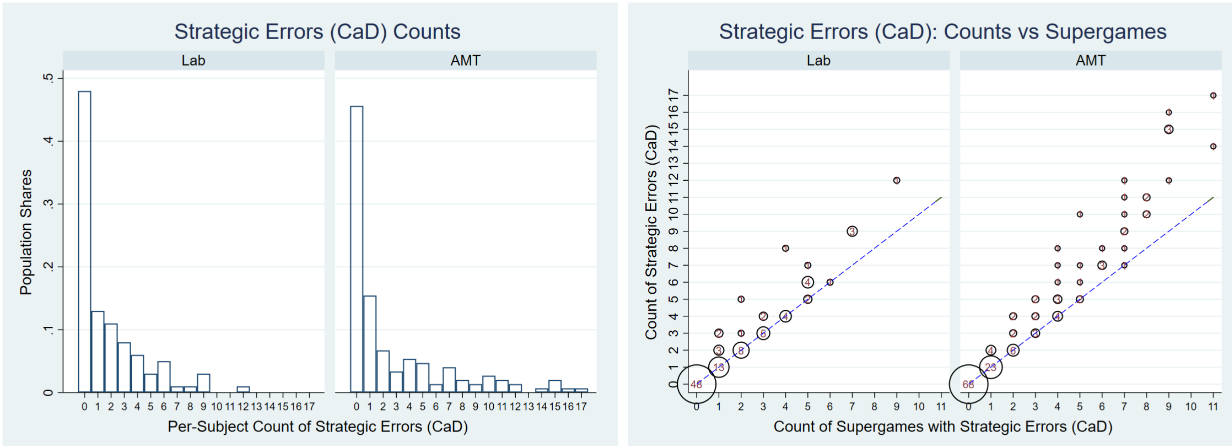

Figure 3: Strategic errors of dominated cooperation after defection (CaD) in 17 relevant supergames (i.e., those lasting longer than one round). Left: Distribution of per-subject counts of instances of cooperation after defection (CaD). Right: Per-subject counts of CaD instances vs. count of supergames with those instances. Bubble size is proportional to the share of subjects (100 Lab subjects, 149 AMT subjects).

intentional, e.g., due to a desire to verify the computer opponent's behavior.[15] Others could be due to a genuine "trembling hand" error of accidentally pressing the "defect" button without noticing it. In either case, an attentive payoff-maximizing subject would likely refrain from repeatedly making dominated choices in multiple supergames.[16]

Given the possibility of trembling hand behavior, we consider the number of supergames where the strategic error of CaD happens at least once. Figure 3 (right) compares the total count of strategic errors (CaD) per subject (vertical axis) versus the count of supergames where such errors were made (horizontal axis). While most strategic errors were made only once in a supergame (as revealed by the bubbles located on the diagonal in the figure), the extent of strategic errors is substantial, with 24% of the Lab subjects and 29.53% of the AMT subjects making errors in at least 3 out of the 17 relevant supergames (those lasting more than one round), suggesting that some of the dominated CaD behavior could instead be due to inattention or a lack of strategic understanding of the game (though recall that subjects had to pass a quiz before proceeding to the play). While the prevalence of CaD

[15]Recall, however, that the computer program's action choice of X or Y for the current round, based on the history of play and following the Grim trigger strategy, was shown to subjects on their decision screen in advance of their entering a choice for that same round so this uncertainty should have been minimized.

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

errors is relatively small, it nevertheless complicates the interpretation of the deviations from the theoretically optimal behavior.

Finding 3. *A majority of subjects (52% in Lab and 54.36% in AMT) made at least one strategic error of choosing to cooperate after defecting (CaD) earlier within the same supergame, i.e., after triggering a "grim" response. Overall, suboptimal, excessive cooperation amounts to 7.58% for Lab and 11.47% for AMT of the relevant observations, with 24% of the Lab subjects and 29.53% of the AMT subjects making dominated choices in at least 3 out of 17 relevant supergames (i.e., those lasting longer than one round).*

4.4 Overall Point Totals

In this section we focus on total awarded points, which is the sum of points earned over all 48 decisions. This serves as a theoretical measure of the payoff consequences resulting from subjects' behavior. First, recall that the theoretically optimal policy of cooperating (defecting) in every round of every supergame when $\delta > (<) 0.5$ (see Hypothesis 1) is derived *ex ante*, before the lengths of each supergame are realised. One can calculate that, given the realization of random supergame terminations, the *ex ante* optimal play would result in an overall total of 4,050 points *ex post*.[17]

Figure 4 reports on subjects' point totals. As this figure shows, the empirical range of overall point totals is $3,285 - 4,185$ points for the Lab sample and $3,195 - 4,185$ for the AMT sample with a mean (st. dev.) of 3,835.05 (203.38) for Lab and 3,766.21 (240.58) for AMT. Note that the lowest *ex post* point total that is achievable here is quite substantial at 2,925 points.[18] This *ex post* lower bound can be seen as the "fixed" component of the

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

[18] This *ex post* theoretical minimum of 2,925 points arises from making the strategic mistake of cooperating after defecting and is calculated as follows. In the supergames lasting only one round one achieves the lowest payoff of 75 by cooperating (7 decisions). But in the supergames lasting longer than one round, the lowest payoff is obtained by defecting in the first round and earning 120 (17 decisions) and earning only 15 from cooperating thereafter (24 decisions).

overall point total, and any overall point total in excess of that amount can be seen as the "variable" component. It is easy to see that, on average, the Lab subjects earned only 80.9% and the AMT subjects only 74.8% of the "variable" component of the overall point totals that could be achieved by following the theoretically optimal policy.

As Figure 4 further reveals, some Lab subjects earned as little as 32% and some AMT subjects as little as 23.1% of the "variable" component which could be achieved by following the optimal policy. Furthermore, 16% of the Lab subjects and 28.86% of the AMT subjects could have *increased* their total point earnings to the level of 3,600 points by simply choosing either to always cooperate (All-C) or to always defect (All-D),[19] which is 60% of the "variable" component that could be achieved by following the theoretically optimal policy. This observation suggests that, among other deviations, strategic errors (CaD) reduce overall point totals.

As Figure 4 further shows, the mode for the Lab sample is at the theoretically optimal point total of 4,050, and this point total is also relatively frequent among the AMT subjects. Yet, strikingly, the maximum point total is still higher in both samples. Overall, 19% of the Lab subjects and 16.11% of the AMT subjects were able to achieve at least the theoretically optimal point value of 4,050, despite only two subjects in the Lab sample and one subject in the AMT sample actually behaving in a way that was fully theoretically optimal (in all 48 choices). Thus, given the random realization of the supergames, some subjects were able to achieve at least as much as the theoretically optimal point total despite pursuing strategies that were not theoretically optimal. A potential explanation for this mystery is provided in the next Section 4.5.

Finding 4. *The Lab subjects earned only 80.9% and the AMT subjects only 74.8% of what could be achieved relative to the* ex post *theoretical minimum by following the* ex ante *optimal policy. 16% of the Lab subjects and 28.86% of the AMT subjects achieved lower point totals than what they could have achieved by either always cooperating or always defecting.*

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

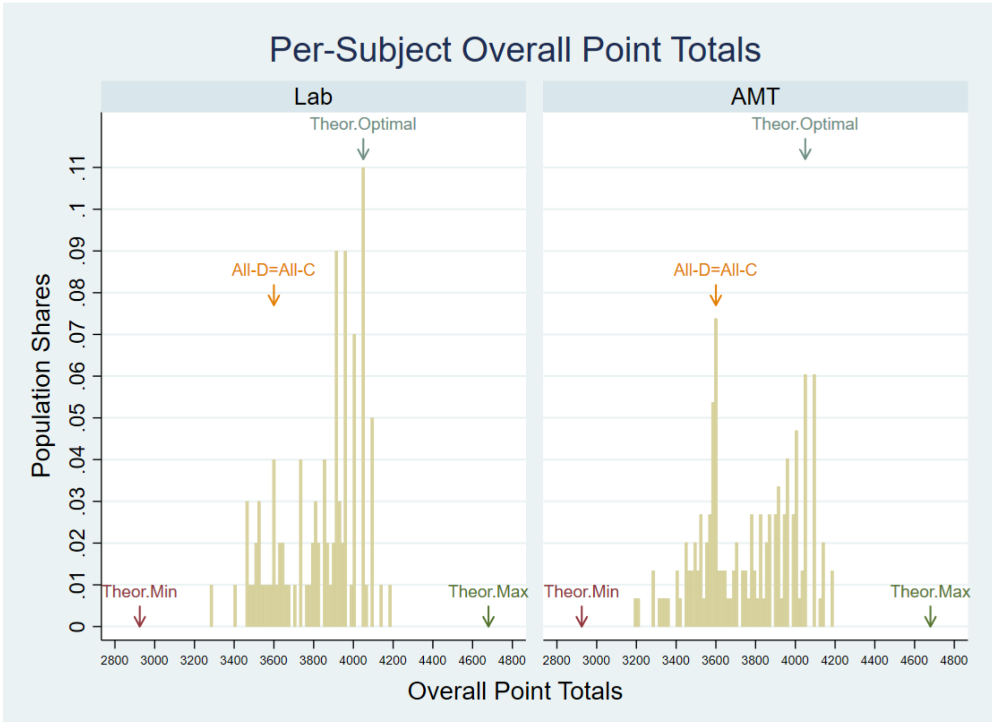

Figure 4: Distribution of overall point totals, or the sum of point earnings across all 48 decisions. *Ex post* theoretical point total from following the *ex ante* optimal policy is 4,050, while the ex post theoretical minimum and maximum point totals are 2,925 and 4,680, respectively.

