UNIVERSITY OF SUSSEX

ECONOMICS

Working Paper Series

No. 86-2016 (*updated June 2017*)

Cooperation in the Finitely Repeated Prisoner's Dilemma

Matthew Embrey[1], Guillaume R. Fréchette[2], Sevgi Yuksel[3]

University of Sussex [1], NYU[2], UCSB[3]

Abstract: More than half a century after the first experiment on the finitely repeated prisoner's dilemma, evidence on whether cooperation decreases with experience–as suggested by backward induction–remains inconclusive. This paper provides a meta-analysis of prior experimental research and reports the results of a new experiment to elucidate how cooperation varies with the environment in this canonical game. We describe forces that affect initial play (formation of cooperation) and unraveling (breakdown of cooperation). First, contrary to the backward induction prediction, the parameters of the repeated game have a significant effect on initial cooperation. We identify how these parameters impact the value of cooperation–as captured by the size of the basin of attraction of Always Defect–to account for an important part of this effect. Second, despite these initial differences, the evolution of behavior is consistent with the unraveling logic of backward induction for all parameter combinations. Importantly, despite the seemingly contradictory results across studies, this paper establishes a systematic pattern of behavior: subjects converge to use threshold strategies that conditionally cooperate until a threshold round; and conditional on establishing cooperation, the first defection round moves earlier with experience. Simulation results generated from a learning model estimated at the subject level provide insights into the long-term dynamics and the forces that slow down the unraveling of cooperation.

Key words: repeated games, prisoner's dilemma, threshold strategies, basin of attraction

JEL classification: C73, C92

We wish to thank the editor, anonymous referees, James Andreoni, Yoella Bereby-Meyer, Russell Cooper, Pedro Dal Bó, Douglas DeJong, Robert Forsythe, Drew Fudenberg, Daniel Friedman, P.J. Healy, Steven Lehrer, Friederike Mengel, John Miller, Muriel Niederle, Ryan Oprea, Louis Putterman, Thomas Ross, Alvin Roth, Andy Schotter, Andrzej Skrzypacz, Charlie Sprenger, Lise Vesterlund, James Walker and Matthew Webb for helpful comments, suggestions, or for making their data sets available; as well as conference and seminar participants at Purdue U., Bonn U., Goethe U., U. of Cologne, LSE, Durham U., the GAMES 2016 conference, MBEES, U. of Michigan, OSU, the Social Dilemmas Conference at Brown U., the Mont Tremblant conference in Political Economy, the Miniconference on Experimental Political Economy at CalTech, the ESA meetings in Fort Lauderdale, the SITE conference and the Behavioral and Experimental Economics class at Stanford, the Workshop on Norms and Cooperation at the U. of Zurich, UC Santa Cruz, and NYU Abu Dhabi. We would further like to thank the National Science Foundation (grant SES-1225779), and the Center for Experimental Social Science for research support. We are responsible for all errors.

1. Introduction

The prisoner's dilemma (henceforth PD) is one of the most extensively studied games in the social sciences. The reason is that the tension at the center of the game–the conflict between what is socially efficient and individually optimal–underlies many interesting interactions, economic and otherwise.[1] Played once, standard equilibrium notions predict the Pareto-dominated, uncooperative outcome. Repeating the game does little to improve the theoretical outlook whenever there is a commonly-known last round; the demands of subgame perfection, where threats to punish uncooperative play must be credible to have bite, result in the unraveling of cooperation via backward induction.

In this paper, we experimentally study the finitely repeated PD to understand the factors that affect (1) the emergence of cooperative behavior; and (2) its possible unraveling with experience. Our results indicate that cooperative behavior in this canonical environment is driven by two behavioral regularities: the role of the value of cooperation and the emergence of threshold strategies. First, we identify a simple-to-compute statistic that captures initial cooperativeness in this game. The statistic neatly summarizes how the parameters of the environment affect the key strategic tension in the game. Importantly, the statistic highlights the role of strategic uncertainty in determining cooperative behavior, and provides a simple measure to assess its impact in different environments. Second, we find evidence for a previously unidentified regularity in learning about strategies. Our results indicate that people learn to use strategies that allow for conditional cooperation early on (creating dynamic game incentives), but switch to defection later (accounting for unraveling). With experience, the defection region grows; and the structure of these strategies provides a backdrop for how backward induction prevails in finitely repeated games. However, it can take time for the full consequences of these strategies to emerge.

Despite more than half a century of research since the first experiment on the PD (Flood 1952), it is difficult to answer whether people learn to cooperate or defect in this game. That is, data from different studies give a seemingly contradictory picture of the evolution of play with experience.[2] Despite the multitude of papers with data on the game, several of which test alternative theories consistent with cooperative behavior, it is still difficult to draw clear conclusions on whether or not subjects in this canonical environment are learning the underlying strategic force identified by the most basic equilibrium concept.

The source of these contradictory results could be the different parameters implemented, in terms of payoffs and horizon, other features of the design, or differences in the analysis. To address this, we collect all previous studies (meeting certain criteria)

[1] Examples include Cournot competition, the tragedy of the commons, team production with unobservable effort, natural resource extraction, and public good provision, to name a few.

[2] For example, Selten & Stoecker (1986) interpret their results to be consistent with subjects learning to do backward induction. They report the endgame effect–the point after which subjects mutually defect–to move earlier with experience. In contrast, Andreoni & Miller (1993) find that behavior moves in the opposite direction; namely, they observe that the point of first defection increases with experience.

and analyze the data within a unified framework.[3] This analysis confirms the apparent contradictory nature of prior results with respect to whether behavior moves in the direction suggested by backward induction. We investigate the topic further with a new experiment.

With respect to the forces that affect initial play (formation of cooperation), and unraveling (breakdown of cooperation) we document the following. For initial play, the parameters of the repeated game have a significant impact on initial cooperation levels, contrary to the prediction of subgame perfection. We confirm that increasing the horizon increases cooperation, in line with a folk wisdom shared by many researchers on how the horizon of a supergame affects play. Namely, that as the horizon increases, cooperation rates increase, and this is attributed, in a loose sense, to the difficulty of reasoning backwards through more rounds.[4]

Our results indicate that the effect of the horizon on cooperation is brought about via a different channel. Increasing the horizon, while keeping the stage-game parameters constant, increases the value of using a conditionally cooperative strategy relative to one that starts out by defection. The trade-off between cooperation and defection can be captured by the size of the basin of attraction of always defect (AD), a simple statistic imported from the literature on infinitely repeated PDs.[5] In a regression analysis of round-one choices in the meta-study, the value of cooperation has significant explanatory power over and above the length of the horizon. The new experiment addresses this point directly by comparing two treatments in which the horizon of the repeated game is varied, but the value of cooperation is kept constant. Round-one cooperation rates remain similar throughout our experiment between these two treatments.

One key new finding is that in our experiment, and in every prior experiment for which we have data, subjects always take time to "learn" to use threshold strategies: strategies that conditionally cooperate until a threshold round before switching to AD. This observation is a crucial part of understanding why prior experiments suggest contradictory patterns with respect to backward induction. Once behavior incompatible with threshold strategies has disappeared, we find consistent evidence in all treatments in our data set

[3] Although we re-analyze the original raw data, rather than collate the results of previous studies, we will refer to this part of our analysis as the meta-study for simplicity.

[4] With *folk wisdom*, we refer to the common conception that cognitive limitations play an important role in explaining divergence from equilibrium behavior in games involving unraveling arguments. In the context of finitely repeated PD, Result 5 of Normann & Wallace (2012) is an example of prior experimental evidence suggesting a positive correlation between cooperation rates and the horizon. In the context of speculative asset market bubbles, Moinas & Pouget (2013) show that increasing the number of steps of iterated reasoning needed to rule out the bubble increases the probability that a bubble will emerge. We can also point to multiple papers using the level-k model to explain behavior in the centipede game, which, as we discuss in Section 7 and Online Appendix A.1, is closely connected to the finitely repeated PD (see Kawagoe & Takizawa (2012), Ho & Su (2013), Garcia-Pola et al. (2016)). In a recent paper, Alaoui & Penta (2016) present a model of endogenous depth of reasoning that can account for how payoff structure affects the degree to which unraveling is observed in this class of games.

[5] These observations can be found in Dal Bó & Fréchette (2011). See, also, Blonski et al. (2011), who provide an axiomatic basis for the role of risk dominance in this context.

that the round of first defection moves earlier with experience. However, early behavior typically involves multiple switches between cooperation and defection, and, thus, learning to play threshold strategies results in a decrease in the rate of early defections. The speed at which each of these two opposing forces happen–which varies with the payoffs and the horizon of the game–make the combined effect look as though subjects either behave in line with learning backward induction or not.