Importantly, 17% of the Lab subjects and 15.43% of the AMT subjects were able to achieve point totals at least as high as the theoretically optimal point totals – without always following the theoretically optimal strategy.

4.5 "End-Timing" (DaC)

Note that in our study, when playing against a robot that uses the Grim trigger strategy, one could potentially earn a higher total of 4,680 points as compared with the theoretically optimal strategy which earns only 4,050 points. Such a higher point total can only be achieved by knowing *exactly when* each supergame would end and using a perfect "end-timing" strategy. This strategy involves cooperating in all rounds before the final round of a supergame, and then defecting in the final round to earn the temptation payoff without triggering the grim punishment. Since subjects in our study did not know when a supergame would end, they could not execute this end-timing strategy perfectly.[20] Nevertheless, some subjects appear to be pursuing this type of strategy. While we refer to this as "end-timing" behavior, that phrasing is short-hand for succumbing to the "gambler's fallacy" – the mistaken belief that the probability of an event occurring is lowered when that event has occurred more frequently in the recent past, even though the probability remains the same (see, e.g., Cowan (1969)). While our subjects did not know when a supergame would end, they may have formed expectations about the possibility of it ending, leading to the use of this timing strategy.

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. Thus this end-timing strategy 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, however, that risk attitudes of subjects cannot in themselves explain this behavior. It can be shown that, under CRRA preferences, for example, defection is increasing in risk aversion rather than risk taking.[21] Further, given that in reality the stopping probability is constant, a rational highly risk-averse subject should defect from the beginning of a supergame or not at all.[22]

Indeed, Figure 5 (left) shows that, for some continuation probabilities δ, some subjects defect for the first time (thus triggering subsequent defection by the automated opponent) *later* in the sequence, rather than in the first round (if ever) as predicted by the theory.[23]

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

[21]Despite this theoretical prediction, Dal Bó and Fréchette (2018) report that existing experiments find no clear link between risk aversion and cooperation.

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

[23]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 final round of a sequence, are both increasing with δ – see Table A2. Mengel

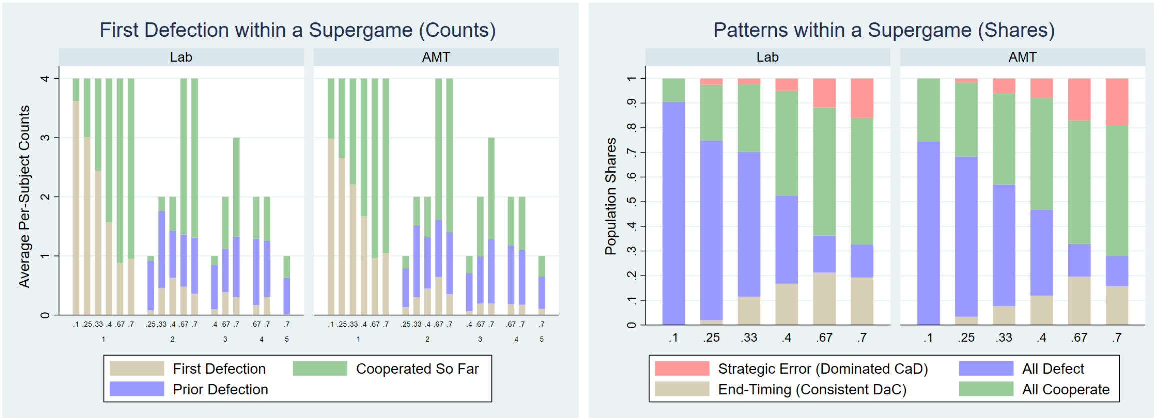

Figure 5: Patterns of cooperation and defection. Left: Average per-subject counts of first defection *within* a supergame by δ and round number. Right: The population shares of the behavioral patterns in a supergame, by δ value. By construction, the four strategies, CaD, DaC, All-D and All-C are mutually exclusive. (Lab: 100 subjects, 2,400 supergames; AMT: 149 subjects, 3,576 supergames.)

Furthermore, as Figure 5 (right) shows, the shares of the supergames where subjects always defected (All-D) are declining as the continuation probability δ increases. However, this does not translate into an increase in the prevalence of the always cooperate (All-C) strategy as the continuation probability δ increases. Instead, as δ (and thus the expected duration of a supergame) increases, both the prevalence of strategic errors (CaD) and "end-timing" (DaC) strategies increase. Note that interpreting All-C strategies is complicated by attrition, as a subject might have *intended* to time their defection, but a supergame ended earlier than expected. Similarly, All-D strategies in low δ supergames may not only be due to theoretically optimal behavior, but might also be observationally equivalent to the use of an end-timing strategy.

Indeed, if the behavior were theoretically optimal, then in the mixed-effects probit regressions reported on in Table 3 (specifications 1-2), the coefficients on the continuation probabilities $\delta = \{0.25, 0.33, 0.4\}$ would have been insignificantly different from the baseline of $\delta = 0.1$, and would only be significantly different for $\delta = \{0.67, 0.7\}$. In addition, the round dummies would all be insignificantly different from the baseline of the first round. Instead, as Table 3 reveals, subjects' tendency to choose cooperation increases with δ, but

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.

decreases significantly with the round number, which is consistent with the use of the end-timing strategy.

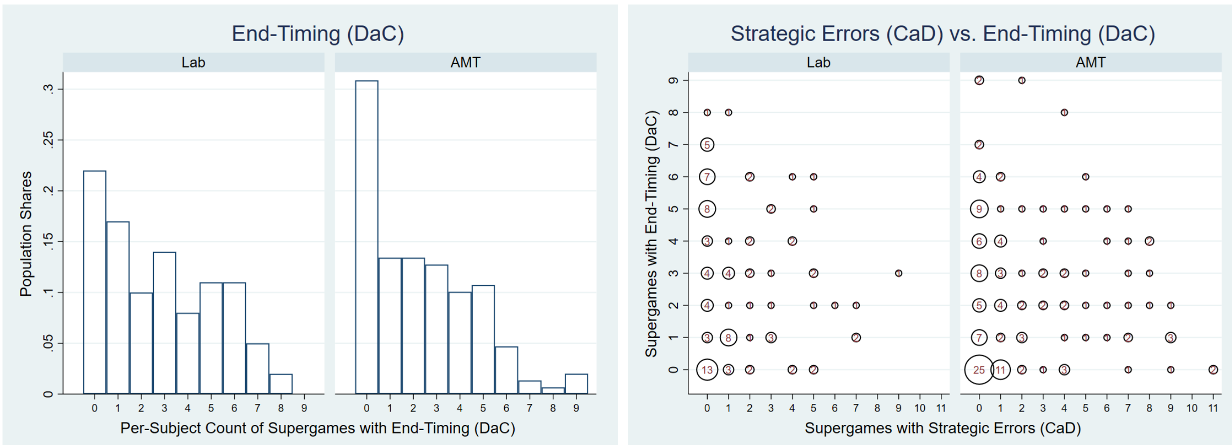

Figure 6: Left: Distribution of per-subject counts of supergames with "end-timing" (DaC), among 17 relevant supergames. Right: Per-subject counts of supergames with strategic errors (CaD) vs. supergames with end-timing (DaC). Bubble size is proportional to the share of subjects (100 Lab subjects, 149 AMT subjects)

The left panel of Figure 6 shows that only 22% of the Lab subjects and 30.9% of the AMT subjects never switched from cooperation to defection within the same supergame (CaD), i.e., never engaged in behavior consistent potentially with the end-timing strategy. However, as the right panel of Figure 6 shows, some potential end-timing behavior may be unintentional, amounting to "mistakes" by subjects who make frequent strategic errors (CaD), i.e., those with higher counts on the horizontal axis of the scatterplots. Yet a few subjects who never or almost never commit strategic errors (CaD) (those closer to zero on the horizontal axis indicating CaD errors) appear to be engaged in end-timing behavior.

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

4.6 Learning Over Time

In this section, we explore whether subjects acquire the ability to play the theoretically optimal strategy and/or make fewer errors such as dominated CaD as they gain experience. As Table 1 shows, subjects in both pools change their behavior in the second half of the experiment (in the last 12 supergames) relative to the first half (first 12 supergames). Specifically, they make fewer dominated CaD errors[24] and cooperate less over time. Interestingly, over time, the AMT subjects move towards the theoretically optimal behavior and earn marginally more total points. However, both subject pools exhibit a pronounced increase in their end-timing activity (DaC), and, overall, the improvement in the frequency of optimal play, while significant, is not very large, rising from 64% (59%) in the first half to just 66% (62%) in the second half for the Lab (AMT) samples.

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

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

Finding 6. *Over time, subjects learn to commit fewer dominated CaD errors, and move*

[24] As Figure B5 in the Appendix shows, while the incidence of dominated CaD errors decline over time, they do not disappear entirely, with the CaD errors comprising 4% for the Lab subjects and 7% for the AMT subjects across the supergames in the second half of play.

closer to theoretically optimal behavior. However, the play of dominated CaD does not disappear over time, and, moreover, the frequency with which subjects engage in end-timing behavior (DaC) increases over time.

4.7 Classifying the Patterns of Play Within Each Supergame

The complex pattern of non-constant intra-supergame play (discussed earlier in Section 4.5) highlights the difficulties in interpreting each type of play in isolation, calling for a more holistic approach. Indeed, our two design innovations allow us to interpret each subject's entire play across all supergames. It turns out that we can classify subjects' patterns of play within each supergame into 6 *mutually exclusive* types: optimal All-C, optimal All-D, suboptimal All-C, suboptimal All-D, strategic errors (CaD), and end-timing (DaC).[25]

As Figure 7 shows, there is no prevalent pattern to subjects' play. The two (out of 100) Lab subjects and eight (out of 149) AMT subjects who always cooperated are represented by dark green and light green bars meeting at the solid red line on the far left side of each panel. Single subjects in each pool who always defected are represented by dark blue and light blue bars meeting at the solid red line on the far right side of each panel. The two Lab subjects and the one AMT subject who always made perfect, theoretically optimal choices are represented by dark green and dark blue bars meeting at the solid red line, towards the right hand side end of each panel.

Furthermore, as Figure 7 shows, both types of non-constant play (CaD and DaC) as well as 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 a notable heterogeneity in subjects' choices to cooperate or defect, without any representative pattern. Only two out of 100 Lab subjects and just one out of 149 AMT subjects always followed the theoretically optimal strategy. Two (one) Lab subjects and*

[25] Note that if the robot opponent was programmed to play Tit-for-Tat instead, the number of possible patterns of play would be at least eight, complicating the interpretation and analysis of subject behavior.

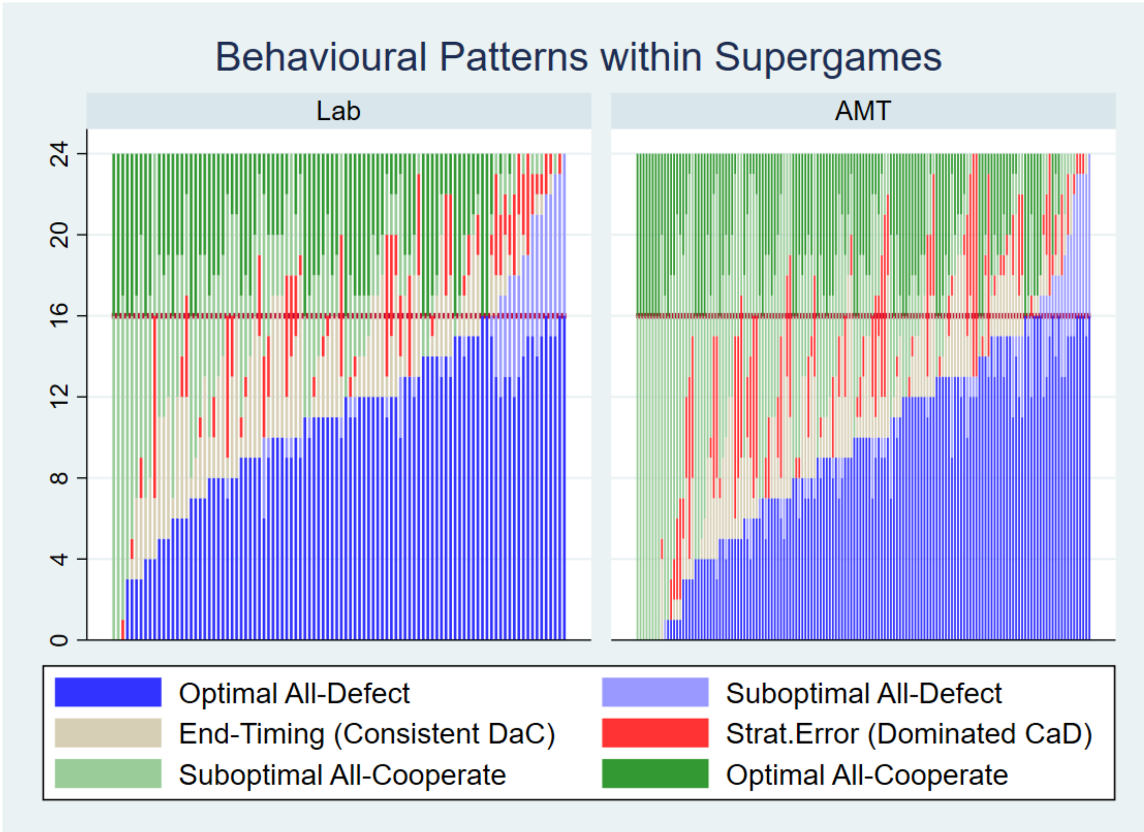

Figure 7: Subject heterogeneity in patterns of choices within supergames, out of 24 supergames, by subject, ordered by the count of supergames with (combined optimal and sub-optimal) All-Defect choices (100 Lab subjects and 149 AMT subjects). The theoretically optimal strategy involves always defecting in 16 supergames and always cooperating in the remaining 8 supergames (represented by the solid red horizontal line).

eight (one) AMT subjects are fully biased towards cooperation (defection). The rest of the subjects appear to pursue strategies that are neither theoretically optimal nor purely biased.

5 Individual Differences and Rational Inattention

In this section we provide an analysis of the individual determinants of subjects' behavior.

5.1 The Effect of Cognitive Abilities

As noted earlier, we asked all of our subjects to answer 7 cognitive reflection test (CRT7) questions as part of the study, and we use subjects' total score on this 7-item test as a measure of their reflective cognitive styles. The mean (st. dev.) of the CRT7 score was 3.78 (2.26) for the Lab subjects and 3.58 (2.17) for the AMT subjects, with a median of 4 for both (see Figure B7, left panel). In terms of CRT7 scores, there is no significant difference between the two subjects pools (two-sided t-test= 0.72, $p = 0.48$).

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

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

As Table 2 shows, the CRT7 score predicts the rational aspects of subjects' behavior rather well. Specifically, it is positively correlated with their overall point totals (specifications 1-2) and negatively with a count of the number of supergames with strategic errors (dominated CaD) (specifications 3-4). Interestingly, for the AMT subjects only, the count of theoretically optimal choices (specification 6) is positively correlated with the CRT7 score. (For the Lab subjects in specification 5, the relationship is marginal, but the overall fit represented by the F statistics, is poor.) For both subject pools, the count of apparent end-timing behavior (DaC) (specifications 7-8) is not correlated with the CRT7 score, or with any other characteristic. Note however, that by construction, these last two counts (theoretically optimal and DaC) are *mutually exclusive*. However. as discussed earlier in Section 4.5, some choices which are consistent with following the theoretically optimal strategy are also observationally equivalent to an end-timing strategy. Thus, it is not a surprise that the combined count of whether subjects follow either the theoretically optimal strategy (which involves either All-D for $\delta < 0.5$ or All-C for $\delta > 0.5$), or engage in end-timing (which involves consistent defection after cooperation, or DaC) *is* positively correlated with the CRT7 score for both subject pools (see specifications 9-10).

In contrast, the CRT7 score on its own has no effect on the choice to cooperate which can be seen in specifications 3-4 of Table 3, which contains average marginals (dy/dx) from mixed-effects probit regressions of the choice to cooperate or defect in all 48 rounds of the prisoners' dilemma game, controlling for demographics, CRT7 score, and other personal characteristics.[26,27]

Finding 8. *Subjects with higher CRT7 scores are more likely to behave in a payoff-maximizing fashion, and engage in end-timing behavior.*

5.2 Inattention

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

To explore these hypotheses, we split our samples in two, according to the median CRT7 score (equal to 4 in both samples). As specifications 7-8 in Table 3 show, subjects with relatively high proxies for cognitive ability (CRT7>4), tend to respond more strongly to the continuation probability δ, less likely to cooperate after defection in the same supergame (i.e., make relatively fewer strategic errors CaD), and exhibit a stronger tendency for following

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

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

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

Table 3: Choices to cooperate: mixed-effects probit regressions, marginals (dy/dx), robust errors in parentheses. (See Table B5 for the corresponding odds.) "Supergame" is the supergame number in the sequence of supergames, "Order Long" is a dummy variable for whether the first supergame in the sequence had $\delta = 0.67$, "Prior Defection" is a dummy variable for whether the subject defected in prior rounds of a given supergame, "Prediction" is the subjects' predictions of the share of their own cooperative choices in Round 1 across all 24 supergames (scaled down by 10). Chi2 and corresponding p-values are from the odds regressions (see Table B5). (Significance * 0.10 ** 0.05 *** 0.01 ***† 0.001.)

the end-timing strategy as their coefficients on the round dummies tend to be negative and strongly significant. In contrast, as specifications 5-6 show, subjects with relatively lower proxies for cognitive ability (CRT7$\leq$4) do not exhibit any systematic sensitivity to the round number.