Although these forces imply unraveling of cooperation in the long run, we find that this process can be very slow. Hence, to complement our results, we estimate a subject-level learning model, and use the estimates to generate simulations of long-run behavior. Our simulations suggest that cooperation rates may remain non-negligible even after ample experience in the case of parameter constellations conducive to high levels of initial cooperation.[6] The estimation of this learning model also allows us to see the evolution of the expected value of various strategies. This helps clarify why unraveling is slower in some treatments than in others. In addition, simulations under counterfactual specifications reveal that the stage-game parameters, rather than the variation in how subjects "learn" across treatments, explain variations in the speed of unraveling.

2. Theoretical Considerations and Literature

The PD is a two-person game in which each player simultaneously chooses whether to cooperate (C) or defect (D), as shown in the left panel of Figure 1(a). If both players cooperate, they each get a reward payoff R that is larger than a punishment payoff P, which they would get if they were both to defect. A tension results between what is individually rational and socially optimal when the temptation payoff T (defecting when the other cooperates) is larger than the reward, and the sucker payoff S (cooperating when the other defects) is smaller than the punishment.[7]

In this case, defecting is the dominant strategy in the stage-game and, by backward induction, always-defect is the unique subgame-perfect equilibrium strategy of the finitely-repeated game.

	C	D
C	R, R	S, T
D	T, S	P, P

(a) Payoff Matrix

	C	D
C	$1, 1$	$-\ell, 1+g$
D	$1+g, -\ell$	$0, 0$

(b) Normalized Matrix

Figure 1: The Prisoner's Dilemma

[6] Even in these cases, simulation results show a slow but continued decline in cooperation.

[7] In addition, the payoff parameters can be restricted to $R > \frac{S+T}{2} > P$. The first inequality ensures that the asymmetric outcome is less efficient than mutual cooperation. The second inequality, which has been overlooked in the literature but recently emphasized by Friedman & Sinervo (2016), implies that choosing to cooperate always improves efficiency. These conditions create a sharp conflict between social efficiency and individual optimality.

One of the earliest discussions of the PD included a small-scale experiment. Dresher and Flood conducted that experiment in 1950 using two economists as subjects (reported in Flood (1952)). That experiment, and others that followed, found positive levels of cooperation despite the theoretical prediction to the contrary. An early paper to offer an explanation for this phenomenon is due to the *gang of four*: Kreps et al. (1982) showed that incomplete information about the type of the other player (either what strategies they can play or their *true* payoffs) can generate cooperation for a certain number of periods in equilibrium. Alternatively, Radner (1986) proposed the concept of epsilon-equilibria–in which agents are content to get *close* to the maximum equilibrium payoffs–and showed that cooperation can arise as part of an equilibrium strategy. Other possibilities that were later explored include learning and limited forward reasoning (see, for example, Mengel (2014*a*), Mantovani (2016), and the references therein). Moving beyond the standard paradigm, social preferences for fairness, altruism or efficiency can also generate cooperation in this game. Although our meta-analysis and experiment are not designed to distinguish between these theories, they provide a backdrop for how cooperation can arise in this environment. Our purpose in this paper is *not* to test these theories directly, but, rather, to take a step back and identify the main forces observed in the data that affect when and how cooperation emerges. We postpone discussion of our results regarding these theories to the final section.

Much of the early experimental literature on the repeated PD came from psychology. That literature is too vast to be covered here, but typical examples are Rapoport & Chammah (1965), Lave (1965), and Morehous (1966). These papers are concerned mainly with the effect of the horizon, the payoffs, and the strategies of the opponent. Some of the methods (for payments, for instance), the specific focus (often horizons in the hundreds of rounds), and the absence of repetition (supergames are usually played only once) limit what is of interest to economists in these studies.

Studies on the finitely repeated PD also have a long history in economics.[8] Online Appendix A.1 provides an overview of the papers on the topic.[9] More specifically, we

[8]Mengel (2014*b*) presents a meta-study that covers more papers and also supplements the existing literature with new experiments. Different from our work, the paper focuses mainly on comparing results from treatments where subjects change opponents after each play of the stage-game (stranger matching) to results from treatments where subjects play a finitely repeated PD with the same opponent (partner matching). For this reason, the meta-study includes more treatments in which the supergame is not repeated, and the new experiments all involve the play of a single supergame. The paper finds that "risk"–an index based on the sucker and punishment payoffs–explains more of the variation in cooperation under stranger matching, whereas "temptation"–an index based on the temptation and reward payoffs–explains more of the variation under partner matching. Thus, the paper is not intended to consider whether behavior moves in the direction of backward induction or to study the impact of experience more generally. Despite these differences, the main conclusion of the paper emphasizing the importance of the stage game parameters, and specifically highlighting how the "risk" and "temptation" parameters can be interpreted to capture the strength of different forces that affect cooperation in environments with strategic uncertainty, is consistent with our results.

[9]Since our interest lies in the emergence and breakdown of cooperation and the role of experience, we focus only on implementations that include an horizon for the repeated game of two or more rounds and

cover all published papers (that we could find) with experiments that include a treatment in which subjects play the finitely repeated PD and in which this is performed more than once.[10,11]

Overall, these papers give us a fragmented picture of the factors that influence behavior in the finitely repeated PD. Most papers are designed to study a specific feature of the repeated game. However, if we try to understand the main forces that characterize the evolution of behavior, it is difficult to draw general conclusions. For instance, the evidence is mixed with respect to whether or not subjects defect earlier with experience. There is evidence consistent with unraveling (experience leading to increased levels of defection by the end of a repeated game), as well as evidence pointing in the opposite direction (mean round to first defection shifting to later rounds with experience).

3. The Meta-Study

The meta-study gathers data from five prior experiments on the finitely repeated PD. Note that we do not rely simply on the results from these studies, but also use their raw data.[12] The analysis includes 340 subjects from 15 sessions with variation in the stage-game parameters and the horizon of the supergame.

To facilitate the comparison of data from disparate experimental designs and to reduce the number of parameters that need to be considered, the payoffs of the stage-game are normalized so that the reward payoff is one and the punishment payoff is zero. The resulting stage-game is shown in the right panel of Figure 1(b), where $g = (T-P)/(R-P) - 1 > 0$ is the one-shot gain from defecting, compared to the cooperative outcome, and $\ell = -(S-P)/(R-P) > 0$ is the one-shot loss from being defected on, compared to the non-cooperative outcome.

have at least one re-matching of partners.

[10]Several recent papers (Mao et al. (2017); Schneider & Weber (2013); Kagel & McGee (2016); Cox et al. (2015); Kamei & Putterman (2015)) study heterogeneity in cooperative behavior and the role of reputation building in the finitely repeated PD. In Online Appendix A.1, we also include a discussion of these papers.

[11]A related game that has been extensively studied in economic experiments is the linear voluntary contributions mechanism (VCM), often referred to as the public goods game. A two player linear VCM where each player has two actions corresponds to a special case of the PD. Using the notation defined in the next section, a binary two players linear VCM is a PD with $g = \ell$. However, few experiments involve repetitions of finitely repeated linear VCMs (with rematching between each supergame): Andreoni & Petrie (2004, 2008), Muller et al. (2008), and Lugovskyy et al. (2017). Andreoni & Petrie (2004, 2008) do not consider the effect of experience on behavior. Hence, results on the evolution of play in that game are few, but certainly Muller et al. (2008) and Lugovskyy et al. (2017) taken together suggest that behavior changes with the stage game and with the amount of experience. To what extent these changes are similar or different from the ones documented in the current study is not yet clear.

[12]Online Appendix A.2 provides more details on the included studies: henceforth, Andreoni & Miller (1993) will be identified as AM1993, Cooper, DeJong, Fosythe & Ross (1996) as CDFR1996, Dal Bó (2005) as DB2005, Bereby-Meyer & Roth (2006) as BMR2006, and Friedman & Oprea (2012) as FO2012.

3.1 The Standard Perspective

Prior studies have focused mostly on cooperation rates, often with particular attention to average cooperation, cooperation in the first round, cooperation in the final round, and the round of first defection. Thus, we first revisit these data using a uniform methodology while keeping the focus on these outcome variables–what we refer to as the *standard perspective*. Table 1 reports these statistics for each treatment. They are sorted from shortest to longest horizon and from largest to smallest gain from defection.

Table 1: Cooperation Rates and Mean Round to First Defection

				Cooperation Rate (%)						Mean Round to First Defection	
				Average		Round 1		Last Round			
Experiment	H	g	ℓ	1	L	1	L	1	L	1	L
DB2005	2	1.17	0.83	14	13	18	14	10	11	1.21	1.20
within	2	0.83	1.17	25	9	32	13	17	5	1.42	1.14
subject	4	1.17	0.83	33	20	44	32	25	8	1.99	1.58
	4	0.83	1.17	31	22	37	34	20	12	1.76	1.61
FO2012	8	4.00	4.00	33	33	43	67	23	3	2.27	3.53
within	8	2.00	4.00	38	34	43	63	30	3	2.77	3.67
subject	8	1.33	0.67	40	48	43	73	37	3	2.83	4.43
	8	0.67	0.67	44	69	50	87	30	23	3.10	6.07
BMR2006	10	2.33	2.33	38	66	61	93	22	7	3.19	7.39
AM1993	10	1.67	1.33	17	47	36	86	14	0	1.50	5.50
CDFR1996	10	0.44	0.78	52	57	60	67	20	27	4.63	5.53

Notes: First defection is set to Horizon + 1 if there is no defection. 1: First Supergame; L: Last Supergame.