These regressions further reveal that both lower and higher CRT7 groups each have a single significant factor that consistently correlates with their decision to cooperate in both subject samples. In the lower CRT7 group (Table 3, specifications 5-6), those with higher self-reported *Patience* tend to cooperate more frequently (and, vice versa, those with higher self-reported impatience tend to defect more frequently)[28] – significantly among the Lab subjects and marginally for the AMT subjects. Interestingly, Fehr and Leibbrandt (2011) find that patience measures are very highly correlated with cooperative behavior in a field experiment. Patience has also previously been found to influence cooperation in laboratory experiments (Davis et al. (2016)).

Importantly, the Patience measure was marginally higher for the higher CRT7 group in the Lab subject pool (one-sided $t = 1.3718$, $p = 0.0866$), and there was no difference between the two CRT7 groups in the AMT subject pool (two-sided $t = 0.2802$, $p = 0.7798$). Furthermore, for the AMT subjects but not for the Lab subjects, being female was significantly negatively correlated with choices to cooperate, and risk-prone AMT subjects cooperated marginally less frequently.

In contrast, in the higher CRT7 group (Table 3, specifications 7-8), there are no personal characteristics that systematically explain subjects' choices to cooperate - as a rational inattention theory would predict. (Though cooperation was significantly negatively correlated with Retribution for the Lab subjects and with Punishment for the AMT subjects.) Instead, subjects' choices in Round 1 (which is the regression baseline) are strongly and systematically correlated with the "Prediction" variable (elicited from subjects before any choices were made), possibly reflecting their understanding of the task.[29]

[28]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 C).

[29]Before any choices were made, subjects were asked what percentage of their Round 1 choices across all 24

Finding 9. *Subject behaviour is broadly consistent with a simple model of inattention. Cooperative choices made by subjects with a lower proxy for cognitive ability (higher attention) correlate with an elicited proxy for their degree of patience. In contrast, cooperative choices made by those with a higher proxy for cognitive ability (lower attention) are more affected by the structure of the game, and do not correlate with their individual characteristics, but with an elicited proxy for their understanding of the structure of the game.*

5.3 Comparison of the Lab and the AMT Subjects

The above regression results are summarized visually in Figure 8, which shows that the two groups of subjects exhibit different patterns of play. Subjects in the lower CRT7 group (the two top panels of Figure 8) make relatively more frequent strategic errors (CaD) and engage in suboptimal consistent defecting behavior (Suboptimal All-D). By contrast, subjects in the higher CRT7 group (bottom two panels) are closer to the theoretically optimal policy and engage in end-timing behavior (DaC) more often.

Interestingly, Figure 8 further reveals differences in patterns of behavior between the two subject pools. The top two panels showing behavior by subjects with CRT7 scores less than or equal to the median score are consistent with the earlier insights of Arechar et al. (2018) and Snowberg and Yariv (2021) that the Lab subjects are less prone to pro-social behaviour and less likely to make mistakes as compared with the AMT subjects. However, the bottom two panels, showing behavior by subjects whose CRT7 scores are strictly above the median suggest that there is hardly any difference in patterns of behavior across the two subject pools, despite apparent individual differences in non-cognitive characteristics. Indeed, this

supergames would be cooperative (see Section 3). The mean (st. dev.) of this "Prediction" variable for the Lab subjects was 60.96 (29.57) with a median of 62, while for the AMT subjects it was 66.64 (27.04) with a median of 72, with no significant difference between the two subjects pools (two-sided t-test= 1.56, $p = 0.12$, see Figure B7, right panel). By comparison, the optimal choice is $\frac{1}{3}$ (33.33%). In a tobit regression, this measure is significantly negatively correlated with the "Altruism" measure (at $p = 0.01$) and marginally (at $p = 0.10$) positively correlated with "Retribution" - but only for the AMT subjects, as for the Lab subjects there is no correlation whatsoever (results available on request). Finally, we note that the Prediction variable is significantly correlated with subjects' actual first round choices ($r = 0.3517, p = 0.0003$ for the Lab sample and $r = 0.2350, p = 0.0039$ for the AMT sample).

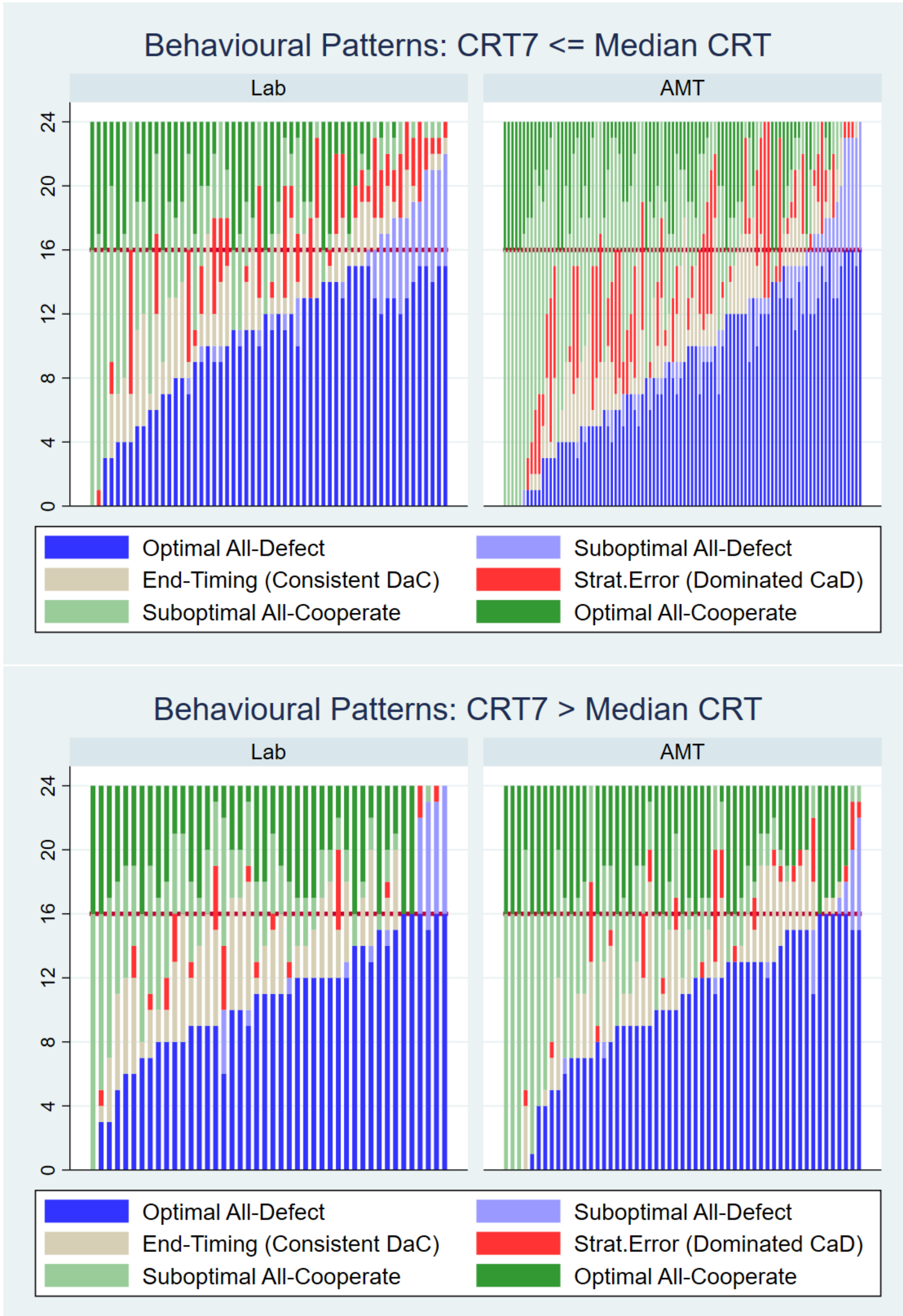

Figure 8: Inattention: subjects' patterns of choices within supergames (out of 24 supergames), split by median $CRT7$. Patterns are presented for each subject, ordered by the count of supergames with (combined optimal and sub-optimal) All-Defect choices. The theoretically optimal strategy involves always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

visual observation is further confirmed by formal t-tests in Table B4.