The first observation that stands out from Table 1 is that, with both inexperienced and experienced subjects, the horizon length (H) and gain from defection (g) organize some of the variation observed in cooperation rates. Cooperation rates increase with the length of the horizon, and decrease with the gain.[13] In this sense, there seems to be some consistency across studies.[14]

Focusing on factors that interact with experience to affect play, the horizon of the repeated game appears to play an important role. Note that the average cooperation rate always increases with experience when the horizon is long ($H > 8$) and always decreases with experience when it is short ($H < 8$). Similarly, the mean round to first defection statistic decreases with experience only if the horizon is very short ($H \leq 4$).

[13]Statistical significance of the patterns reported here are documented in Online Appendix A.2. Tests reported in the text are based on probits (for binary variables) or linear (for continuous variables) random effects (subject level) regressions clustered at the paper level for the meta and the session level for the new experiment. Exceptions are cases in which tests are performed on specific supergames, where there are no random effects. Clustering is used as a precaution against paper or session specific factors that could introduce un-modelled correlations (see Fréchette 2012, for a discussion of session-effects). Two alternative specifications are explored to gauge the robustness of the results in Appendix A.4. The different specifications do not change the main findings.

[14]The results of AM1993 look slightly different from those of other studies with similar parameters, but the differences become less pronounced with experience. Furthermore, the AM1993 data consist of a single session and, thus, it can be expected to be noisier.

Other aspects of behavior that previous studies have focused on are round-one and last-round cooperation rates. The horizon of the repeated game and the gain from defection appear to play a role in how these measures evolve with experience. Figure 2 traces the evolution of these cooperation rates over supergames separated by horizon and payoffs. In most treatments, last-round cooperation rates are close to zero or reach low levels quickly.[15] The evolution of round-one cooperation rates depends on the horizon. With $H = 2$, cooperation rates in round one are close to zero, and when $H = 4$, they are low and decreasing, though at a negligible rate when the gain from defection is small. The round-one cooperation rate moves in the opposite direction as soon as the horizon increases further. With both $H = 8$ and $H = 10$, round-one cooperation increases with experience.

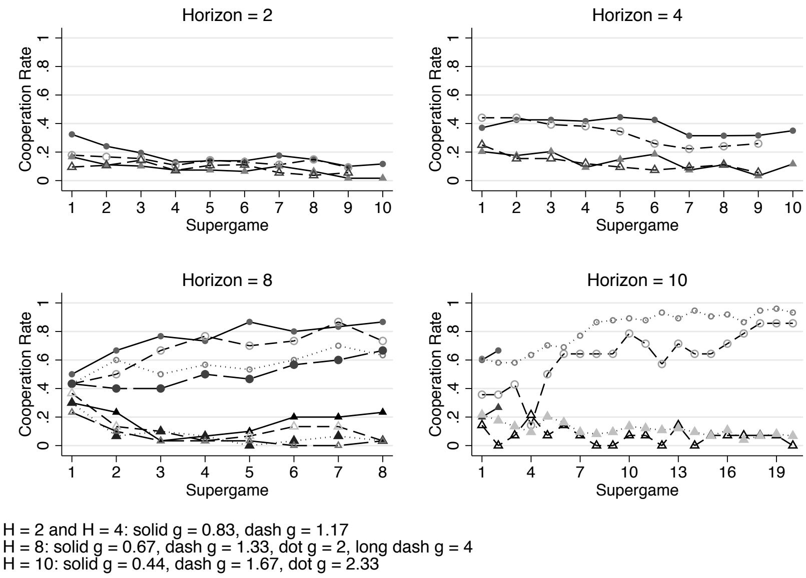


Figure 2: Cooperation Rates: Round One (Circles) and Last Round (Triangles)

3.2 The Value of Cooperation and Round-One Choices

One consistent result to emerge from the *standard perspective* is that average cooperation and round to first defection increase with the horizon. Both observations are consistent with subjects having difficulty–or believing that their partners are having difficulty–making more than a small number of steps of backward induction. However, if the stage-game is kept constant, increasing the horizon also increases the difference in

[15]The decline in cooperation in the last round could be due to multiple factors. If cooperation is driven by reciprocity, the decline could be associated with more pessimistic expectations about others' cooperativeness in the last round. Alternatively, if cooperation is strategic, the decline could be associated with the absence of any future interaction with the same partner. Reuben & Seutens (2012) and Cabral et al. (2014) use experimental designs to disentangle these two forces and find cooperation to be driven mainly by strategic motives.

the value of joint cooperation versus joint defection. Cooperation becomes more valuable since more rounds generate the higher payoffs from joint cooperation. On the other hand, the risk associated with being defected on does not change: when using a conditionally cooperative strategy, there is, at most, one round in which a player can suffer the sucker payoff, irrespective of the length of the horizon. Hence, the value increases but the risk does not.

Experiments on the *infinitely* repeated PD suggest that subjects react to changes in the stage-game payoffs and the discount factor according to how they affect the value of cooperation. However, it is not the case that the value of cooperation, as captured by cooperation being subgame perfect, predicts on its own whether or not cooperation emerges. The decision to cooperate seems to be better predicted by the size of the basin of attraction of always defecting–henceforth $sizeBAD$–against the grim trigger strategy (Dal Bó & Fréchette (2011)).[16] Hence, the strategic tension is simplified by focusing on only two extreme strategies: grim trigger and AD. Assuming that these are the only strategies considered, $sizeBAD$ is the probability that a player must assign to the other player playing grim so that he is indifferent between playing grim and AD.

This measure can be adapted for the finitely repeated PD and used to capture the value-risk trade-off of cooperation. In this case, it can be calculated directly as:[17]

$$sizeBAD = \frac{\ell}{(H-1)+\ell-g}.$$

Values close to one suggest that the environment is not conducive to supporting (non-equilibrium) cooperation since a very high belief in one's partner being conditionally cooperative is required. The opposite is true if the value is close to zero. As can be seen, $sizeBAD$ is increasing in g and ℓ, but decreasing in H.

Table 2 reports the results of a correlated random effects probit investigating the correlation between round-one choices and design parameters such as the $sizeBAD$, stage-game payoffs, and the horizon.[18] The first specification controls for the normalized stage-game parameters, g and ℓ, and H.[19] As can be seen, both g and H have a significant effect

[16]The grim trigger strategy first cooperates and cooperates as long as both players have always cooperated; and defects otherwise.

[17]In the finitely repeated PD, AD (Grim) results in a payoff of 0 ($-\ell$) against a player following AD or a payoff of $1+g$ (H) against a player following Grim. $sizeBAD$ corresponds to the probability, p, assigned to the other player playing Grim that equalizes the expected payoff associated with either strategy, given by $pH-(1-p)\ell = p(1+g)$. Unlike in infinitely repeated games, this calculation is not the best-response to such a population, since adapting the grim strategy to defect in the last round would always achieve a higher payoff.

[18]Note that although there is variation in $sizeBAD$, it is highly correlated with the horizon in these treatments (see Online Appendix A.2).

[19]In addition, there are four regressors that interact the supergame with dummies for the horizon and a regressor for the choice in round one of the opponent in the preceding supergame; and the mean of the random effects is allowed to vary with a subject's choice in round one of supergame one. We report this specification, as it makes the effect of the regressors of interest easy to read. However, a more complete

Table 2: Marginal Effects of Correlated Random Effects Probit Regression of the Probability of Cooperating in Round One

	(1)	(2)
g	-0.04^{***} (0.009)	-0.03^{***} (0.006)
ℓ	-0.02^{***} (0.005)	0.00 (0.005)
Horizon	0.03^{***} (0.004)	0.01 (0.005)
sizeBAD		-0.24^{***} (0.025)
Observations	5398	5398

Notes: Standard errors clustered (at the study level) in parentheses. ***1%, **5%, *10% significance.
Additional controls include experience variables (supergame interacted with Horizon dummies) and choice history variables (whether the player cooperated in the first supergame and whether the player they were matched with cooperated in round one of the last supergame).
Complete results reported in Online Appendix A.2.

on round-one choices, and in the predicted direction: when there is more to be gained from defecting if the other cooperates and when the horizon is short, it is less likely that a subject will make a cooperative round-one choice. The second specification includes the $sizeBAD$ statistic. The new variable–which is a non-linear combination of g, ℓ, and H–has a significant negative impact on cooperation, as would be expected if the value of cooperation considerations outlined above were important. Furthermore, the effect of the design parameters seems to be accounted for, in large part, by the $sizeBAD$ variable, with ℓ and g having a smaller magnitude.[20]

In summary, by combining data sets from prior studies, we are able to investigate the impact of stage-game and horizon parameters on cooperation, as well as the interaction of these with experience. However, a clear understanding of behavior is still not possible using the meta-analysis alone. First, since the majority of experiments do not vary parameters within their designs, much of the variation comes from comparing across studies, where many other implementation details vary. Second, the payoff parameters are, for the most part, constrained to a small region, resulting in a high correlation between the size of the basin of attraction of AD and the length of the supergame. Finally, very few of the studies give substantial experience to subjects.

4. The Experiment

To address the issues identified in the meta-study, we designed and implemented an experiment that separates the horizon from other confounding factors, systematically varying the underlying parameters within a unified implementation. In addition, the new

specification would interact supergames with dummies, not only for each H, but also for all g, ℓ, and H. Those results are presented in Online Appendix A.2, but interpreting the effect of a change in the regressors of interest is complicated by the complex interactions with supergames.

[20] Another way to asses to what extent $sizeBAD$ captures the relevant variation is to compare a measure of fit between specification (1), which does not include $sizeBAD$, and an alternative specification that does include $sizeBAD$, but excludes g, ℓ, and $Horizon$. To give a sense of this, we estimate these two specifications using random effects regressions and report the R^2. It is 0.34 without $sizeBAD$ and 0.35 with $sizeBAD$ but without g, ℓ, and $Horizon$.

sessions include many more repetitions of the supergame than are commonly found in prior studies. The experiment is a between-subjects design with two sets of stage-game payoffs and two horizons for the repeated game: a 2×2 factorial design.

The first treatment variables are the stage-game payoffs. In the experiment, participants play one of two possible stage-games that differ in their temptation and sucker payoffs, as shown in Figure 3.[21] The payoffs when both players cooperate or both players defect are the same in both stage-games. As a consequence, the efficiency gain from cooperating is the same in both sets of parameters: 31%.

	C	D
C	$51, 51$	$5, 87$
D	$87, 5$	$39, 39$

(a) Difficult PD

	C	D
C	$51, 51$	$22, 63$
D	$63, 22$	$39, 39$

(b) Easy PD

Figure 3: Stage-Games in the Experiment

The first stage-game is referred to as the difficult PD, since the temptation payoff is relatively high and the sucker payoff low, while the second stage-game is referred to as the easy PD, for the opposite reason. In terms of the normalized payoffs described in Section 3, the (g, ℓ) combination is $(2.5, 2.83)$ for the difficult PD and $(1, 1.42)$ for the easy PD. As shown in Online Appendix A.2, the easy parameter combination is close to the normalized parameter combinations of a cluster of prior experiments from the meta-analysis. The difficult parameter combination has larger values of both g and ℓ than has been typically implemented.

The second treatment variable is the horizon of the repeated game. To systematically vary the number of steps of reasoning required for the subgame perfect Nash equilibrium prediction, we implement short-horizon and long-horizon repeated games for each stage-game. In the shorter horizon, the stage-game is repeated four times and in the longer horizon, eigth times. Combining the two treatment variables gives the four treatments: D4, D8, E4 and E8, where D/E refer to the stage-game, and 4/8 to the horizon.[22]

Following the intuition that cooperation is less likely in the difficult stage-game, and that the unraveling of cooperation is less likely with a longer horizon, cooperation is expected to be higher as one moves to an easier stage-game and/or to a longer horizon. However, the comparison between D8 and E4 is crucial, as it mixes the difficult stage-game parameters with the longer horizon and vice-versa. Indeed, the parameters have been designed such that this mix gives precisely the same *sizeBAD* in both treatments. Hence, if a longer horizon increases cooperation beyond its impact through the changes

[21] Payoffs are in experimental currency units (ECU) converted to Dollars at the end of the experiment.

[22] The parameters were selected such that, based on the meta-study, we could expect that in the short run, the aggregate statistics would move in the direction of backward induction, at least for D4, and in the opposite direction, at least for E8. Other considerations were that the two values of H did not result in sessions that would be dramatically different in terms of time spent in the laboratory.

in the value of cooperation captured by *sizeBAD*–possibly because there are more steps of iterated reasoning to be performed–treatment D8 should generate more round-one cooperation than treatment E4.[23]

4.1 Procedures

The experiments were conducted at NYU's Center for Experimental Social Science using undergraduate students from all majors, recruited via e-mail.[24] The procedures for each session were as follows: after the instruction period, subjects were randomly matched into pairs for the length of a repeated game (supergame). In each round of a supergame, subjects played the same stage-game. The length of a supergame was finite and given in the instructions so that it was known to all subjects. After each round, subjects had access to their complete history of play up to that point in the session. Pairs were then randomly rematched between supergames.

A session consisted of 20 or 30 supergames and lasted, on average, an hour and a half.[25] At the end of a session, participants were paid according to the total amount of ECUs they earned during the session. Subjects earned between \$12.29 and \$34.70. Three sessions were conducted for each treatment.[26] Throughout, a subject experienced just one set of treatment parameters: the stage-game payoffs and the supergame horizon.

4.2 The Standard Perspective in the Experiment

Table 3 provides a summary of the aggregate cooperation rates across treatments. For each treatment, the data are split into two subsamples: supergames 1-15 and supergames 16-30. Four measures of cooperation are listed: the cooperation rate over all rounds, in the first round and in the last round, as well as the mean round to first defection. The first observation is that our treatments generate many of the key features observed across the different studies of the meta-analysis. This can be seen most clearly with respect to first-round cooperation and mean round to first defection. First-round cooperation in the long-horizon treatments is significantly higher in later than in earlier supergames. Although none of the differences are significant, the mean round to first defection shows the same pattern as initial cooperation.

---

[23]Other statistics have also been found to correlate with cooperation rates in finitely repeated PDs, as in Mengel (2014*b*). Murnighan & Roth (1983) also discuss ten indexes that can be applied to these games. One key difference between these and *sizeBAD* is that they depend only on the payoffs of the stage-game. Hence, their predictions for our E4 and E8 treatments, as well as for D4 and D8, are the same.

[24]Instructions were read aloud. Subjects interacted solely through computer terminals. Instructions are provided in Online Appendix A.6. The computer interface was implemented using zTree (Fischbacher (2007)).

[25]The first session for each treatment consisted of 20 supergames. After running these, it was determined that the long-horizon sessions were conducted quickly enough to increase the number of supergames for all treatments. Consequently, the second and third sessions had 30 supergames. The exchange rate was also adjusted: 0.0045 \$/ECU in the first session and 0.003 \$/ECU in the second and third sessions.

[26]More details about each session are provided in Online Appendix A.3.

Table 3: Cooperation Rates: Early Supergames (1–15) vs Late Supergames (16–30)

	All rounds			Round 1			Last Round			First defect		
Treatment	1–15		16–30	1–15		16–30	1–15		16–30	1–15		16–30
D4	15.4	$>^{**}$	9.0	29.1	$>$	19.5	4.1	$>^{**}$	3.2	1.5	$>$	1.3
D8	34.6	$>$	33.2	49.3	$<^{***}$	57.1	7.9	$>^{***}$	4.0	2.8	$<$	3.1
E4	28.0	$>^{***}$	21.2	49.0	$>$	45.2	10.4	$>^{***}$	3.8	1.9	$>^{**}$	1.7
E8	60.1	$>^{***}$	55.2	79.7	$<^{***}$	88.2	9.0	$>^{***}$	3.0	5.3	$\sim$	5.3
All	37.8	$>^{***}$	33.5	51.1	$<$	51.6	8.0	$>^{***}$	3.6	2.8	$>$	2.7

Notes: Significance reported using subject random effects and clustered (session level) standard errors. ***1%, **5%, *10%.

For the average over all rounds, cooperation is lower during the later supergames and significantly so for the easy stage-game. This observation is in contrast to some of the studies in our meta data that find that the average cooperation rate increases with experience. However, subjects played 30 supergames in our experiment, which is substantially more than in any of the studies in our meta data. To provide a more complete comparison with the studies from the meta-analysis, Table 4 reports the cooperation rate at the supergames corresponding to the length of the various studies in our meta data, as well as in our first and last supergame. For E8, there is a clear increase in cooperation rates early on, followed by a decline. Indeed, the parameters used in this treatment are the closest to the studies in which aggregate cooperation is found to be increasing with experience–namely, those with a longer horizon. The non-monotonicity observed in this treatment, with respect to the evolution of aggregate cooperation rates with experience, suggests that experimental design choices, such as the number of repetitions of the supergame in a session, can significantly alter the type of conclusions drawn from the data.

Table 4: Cooperation Rate for All Rounds in Supergames 1, 2, 8, 20 and 30

	Supergame				
Treatment	1	2	8	20	30
D4	31.5	21.0***	12.5***	11.5***	6.6***
D8	36.3	36.3	36.8	35.6	32.6**
E4	28.2	29.8	30.2	19.4	20.0*
E8	47.6	53.8**	61.4***	51.4	51.6

Notes: Statistical test is for difference from Supergame 1.
For E8, decline from Supergame 8 to 30 is significant at the 1% level.
Significance reported using subject random effects and clustered (session level) standard errors. ***1%, **5%, *10%.