Finding 10. *AMT subjects with relatively high cognitive costs (CRT7 scores weakly below the median score) behave markedly different from the corresponding group of low CRT7 Lab subjects. In contrast, there is little difference in the behaviour of subjects with relatively low cognitive costs (CRT7 scores above the median score) across the two subject pools, despite differences in their non-cognitive individual characteristics.*

6 Conclusion

We have reported on an experiment where we test the most elemental aspects of the standard, game-theoretic model of repeated interactions employing both student subjects and a more representative sample of AMT subjects. In our experiment, subjects play the repeated prisoner's dilemma game against a robot player known to play the Grim trigger strategy. This design converts the original strategic situation into a single-person decision problem for which there is a unique optimal strategy, and eliminates other confounding factors such strategic uncertainty, social preferences and multiple equilibria. We use a within-subject design in which subjects play many different supergames with differing continuation probabilities. Our design enables us to classify subjects' within-supergame play according to one of 6 mutually exclusive patterns, and to separate theoretically optimal behaviour from bias. We can therefore identify systematic errors made by subjects and relate them to individual characteristics, and, in particular, to their cognitive abilities.

Rather shockingly, we find that only 1-2% of subjects behave consistently with the rational choice predictions, and only 3-5% behave consistently with those predictions more than 95% of the time. These very low frequencies essentially amount to noise and lead to the inescapable conclusion that the dynamic, rational choice framework for understanding cooperative behavior may not so empirically relevant.

When reporting results that are at odds with theoretical predictions, experimentalists are often asked: why do subjects make mistakes? The alternative possibility – that the theory may not be empirically relevant – is not usually up for discussion. Thus, it is important to provide a convincing answer to the question as to why subjects make mistakes. As noted, our stripped-down individual-choice experimental design enables us to "zero-in" on the nature of subjects' mistakes at a granular level and to further explore whether and how individual subject characteristics may play a role in explaining these mistakes.

We find first that a majority (52-54%) of our subjects make at least one strategic error of cooperating after defection. Second, some subjects employ an end-timing strategy, consistently defecting after initially choosing to cooperate (DaC) in the same supergame that can yield them higher payoffs than the theoretically optimal strategy. This finding reveals how subjects can be clever in ways that go beyond the confines of the standard theory and thus why it is important to conduct an experimental evaluation. Third, we show that these different behaviors are correlated with our proxy measure for cognitive ability. Finally, we find a qualitative difference between subjects with high and low proxies for cognitive abilities, and argue that this is consistent with a simple model of inattention.

We hope that our findings stimulate further theoretical and empirical work on how players interact in repeated, strategic settings and that our findings will be useful for differentiating intentional strategies from errors in repeated games more generally.

References

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

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.

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 (2022): "Using team discussions to understand behavior in indefinitely repeated prisoner's dilemma games," *American Economic Journal: Microeconomics, forthcoming.*

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 (2019): "Cognitive uncertainty," *National Bureau of Economic Research Working Paper No. 26518*.

Falk, A., A. Becker, T. Dohmen, B. Enke, D. Huffman, and U. Sunde (2018): "Global Evidence on Economic Preferences," *Quarterly Journal of Economics*, 133, 1645–1692.

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

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

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

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

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

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

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

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

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

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

MENGEL, F., S. WEIDENHOLZER, AND L. ORLANDI (2022): "Match Length Realization and Cooperation in Indefinitely Repeated Games," *Journal of Economic Theory*, 200, 105416.

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

NORMANN, H.-T. AND M. STERNBERG (2021): "Human-Algorithm Interaction: Algorithmic Pricing in Hybrid Laboratory Markets," Tech. rep.

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 Experimental Design: Continuation Probabilities and Realizations

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

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

Table A1: Continuation probabilities δ and the number of rounds played for each of the 24 sequences, both treatment orders (one order is just the reverse of the other).

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

Table A2: The distribution of the supergames, split by continuation probabilities δ. The average theoretical and realized supergame durations are 1.99 rounds and 2 rounds, respectively.

B Further Results

B.1 Order Effects

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

We start with the Lab subject pool. While the mean (st. dev.) first round per-subject counts of cooperation in the reverse and long orders are 10.96 (6.48) and 12.10 (4.86), respectively (out of 24), this difference is insignificant ($t = 1.00$, Kolmogorov-Smirnov one-sided $p = 0.278$). The corresponding mean (st.dev.) overall counts are, respectively, 25.52 (9.29) and 22.66 (12.83) (out of 48), with the difference remaining insignificant ($t = 1.28$, Kolmogorov-Smirnov one-sided $p = 0.198$).

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

We now turn to the AMT subject pool. While the mean (st. dev.) first round per-subject counts of cooperation in the reverse and long orders are 11.96 (5.96) and 12.99 (6.16), respectively (out of 24), this difference is insignificant ($t = 1.03$, Kolmogorov-Smirnov one-sided $p = 0.217$). The corresponding mean (st.dev.) overall counts are, respectively, 25.96 (11.21) and 27.38 (11.97) (out of 48), with the difference remaining insignificant ($t = 0.75$, Kolmogorov-Smirnov one-sided $p = 0.410$).

As for the optimal choices, the first round counts are actually lower in the long treatment, with mean (st. dev.) being, respectively, 16.45 (4.45) and 14.53 (4.40), which is significant ($t = 2.68, p = 0.0042$, Kolmogorov-Smirnov one-sided $p = 0.011$). The overall optimal choice counts are, again, lower in the long order treatment (with mean (st.dev.) of 27.12 (9.03) in long order, and 30.71 (9.21) in reverse), and this is significant ($t = 2.40, p = 0.0089$, Kolmogorov-Smirnov one-sided $p = 0.038$.

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

Finding 11. *There is no consistent order effect in either subject sample.*

B.2 Choices to Cooperate vs. Theoretically Optimal Choices

Given the parameters of our design, the theoretically optimal strategy involves choosing to cooperate in all rounds of the 8 supergames where continuation probability $\delta = \{0.67, 0.7\}$, and to defect in all rounds of the other 16 supergames. Thus, perfect theoretically optimal behavior involves exactly 8 counts of cooperation across all first rounds of 24 supergames, and exactly 18 counts of cooperation in the subsequent 24 decisions, amounting to exactly 26 counts of cooperation overall, out of 48 choices (see Figure 2, also Table A2). That is, by design, the theoretically optimal choices should be skewed towards defection initially, since most δs are less than 0.5 and then skewed towards cooperation later on, as it is in the longer games where cooperation is the optimal policy.

Following the previous literature, we look closely at what subjects do in the first round of each supergame (before any knowledge of the opponent's actual play). As Figure B1 shows, only a small fraction of subjects behave according to the equilibrium predictions in the first rounds across all 24 supergames. Specifically, only 2 subjects out of 100 (2% of subjects) in the Lab sample and only 5 subjects out of 149 (3.36% of subjects) in the AMT sample behaved perfectly as predicted for round 1. Furthermore, in both subject pools, no more than 17% of subjects made no more than 3 mistakes out of 24 first round choices (i.e., at least 21 optimal choices out of 24, or 87.5% of choices). The Lab subjects behaved significantly more optimally than the AMT subjects (Kolmogorov-Smirnov one-sided test $D = 0.2197, p = 0.003$), yet this difference is less pronounced at the upper end of optimal choices.

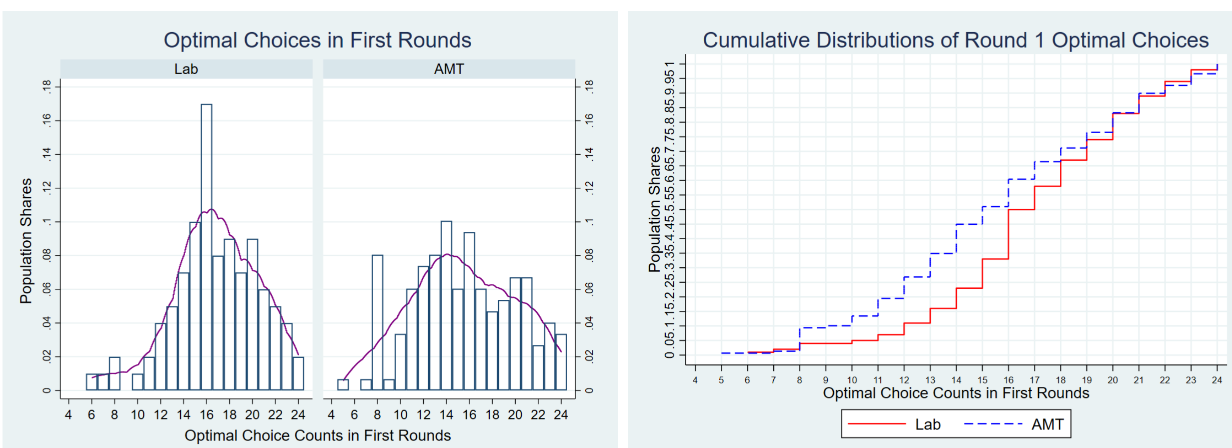

Figure B1: Left: Frequency distributions and kernel densities of per-subject counts of optimal choices in first round choices across all 24 supergames (24 decisions total): Lab (N=100) vs. AMT (N=149) subjects. Right: Both cumulative distributions on the same graph.