Figure 4 provides some insight into the underlying forces generating the differences in the aggregate results documented above. The figure shows the rate of cooperation in each round, averaged over the first five supergames, supergames 13 to 17 and the last five supergames. In the long-horizon treatments, especially in E8, cooperation in early rounds increases with experience. The line associated with the first five supergames lies below the one associated with the last five for early rounds. This pattern contrasts with the short-horizon sessions in which the first-five average is at least as large as the last-five average. For later rounds, in all treatments, cooperation in the last five supergames is lower than in the first five. With a short horizon, cooperation rates fall quickly after the first round. When the horizon is long, this decline does not happen until later, coming after six rounds in early supergames and after four or five rounds in later supergames.[27]

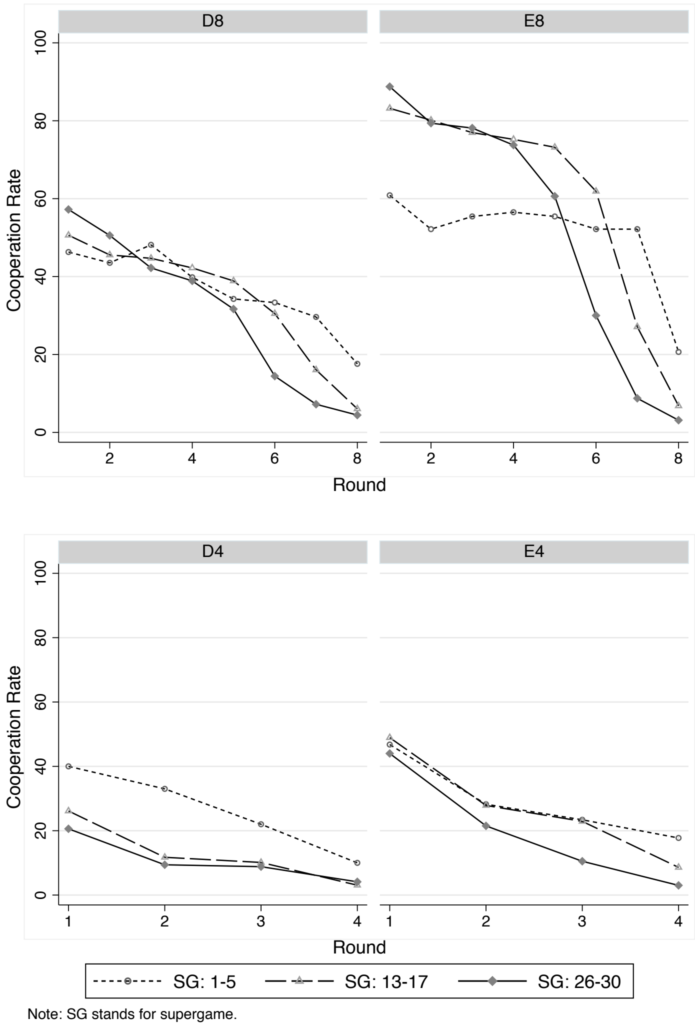


Figure 4: Cooperation Rate by Round Separated in Groups of Five Supergames

4.3 Determinants of Initial Cooperation

Figure 5 shows, for each treatment, the round-one cooperation rate by supergame. The treatments generate very different dynamics with respect to how initial cooperation evolves with experience, again emphasizing how critical the parameters of the stage-game and the horizon can be in determining the evolution of play. The D4 treatment

[27]Online Appendix A.3 includes a complete pairwise comparison of the cooperation measures by treatment.

shows decreasing initial cooperation rates, whereas the E8 treatment shows a notable increase over supergames. The cooperation rates for D8 and E4 look very similar. Neither treatment suggests a trend over supergames, and cooperation rates stay within the 40-60% range for the most part. In fact, round-one cooperation rates are not statistically different across the two treatments. Moreover, cooperation rates in supergames one and 30 are statistically indistinguishable between the two treatments.[28]

Remember that in our experiment, the horizon and the stage-game payoffs were chosen so that the *sizeBAD* for E4 and D8 are identical. The equivalence of initial cooperation rates between the two treatments suggests that, from the perspective of the first round, the horizon of the repeated game has an effect on cooperation mainly through its impact on the value of cooperation. An important implication of this is that our findings run counter to the folk wisdom described earlier, which attributes higher cooperation rates in longer horizons to the difficulty of having to think through more steps of backward induction.

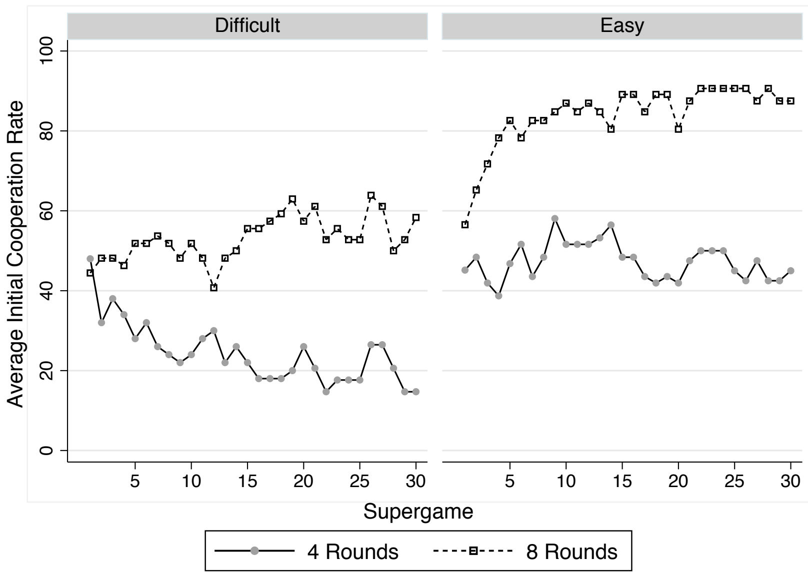


Figure 5: Average Cooperation Rates in the First Round

5. The Breakdown of Cooperation

Since the E8 treatment provides the starkest contrast to the backward induction prediction, we first provide a more detailed description of behavior in this treatment. We

[28]In addition to being true for all supergames pooled together, this is true for most supergames taken individually, except for a few outliers. E4 is higher in supergames 12 and 14 (at the 10% and 5% level, respectively) and D8 is higher in supergames 18, 19, and 21 (at the 10%, 5%, and 5% level, respectively). Pooling across supergames from the first and second half of a session separately, cooperation rates are not significantly different between E4 and D8 (see Online Appendices A.3 and A.4 for details).

then apply the key findings from this section to the other treatments and to the other studies in our meta-analysis in the following section.

5.1 Behavior in the E8 treatment

Figure 6 tracks cooperation rates across supergames, with each line corresponding to a specific round of the supergame. The selected rounds include the first round and the final three rounds.[29] Looking at the last round, the trend toward defection is clear. The round before that shows a short-lived increase in cooperation followed by a systematic decline. Two rounds before the end, the cooperation rate increases more dramatically and for a longer time, but this is eventually followed by a gradual decline. Cooperation rates in round one increase for most of the experiment but then stabilize towards the end, at a high level close to 90%. Hence, confirming the results from prior studies with longer horizons, cooperation early in a supergame increases with experience, but cooperation at the end of a supergame decreases with experience. In addition, non-monotonicity in cooperation rates for intermediate rounds suggests that the decline in cooperation slowly makes its way back from the last round. On the whole, there is a compelling picture of the unraveling of cooperation. However, the process is slow, and, even by the thirtieth supergame, cooperation is not decreasing in the first round.

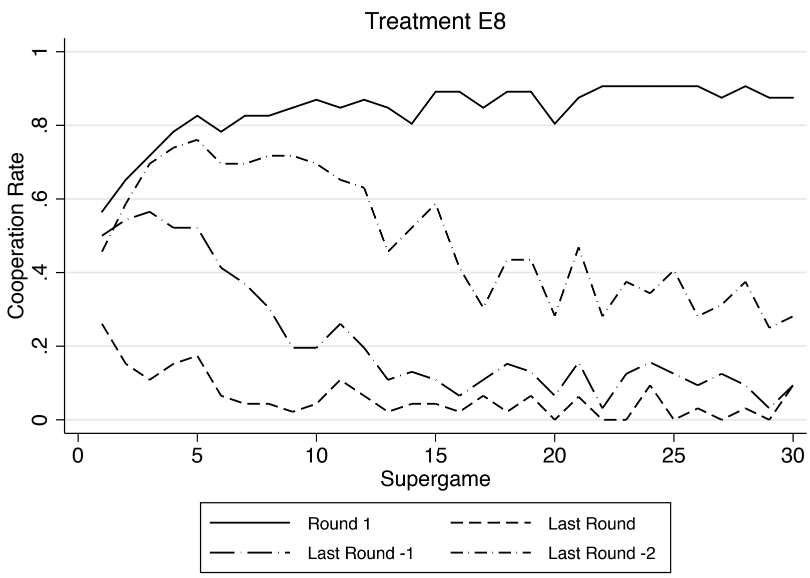


Figure 6: Mean Cooperation Rate by Round

Thus, we have conflicting observations: behavior at the end of a supergame moves slowly in the direction suggested by backward induction, while cooperation in early rounds increases with experience. To reconcile the conflict, consider the aggregate measure, mean round to first defection. This measure is a meaningful statistic to represent the unraveling

[29]Online Appendix A.3 replicates Figure 6, including all rounds.

of cooperation, primarily because we think of subjects using threshold strategies. That is, we expect defection by either player to initiate defection from then on. Hence, the typical description of backward induction in a finitely repeated PD implicitly involves the use of threshold strategies: (conditionally) cooperative behavior in the beginning of a supergame that is potentially followed by noncooperative behavior at the end of the supergame. Indeed, reasoning through the set of such strategies provides a basis for conceptualizing the process of backward induction.