Figure B2 provides an analogue of the right panel of Figure 2, separating optimal and suboptimal choices for each continuation probability and each round, and distinguishing dominated cooperation after defecting earlier within the same supergame (CaD).

For each subject pool, Figure B3 presents the frequency distributions of choices to co-

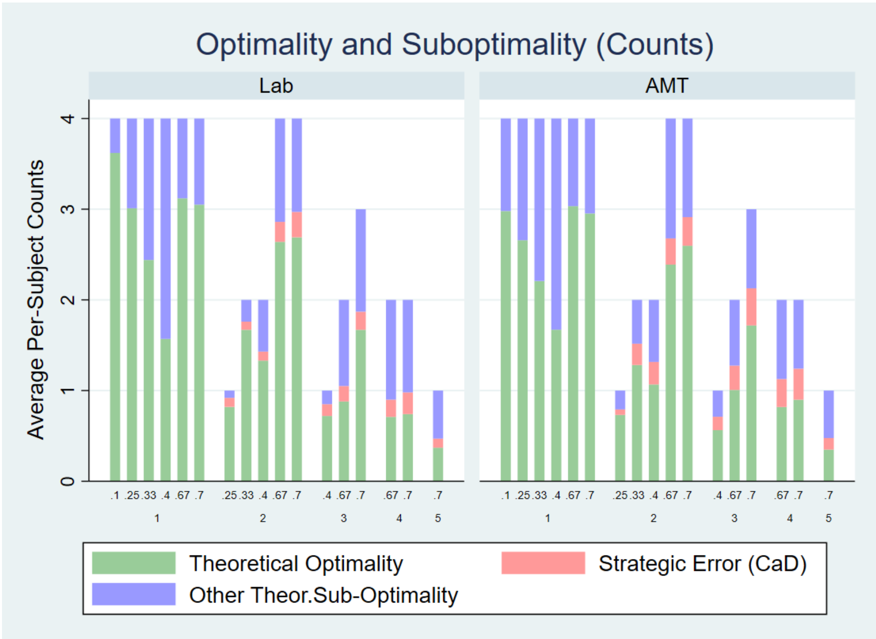

Figure B2: Patterns of cooperation and defection: Average per-subject counts of optimality versus sub-optimality, split by continuation probability δ (the first row of the horizontal axis scale) and by the round number of a sequence (the second row of the horizontal axis scale). The counts of suboptimal choices are divided between the error of CaD and other suboptimal choices. See Figure B2 for the corresponding co-operative and non-cooperative choices. (Lab: 100 subjects, 2,400 supergames, AMT: 149 subjects, 3,576 supergames).

operate (bottom left panels) and of the interpretation of these choices in terms of whether cooperation was the equilibrium response to the Grim Trigger strategy given the continuation probability (top panels). The bottom right panels further provide two-dimensional distributions of the cooperative and optimal choices, where the possible choice combinations are restricted to the polygons delineated by the dashed lines. As the histograms show, subjects in both pools tend to excessively cooperate in the first round of each supergame, far above the theoretical prediction of 8. The mean (st. dev.) count of cooperative choices is 11.53 (5.73) for Lab and 12.47 (6.06) for AMT (with no significant difference across the two pools, see Table B3). As a result, the mean (st.dev.) count of theoretically optimal choices per subject is 16.81 (3.74) for Lab, which is significantly higher than 15.50 (4.52) for AMT. Both pools are prominently short of the theoretical prediction of 24.

Finding 12. *For both subject pools, in the first rounds, subjects cooperate excessively – on average by 44.1% by Lab and by 55.9% by AMT – compared to the theoretical optimum.*

Turning to the overall choice counts, Figure B4 presents the two-dimensional distributions of cooperative and optimal choices and the corresponding marginal distributions for all 48 choices in all 24 supergames. The mean (st. dev.) of the overall count of cooperative choices is only 24.09 (11.24) for Lab, which is significantly lower than the theoretical prediction of 26 (one-sided $t = 1.670, p = 0.046$). Interestingly, the AMT subjects' choice to cooperate are

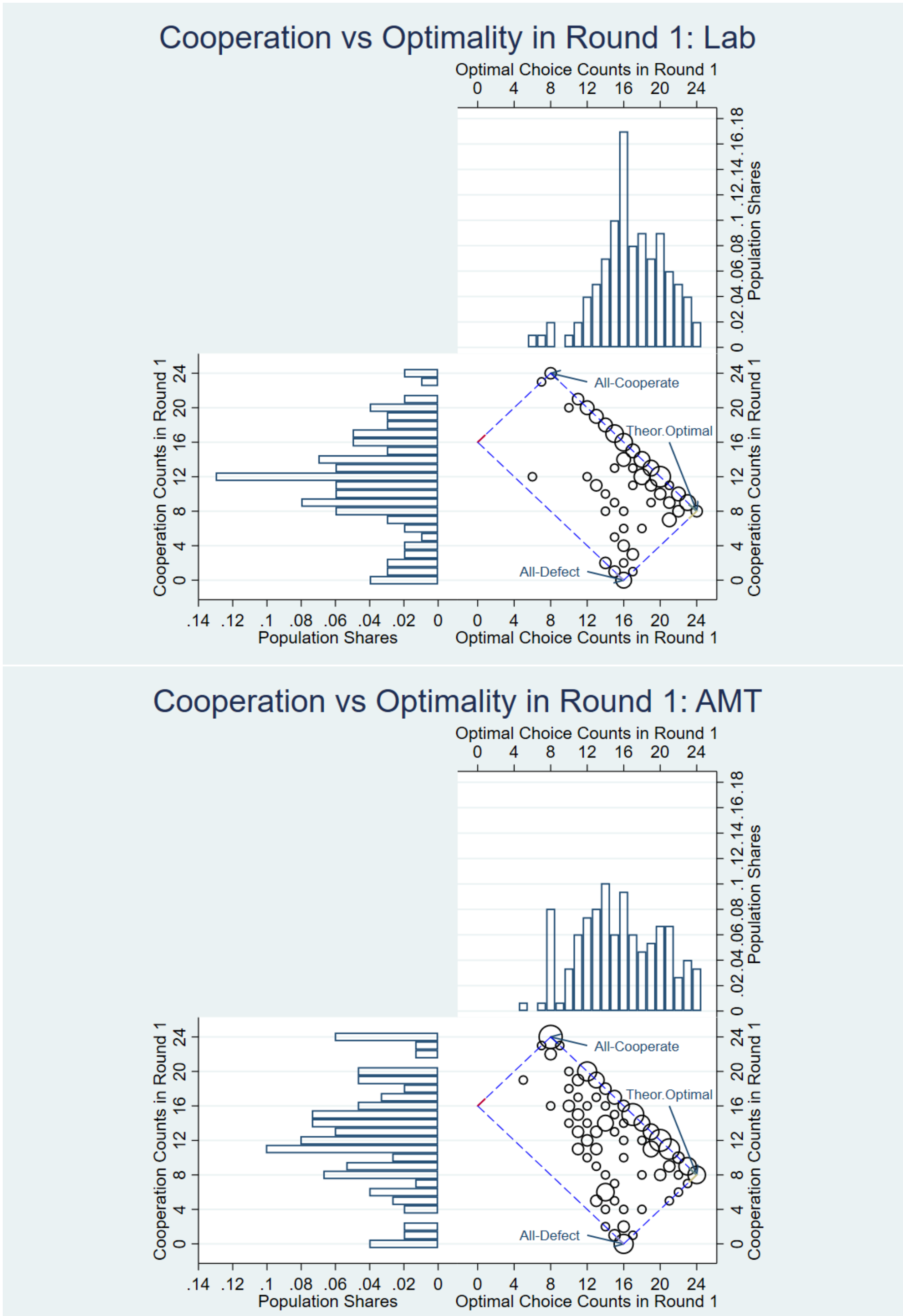

Figure B3: Choices in the first rounds each supergame: Lab (N=100) vs. AMT (N=149) subjects. Two-dimensional distributions of per-subject counts of cooperation and of optimal choices across all 24 supergames, together for the frequency distributions of cooperative (bottom left) and optimal (top right)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