A threshold m strategy is formally defined as a strategy that defects first in round m, conditional on sustained cooperation until then; defection by either player in any round triggers defection from then on. Consequently, this family of strategies can be thought of as a mixture of Grim Trigger (Grim) and AD. They start out as Grim and switch to AD at some predetermined round m. The family of threshold strategies includes AD, by setting $m = 1$. It also includes strategies that always (conditionally) cooperate, as we allow for the round of first defection, m, to be higher than the horizon of the supergame. Thus, it is possible to observe joint cooperation in all rounds of a supergame if a subject following a threshold strategy with $m > horizon$ faces another subject who follows a similar strategy. However, any cooperative play in a round after the first defection in the supergame, regardless of who was the defector, is inconsistent with a threshold strategy. Threshold strategies also have the property that a best response to a threshold strategy is also a threshold strategy.[30]

If subjects use threshold strategies, then it would be equivalent to measure cooperation using the mean round to first defection or the mean round to last cooperation, as threshold strategies never cooperate after a defection.[31] These different statistics are presented in the same graph in the left panel of Figure 7. Two key observations are immediately apparent. First, the two lines are very different to start with but slowly converge. Second, mean round to last cooperation is decreasing with experience while mean round to first defection is increasing (at least in the early parts of a session).

If, instead of mean round to first defection, one considers mean round to last cooperation, then it appears as if subjects move in the direction suggested by backward induction in all treatments, including E8. The gap between the two lines suggests that the use of threshold strategies becomes dominant over the course of the experiment. This suggestion is confirmed in the right panel of Figure 7, which shows the fraction of choice sequences perfectly consistent with the use of a threshold strategy.

Hence, aggregate measures such as the average cooperation rate and mean round to first defection confound multiple forces. Subjects learning to play threshold strategies

[30]Threshold strategies are potentially different from conditionally cooperative strategies which other studies of repeated social dilemmas have focussed on. Threshold strategies are by definition conditionally cooperative only up to the threshold round (except if $m > horizon$). Always defect is not a conditionally cooperative strategy, but is a special case of the threshold strategies (where $m = 1$).

[31]More precisely, for a subject using a threshold strategy, the last round of cooperation is the round before the first defection, regardless of the opponent's strategy. Hence, when we directly compare the mean round to first defection and the mean round to last cooperation, we add one to the latter.

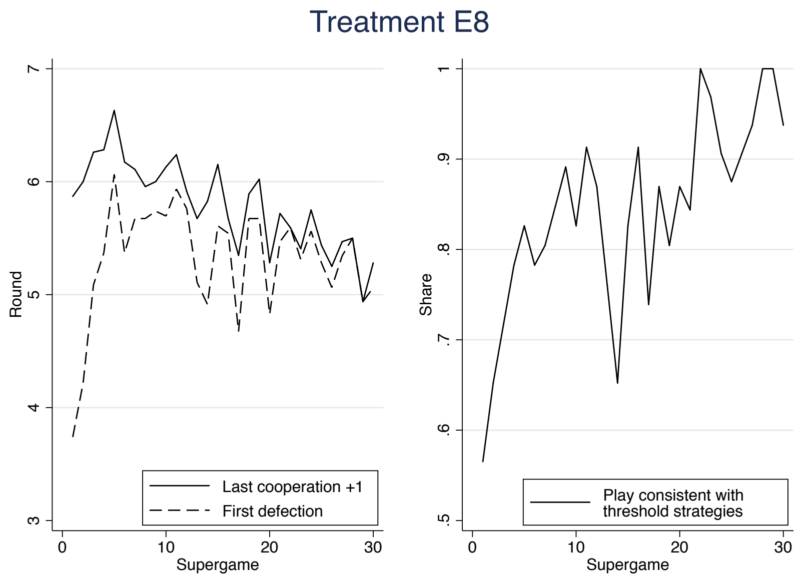


Figure 7: Evolution of Threshold Strategies

can increase their cooperation at the beginning of a supergame, even if the strategy they are learning is not more cooperative. To illustrate this effect, consider a subject who, on average, over the course of a session, plays a threshold strategy that (conditionally) cooperates for the first four rounds and defects from round five onwards ($m = 5$). However, the probability that the subject implements the strategy correctly is only 0.6 in early supergames, whereas it is 1 in later supergames. If we assume that the distribution of strategies used by the other subjects remains constant, and a sufficient share of them play cooperative strategies, mean round to first defection will increase with experience for the subject learning to use this threshold strategy. This is because the subject will sometimes defect before round five in early supergames, even in the absence of any defection by her partner, but never in later supergames. This type of learning behavior would also lead to increasing cooperation rates in round one. In addition, it would generate a decreasing round to last cooperation over supergames.[32]

For subjects who have settled on threshold strategies, it is possible to identify two additional forces, each pulling in the opposite direction. If a subject believes that his partner is likely to defect starting in round five, then he would want to start defecting at round four. This is captured by the fact that a best response to a threshold m strategy is a threshold $m - 1$ strategy. This reasoning is exactly the building block for the logic of backward induction and leads to lower cooperation rates, a decrease in the round to first defection among subjects using threshold strategies, and a decrease in the last round

[32]Burton-Chellew et al. (2016) make a related observation in the context of a public goods game. By comparing how subjects play against other subjects vs. computers, they show that cooperative behavior often attributed to social preferences in such contexts are better explained as misunderstandings in how to maximize income.

of cooperation. However, even if every subject uses threshold strategies, if there is heterogeneity in thresholds to start with, some subjects may realize over time that enough of their partners use higher thresholds than they do and, thus, may want to defect later. Such adjustments would lead to increases in some of the cooperation measures. Consequently, the overall effect on cooperation is ambiguous. These considerations highlight the problems arising from restricting attention to these aggregate measures. They confound the learning taking place on different levels: learning to use threshold strategies, updating beliefs about the strategies of others, and best responding to the population.

Figure 8 provides further evidence for this interpretation. The graph on the left compares the evolution of mean round to first defection for the whole sample to that of the subset of pairs that jointly cooperate in Round 1. As expected, the line conditional on Round 1 cooperation is higher, but the gap between the two lines shrinks as round 1 cooperation rates increase over time. Most importantly, conditional on achieving cooperation in the first round, mean round to first defection actually decreases over time.[33] The graph on the right demonstrates this in another way, by plotting the distribution of the first defection round for the first, the second and the last ten supergames. If the breakdown of cooperation is defined as the first defection for a pair, then cooperation is most likely to break down at the beginning or towards the end of the supergame. With experience, the probability of breakdown at the beginning of a supergame decreases, but conditional on surviving the first round, cooperation starts to break down earlier. The shift is slow but clearly visible. The modal defection point (conditional on being higher than 1) shifts earlier by one round for every ten supergames.

5.2 Breakdown of Cooperation in Other Treatments

Figure 9 illustrates, for the three other treatments, the evolution of cooperation for the first and last three rounds. D8 has a similar increase in initial cooperation with experience, as noted for E8, but it is less pronounced. Initial rates of cooperation are below 60% for nearly all supergames and are mostly comparable to those observed in E4. The lowest rates of initial cooperation are, as expected, in the D4 treatment. The rate drops quickly from a starting point similar to the other treatments to a rate of about 20%, where it remains for the majority of the supergames.

For all treatments (including E8), cooperation in the last round is infrequent, especially after the first ten supergames. We observe a similar pattern for cooperation in the penultimate round, although for the easy stage-games, cooperation either starts much higher or takes more supergames to start decreasing. The treatments display more important differences in behavior for the third from last round. Here, cooperation rates drop consistently below the 20% mark in the difficult stage-game treatments and take longer to start decreasing in the long-horizon treatments. Cooperation in this round drops quickly to very low levels in D4, hovers around the 20% mark in E4, and starts higher in D8 before dropping below 20%. Overall, this confirms the tendency of decreasing cooperation

[33]Note, however, that this decrease slows down considerably after supergame 16.

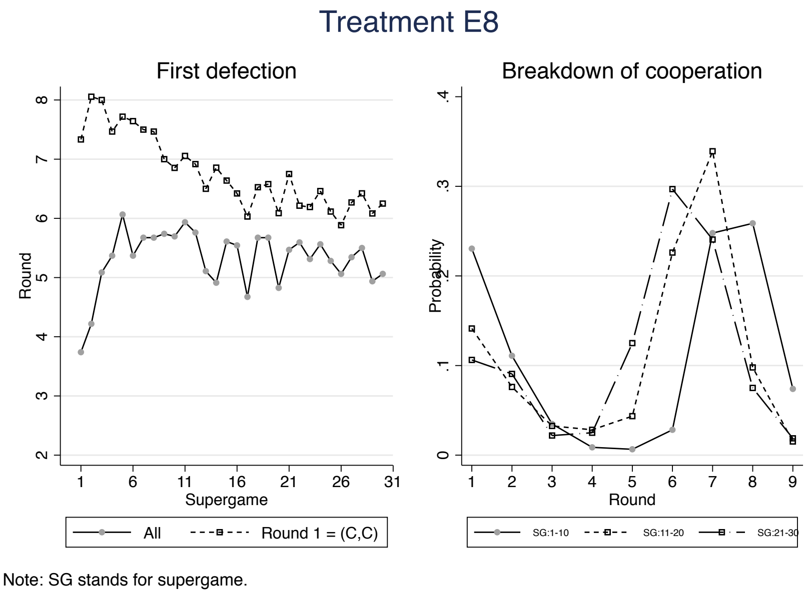


Figure 8: (Left Panel) Mean Round to First Defection: All Pairs Versus Those That Cooperated in Round One; (Right Panel) Probability of Breakdown in Cooperation

rates to start from the last round and gradually shift to earlier rounds. However, this also highlights that this process can be slow, as cooperation rates in round one decrease over the 30 supergames in only one of the four treatments.

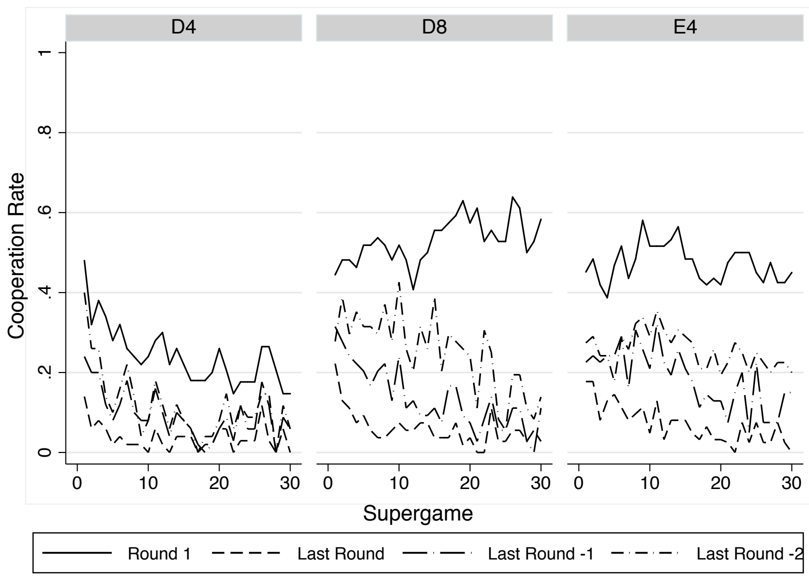


Figure 9: Cooperation Rates in Selected Rounds Across Supergames

Figure 10 confirms the observations that not everyone plays threshold strategies at the start of the experiment and that the use of threshold strategies grows with experience. In the D8 treatment, the gap between round to first defection and last round of cooperation + 1 is originally comparable in size to what is observed in the E8 treatment. With experience, the two become closer. However, by the end, they are still not identical. For the treatments with the short horizon, the gap is small to start with and even smaller by the end. Note that with a shorter horizon, there are fewer possible deviations from a threshold strategy. Moreover, with a longer horizon, there is more incentive to restore cooperation after a defection is observed.[34] These suggest that convergence to threshold strategies would happen faster in shorter horizon games.

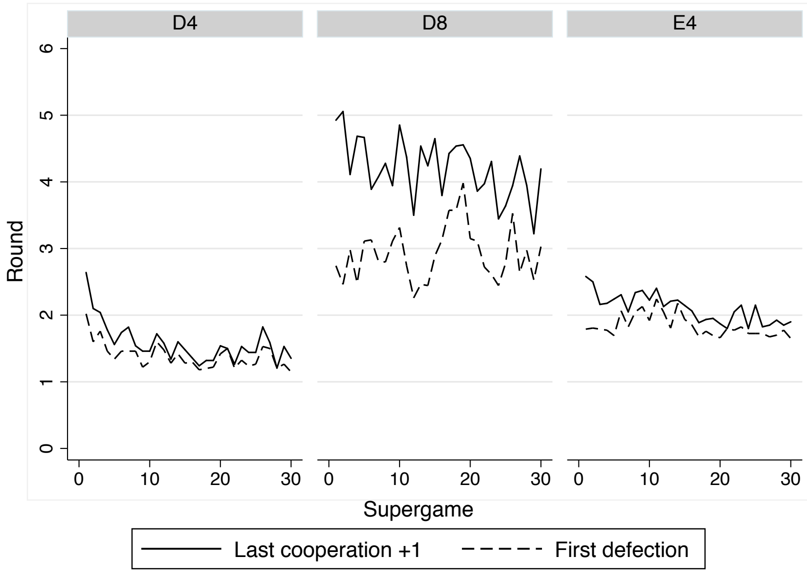


Figure 10: Evolution of First Defection Versus Last Cooperation Across Supergames

What about experiments in the meta? Are there also indications of unraveling in these once behavior is considered in a less aggregated form? To investigate this, we replicate Figures 7 and 8 in Online Appendix A.2 for the the two experiments that allowed subjects to play a substantial number of supergames: AM1993 and BMR2006. Both experiments show patterns consistent with our experimental results. Cooperation in the last round quickly decreases, whereas cooperation rates in earlier rounds first increase. The increase is followed by a decrease once the next round's cooperation rate is low enough. In both studies, there is a steady increase in round-one cooperation that does not reach the point where it starts decreasing.

Perhaps the most striking regularity to emerge across all the papers in the meta-study and our own experiment is the universal increase in the use of threshold strategies when we

[34]Indeed, H is negatively correlated with play consistent with a threshold strategy in the first supergame. This does not reach statistical significance if only considering our experiment ($p = 0.11$) but it is significant at the 1% level when considering the entire meta data. Note that it is not statistically related to g or ℓ.

compare the beginning of an experiment to the end (see Table 5). In the first supergame of all studies with $H \geq 8$, less than 50% of play is consistent with a threshold strategy. However, this number is higher than 75% in all but one treatment by the last supergame (in many, it is more than 85%). Even in experiments with $H = 4$, which already begin with 68% play of threshold strategies, they are more popular at the end. This suggests a non-negligible amount of experimentation or confusion at the beginning of a session, followed by a universal convergence to using threshold strategies.[35,36]

Table 5: Consistency of Play with Threshold Strategies

				Play Consistent With Threshold Strategy	
Experiment	Horizon	g	ℓ	First Supergame	Last Supergame
DB2005	2	1.17	0.83	–	–
	2	0.83	1.17	–	–
	4	1.17	0.83	0.68	0.80
	4	0.83	1.17	0.68	0.78
FO2012	8	4.00	4.00	0.43	0.90
	8	2.00	4.00	0.43	0.90
	8	1.33	0.67	0.37	0.77
	8	0.67	0.67	0.47	0.87
BMR2006	10	2.33	2.33	0.42	0.81
AM1993	10	1.67	1.33	0.29	0.79
CDFR1996	10	0.44	0.78	0.30	0.50
Meta All	.	.	.	0.52	0.79
EFY (D4)	4	3.00	2.83	0.66	0.94
EFY (D8)	8	3.00	2.83	0.50	0.65
EFY (E4)	4	1.00	1.42	0.66	0.94
EFY (E8)	8	1.00	1.42	0.57	0.89
EFY All	.	.	.	0.60	0.85

Notes: Supergame refers to supergame within a set of payoff and horizon parameters.

6. Long-run Behavior

The results of the last sections are highly suggestive that unraveling is at work in all treatments. However, for some treatment parameters, the process is slow enough that it would take too long for cooperation to reach close to zero levels in a reasonable amount of time (for subjects to be in a laboratory). Hence, we now estimate a learning model that will allow us to consider what would happen with even more experience. Using estimates obtained individually for each subject, we can simulate behavior for many more supergames than can be observed during a typical lab session. This can help us gain

[35] The only study in which consistency with the threshold strategy is still low (at only 50%) by the last supergame is CDFR1996, in which the last supergame is only the second supergame.

[36] Statistical significance is established in the regressions reported in Online Appendix A.2.

insight into whether the unraveling would eventually move back to round one or whether it would stop short of going all the way. It can also give us a sense of the speed at which this might happen, as well as providing structural estimates for a counterfactual analysis and an exploration of the expected payoffs of different strategies conditional on the distribution of play.