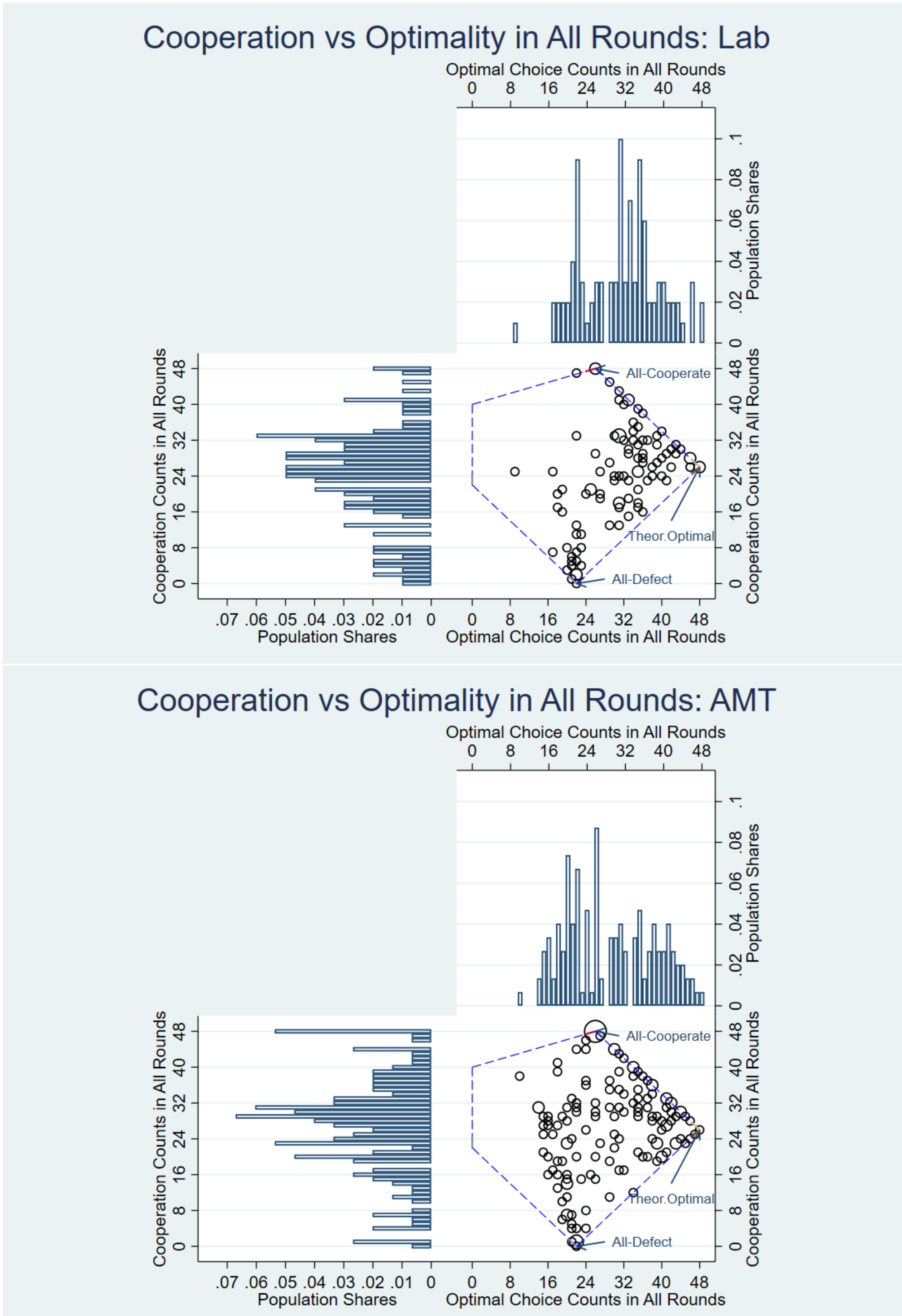

Figure B4: Choices in all rounds: Lab (N=100) vs. AMT (N=149) subjects. Two-dimensional distributions of per-subject counts of cooperation and of optimal choices across all 24 supergames, together for the frequency distributions of cooperative (bottom left) and optimal (top right) choices. Bubble size is proportional to the share of subjects.

26.66 (11.58), and thus on average are not significantly different from the predicted value of 26 (two-sided $t = 0.701, p = 0.4847$). However, for both subject pools, cooperative choices are often sub-optimal – as depicted by the two-dimensional distributions in the bottom right panels for each subject pool in Figure B4. (The shapes of the polygons for the overall choices in the bottom right panels are due to the possibility of dominated CaD choices, described in Section 4.3.) Indeed, for both subject pools, the overall optimal choice counts are significantly short of the theoretical prediction of 48, with a mean (st. dev.) of 31.05 (8.06) for Lab, which is marginally greater than 28.93 (9.27) for AMT (see Figure B3).

In other words, while subjects start by cooperating excessively in the first rounds, this early excessive cooperation in the first rounds is followed by the subsequent defection. The mean (st. dev.) of cooperation counts in the subsequent rounds (given by the difference in the overall and first round cooperation counts) is only 12.56 (6.29) for Lab and 14.19 (6.58) for AMT, both significantly lower than the theoretical prediction of 18 ($p = 0.000$). Note that this is despite the excessive cooperation of 1.82 counts per Lab subject and 2.75 counts per AMT subject on average due to strategic (CaD) errors.

Finding 13. *For both subject pools, compared with the theoretical predictions, on average, subjects cooperate too much at the beginning of supergames with* $\delta < 0.5$ *and stop cooperating too early in supergames with* $\delta > 0.5$*, with only 64.69% for Lab and 60.27% for AMT of all choices being theoretically optimal.*

As Figures B3 and B4 show, there is a significant heterogeneity in subjects' behavior (particularly for AMT), without any clear "representative" pattern. The initial heterogeneity of play in the first rounds in Figure B3 (bottom right panels for each subject pool) is further amplified by the heterogeneous strategies employed by the subjects in the subsequent rounds, depicted in the corresponding panels in Figure B4.

As the bottom right bottom panels of Figure B4 for each subject pool show, there are three similarly sized clusters (about 3-7% of each subject pool) at each of the three corners of the polygon. One can see that there are only two (out of 100) and one (out of 149) subjects who made perfect theoretically optimal choices, located in the far right corners of the polygons for the Lab and AMT pools, respectively. There are further only three and five such subjects, respectively, who made up to three suboptimal choices.

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

Finding 14. *In both subject pools, perfect and near-perfect theoretically optimal behavior is rare, with only 5% of the Lab and 4.69% of the AMT subjects making no more than 3 theoretically sub-optimal choices. These shares are of similar order of magnitude as the shares of subjects who defected no more than 3 times in both pools (4% Lab and 6.71% AMT), and the share of subjects who cooperated no more than 3 times in both pools (5% Lab and 3.36% AMT).*

B.3 Learning

In Figure B5, one can observe an increase in the "end-timing" activities in the both subject pools by comparing those in the first few supergames to that in the last few (further supported by the tests in Table 1). As this same Figure reveals, while the incidence of dominated CaD errors decline over time, they do not disappear entirely. Figure B6 further presents the patterns of intra-supergame play across all 24 sequences, split by the sequence order.

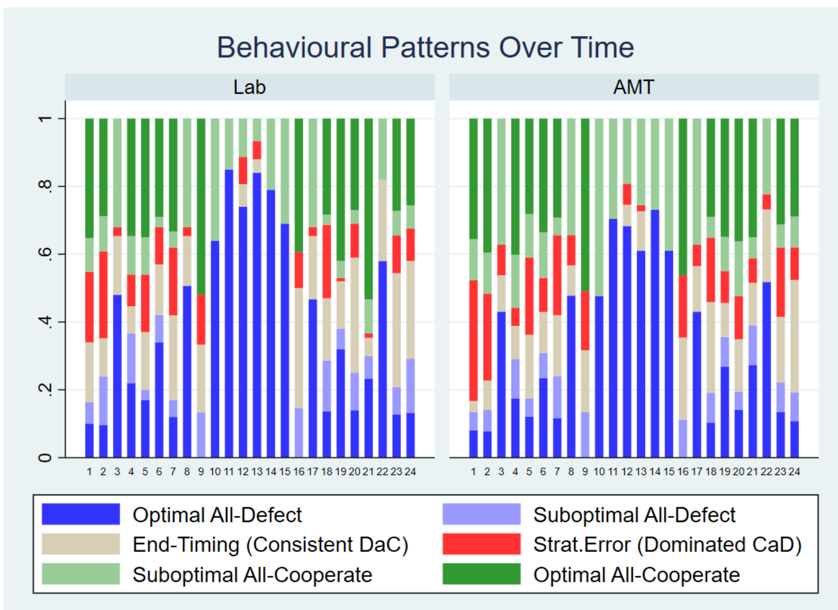

Figure B5: Subject behavior over the sequence of 24 supergames, by subject pool. Both supergame sequence orders are pooled together so a supergame with a given number could involve different continuation probabilities δ and corresponding optimal actions, depending on the sequence order.

B.4 Comparison of Lab and AMT

Table B3 compares the two subject pools, Lab and AMT. Table B4 compares the two subject pools, Lab and AMT, split according to the median CRT7.

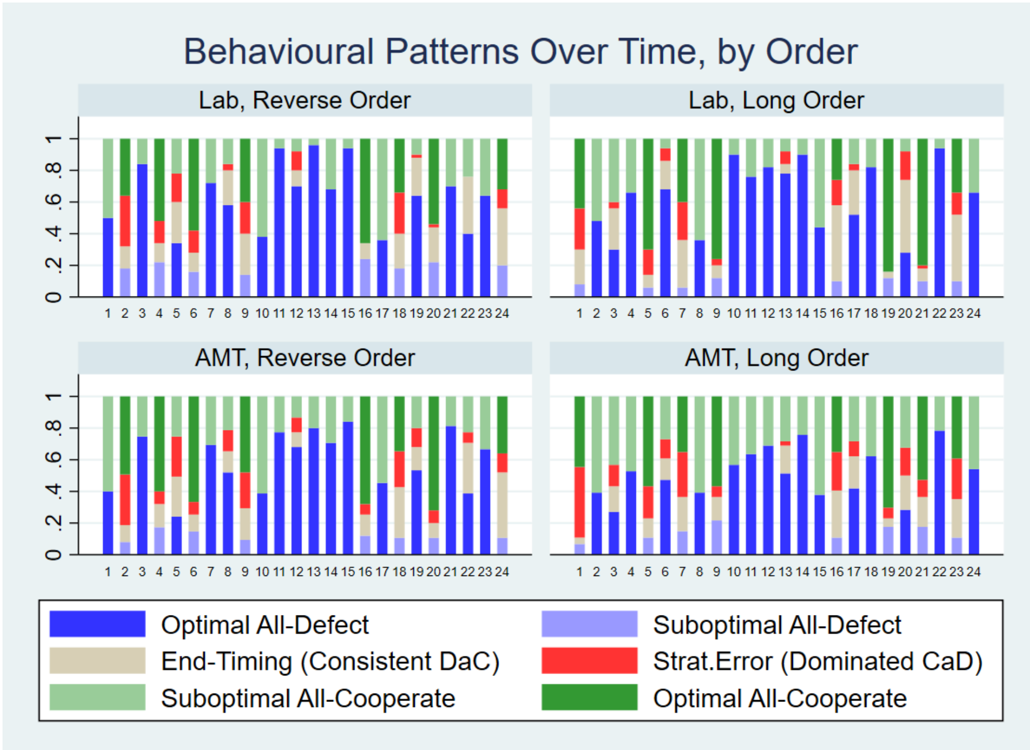

Figure B6: Subject behavior over the sequence of 24 supergames, by sequence order and subject pool.

B.5 Rational Inattention: Further Results

Figure B7 presents the distributions of the two key variables for the inattention model, $CRT7$ and Prediction variable. As Table B3 shows, the means of these two variables are not statistically different between the two subject pools. Table B5 presents the odds corresponding to the regressions of Table 3.

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

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

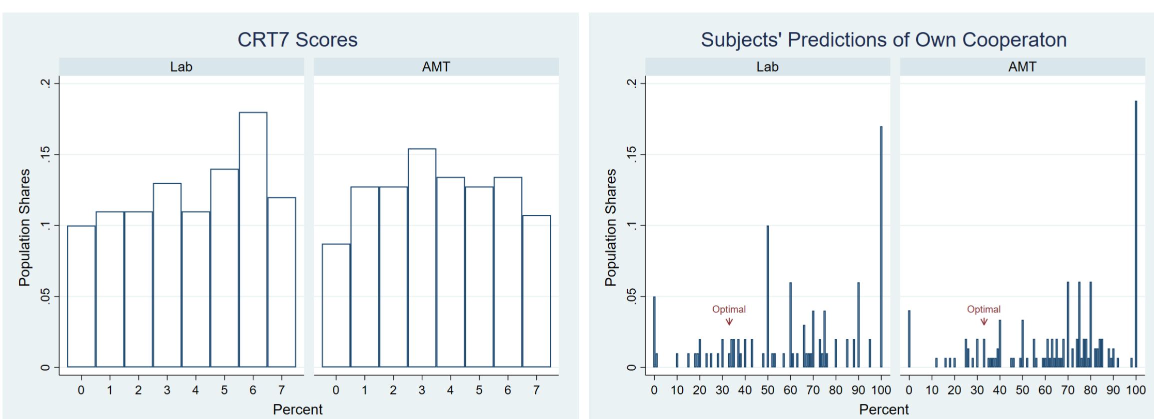

Figure B7: Frequency distributions of CRT7 scores and of the "Prediction" variable (which is the subjects' predictions of the share of their own cooperative choices in Round 1 across all 24 supergames and which is the proxy for subjects' understanding the structure of the game.

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

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

Table B4: For each subject pool: Means and standard deviations of key variables, and t-tests of differences between the means for two pools (all equal variance tests except for Age and Quiz Errors). *df* stands for degrees of freedom or Satterthwaite's degrees of freedom in case of unequal variances for Age and Quiz Errors, *p*value stands for $Pr(|T| > |t|) = 0$. (Significance * 0.10 ** 0.05 *** 0.01 ***† 0.001.)

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

Table B5: Choices to cooperate: mixed-effects probit regressions, odds, robust errors in parentheses. (See Table 3 for the corresponding marginals.) "Supergame" is the supergame number in the sequence of supergames, "Order Long" is a dummy variable for whether the first supergame in the sequence had $\delta = 0.67$, "Prior Defection" is a dummy variable for whether the subject defected in prior rounds of a given supergame, "Prediction" is the subjects' predictions of the share of their own cooperative choices in Round 1 across all 24 supergames (scaled down by 10). (Significance * 0.10 ** 0.05 *** 0.01 ***† 0.001.)

C Appendix: Experimental Questions and Instructions

Personality Questions

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.[30]

Questionnaire

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

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

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

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

CRT questions

Subjects were asked to provide numerical answers to the following cognitive reflection test (CRT) questions.[31]

[30] Taken from Falk et al. (2018).
[31] Based on Toplak et al. (2014) and Ackerman (2014).

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

Repeated PD Game Instructions

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

In each round, you play a game.

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

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

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

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

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

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

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

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

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

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

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

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

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

$0.01 in US dollars, that is, the greater are your point earnings, the greater are your money earnings.

Comprehension quiz

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

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

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

Questions

1. If, in a round, you chose X and the computer program chose X, what is your payoff in points for the round? What is the computer program's payoff?
2. If, in a round, you chose Y and the computer program chose X, what is your payoff in points for the round? What is the computer program's payoff?
3. If, in a round, you chose Y and the computer program chose Y, what is your payoff in points for the round? What is the computer program's payoff?
4. If you have chosen Y in any prior round of the current sequence, what will the computer program choose in the current round of the sequence? Choose: X or Y
5. True or false: At the start of each sequence, you will know exactly how many rounds will be played in the sequence. Choose: True or False
6. True or false: If, in a sequence, the continuation probability is 75%, then you can expect that there will be more rounds in that sequence, on average, than in a sequence with a continuation probability of 25%. Choose True or False

Belief Elicitation

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

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. For instance, the screenshots shown in Figures C1-C2 provide an illustration of the screens that subjects faced in the first round (Figure C1) and in continuation rounds (Figure C2) of the first supergame of the "orderlong" treatment, which had a continuation probability of 0.67 and lasted for 4 rounds. In this illustration, the subject chooses Y (defect) in all 4 rounds and the computer program responds accordingly. Note that subjects were always informed in advance about the computer opponent's decision for each round based on the round number, the history of play and the prescriptions of the Grim trigger strategy. For instance, in round 1 (Figure C1) the subject is instructed: "Since this is the first round of a sequence the computer will always choose X." After the subject chose Y in the first round of Sequence 1, in the second round of the sequence (Figure C2) the subject is instructed: "Based on your choices in previous rounds of this sequence the computer will choose Y".

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

Figure C1: (Top) Start screen for a new sequence. (Middle) Main decision screen for a period in the sequence. (Bottom) Results screen for a period in the sequence.

Sequence 1, round 2

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

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

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

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

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

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

History of Rounds in this Sequence

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

Results of sequence 1, round 4

You chose Y this round.

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

Therefore, your payoff this round is 30.0 points.

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

Next

History of Rounds in this Sequence

Sequence	Chance to Continue	Round	Your choice	Computer's Choice	Payoff
1	0.67	1	Y	X	120.0 points
1	0.67	2	Y	Y	30.0 points
1	0.67	3	Y	Y	30.0 points
1	0.67	4	Y	Y	30.0 points

Figure C2: (Top) Decision screen for a continuation period in the sequence, noting what the robot player will do, based on the history of play. (Bottom) Screen for the final period of a sequence noting that based on the random drawn, the sequence has ended.