6.1 Model

The general structure of the learning model we adopt is motivated by the following observations documented in the previous sections: (1) cooperation rates in the first round of a supergame are decreasing in the size of the basin of attraction of AD; (2) choices respond to experiences with other players in previous supergames; and (3) a majority of subjects converge to using thresholds strategies. These observations suggest that subjects are influenced by their beliefs over the type of strategy their partners are following (point 2 above), and by the implied value of cooperation given these beliefs, which is also a function of the stage-game payoffs and the supergame horizon (point 1 above). We specify a simple belief-based learning model that can capture these key features.[37]

Each subject is assumed to start the first supergame with a prior over the type of strategies her partner uses. The set of strategies considered in the learning model consists of all threshold-type strategies along with TFT and Suspicious Tit-for-Tat (STFT).[38] Note that, contrary to the threshold strategies, TFT and STFT allow for cooperation to re-emerge after a period of defection within a supergame. We have included all strategies for which there is evidence of systematic use in the data.[39]

Beliefs evolve over time, given a subject's experience within a supergame. After every

[37]This is similar to the recent use of learning models to investigate the evolution of behavior in dynamic games. Dal Bó and Fréchette (2011) do this in the context of indefinitely repeated games experiments; Bigoni et al. (2015) use a learning model to better understand the evolution of play in their continuous-time experiments. In both cases, however, the problem is substantially simplified by the fact that strategies take extreme forms–immediate and sustained defection or conditional cooperation (sustained or partial). In the first paper, restricting attention to initial behavior is sufficient to identify strategies; in the second paper, initial and final behavior are sufficient to discriminate among the strategies considered. This will not be possible here, and, hence, estimating a learning model poses a greater challenge.

The approach described here is closest to that of Dal Bó & Fréchette (2011). The model is in the style of Cheung & Friedman (1997). The reader interested in belief-based learning models is referred to Fudenberg (1998). There are many other popular learning models: some important ones are found in Crawford (1995), Roth & Erev (1995), Cooper et al. (1997), and Camerer & Ho (1999).

[38]The set of threshold strategies includes a threshold strategy that cooperates in every round if the other subject cooperates (threshold is set to horizon +1), as well as AD (threshold is set to 1). TFT and STFT replicate the other player's choice in the previous round; TFT starts by cooperating, whereas STFT starts by defecting.

[39]Cooperating all the time, irrespective of the other's choice, is not included in the strategy set because there is no indication in the data that subjects follow such a strategy. More specifically, even the most cooperative subject in our dataset defected at least 34 times throughout the session, and at least 15 times in the last ten supergames.

supergame, a subject updates her beliefs as follows:

$$\beta_{it+1} = \theta_i \beta_{it} + L_{it} \tag{1}$$

where β_{it}^k can be interpreted as the weight that subject i puts on strategy k to be adopted by his opponent in supergame t.[40] θ_i denotes how the subject discounts past beliefs ($\theta_i = 0$ gives Cournot dynamics; $\theta_i = 1$ fictitious play), and L_{it} is the update vector given play in supergame t. L_{it}^k takes value 1 when there is a unique strategy that is most consistent with the opponent's play within a supergame; for all other strategies, the update vector takes value 0. When there are multiple strategies that are equally consistent with the observed play, threshold strategies take precedence, but there is uniform updating among those.[41]

Given these beliefs, each subject is modeled as a random utility maximizer. Thus, the expected utility associated with each strategy can be denoted as a vector:

$$\vec{\mu}_{it} = \vec{u}_{it} + \lambda_i \vec{\epsilon}_{it} \tag{2}$$

$\vec{u}_{it} = \vec{U}\beta_{it}$, where $\vec{U}$ is a square matrix representing the payoff associated with playing each strategy against every other strategy. Note that $\vec{U}$ is a function of the horizon of the repeated game, as well as of the stage-game payoffs. The parameter λ_i is a scaling parameter that measures how well each subject best-responds to her beliefs, and ϵ_{it} is a vector of idiosyncratic error terms. Given standard distributional assumptions on the error terms, this gives rise to the usual logistic form. In other words, the probability of choosing a strategy k can be written as:

$$p_{it}^k = \frac{\exp(\frac{u_{it}^k}{\lambda_i})}{\sum_k \exp(\frac{u_{it}^k}{\lambda_i})} \tag{3}$$

The structure of the learning model that we adopt is typical. What is unusual in our case is that, on this level, it describes choices over strategies rather than actions. It captures the dynamics of updating beliefs across supergames about the strategies adopted by others in the population and, consequently, describes learning about the optimality of different strategies.[42]

[40]Note that the sum of the components of β_{it} need not sum to 1. This sum can be interpreted as the *strength* of the prior. In combination with θ_i, the sum captures how much emphasis is given to new experiences in updating beliefs.

[41]The tie-breaking rule, which favors threshold strategies in the belief updating, eliminates the possibility of emergence of cooperation via TFT-type strategies in an environment in which all subjects have settled on threshold strategies, as observed towards the end of the sessions in our data.

[42]Cox et al. (2015) propose a similar model of boundedly rational behavior in the finitely repeated PD, in which they relax the assumption that players' prior beliefs are consistent with their opponent's best response. They explore bounds on initial beliefs that can sustain cooperative behavior and highlight parallels between the theoretical predictions of the model and experimental results. However, they do

Not all behavior within a supergame is perfectly consistent with subjects following one of the strategies that we consider. Allowing for other behavior is important to describing the evolution, but it comes at the cost of more parameters to estimate. Given that our data suggest that threshold strategies become dominant over time, we follow a parsimonious approach, and instead of expanding the set of strategies considered, we augment the standard model by introducing an implementation error.[43]

The implementation error introduces noise into how strategies are translated into actions within a supergame. In every round, there is some probability that the choice recommended by a strategy is incorrectly implemented. As the results have shown, in some treatments, all choices quickly become consistent with threshold strategies, while in others, the choices inconsistent with threshold strategies disappear more slowly. To account for this, the implementation error is specified as $\sigma_{it} = \sigma_i^{t^{\kappa_i}}$, where t is the supergame number and $0 \leq \sigma_{it} \leq 0.5$. Such a specification allows for extremely rapid decreases in implementation error (high κ) as well as constant implementation error ($\kappa = 0$). Specifically, given her strategy choice and the history of play within a supergame, σ_{it} represents the probability that subject i will choose the action that is inconsistent with her strategy in a given round.[44]

In summary, for each subject, we estimate $_{i0}$, $_{i}$, σ_i, which describe initial beliefs, noise in strategy and action choice implementation, and $_{i}$, κ_i, which describe how beliefs are updated with experience and how execution noise changes over time.[45] The estimates are obtained via maximum likelihood estimation for each subject separately.[46] We provide summary statistics of the estimates in Online Appendix A.5.

It is important to clarify that the model allows for a great range of behavior. Neither convergence to threshold strategies, nor unraveling of cooperation is structurally imposed, although both are potential outcomes under certain sets of parameters.

6.2 Simulations

We first use individual-level estimates in conducting simulations to determine if the learning model captures the main qualitative features of the data. Then, we use the

---

not structurally estimate the model to study long-term behavior.

[43]There is also no other strategy for which there is sufficient evidence to indicate that a portion of the population might be using it systematically.

[44]The implementation noise affects play within a supergame in two possible ways. The first is the direct effect; in every round, it creates a potential discrepancy between intended choice and actual choice. The second is the indirect effect; it changes the history of play for future rounds.

[45]When $H = 4$, this represents 11 parameters for 120 observations (30 supergames of four rounds), and when $H = 8$, it is 15 parameters for 240 observations (30 supergames of 8 rounds). Except in the two sessions of 20 supergames, where there are 80 and 160 choices per subject for the short and long horizons, respectively.

[46]An alternative would be to pool the data. However, for the purpose of this paper and given the number of observations per subject, obtaining subject-specific estimates is useful and reasonable. Fréchette (2009) discusses issues and solutions related to pooling data across subjects in estimating learning models and, more specifically, with respect to hypothesis testing.

simulations with more repeated games to understand how cooperation would evolve in the long run.[47] These simulations consist of 100,000 sessions by treatment.[48]

The learning model fits the data well in terms of capturing the differences between the treatments with respect to aggregate cooperation rates, mean round to first defection, and evolution of behavior within a session in all treatments. This is illustrated for the E8 treatment in Figure 11, which compares the average simulated cooperation rate for each round of the repeated game with the experimental results.[49] The simulation results capture the key qualitative features of behavior observed in the data remarkably well. In particular, cooperation rates are clearly increasing in the early rounds of a supergame, while decreasing in later rounds, as observed in the data. For rounds in the middle, such as rounds 5 and 6, there is non-monotonicity in cooperation rates, as they first increase and later decrease. Note that these features are recovered in a model in which there are no round- or supergame-specific variables, and updating occurs over beliefs about strategies only between supergames.

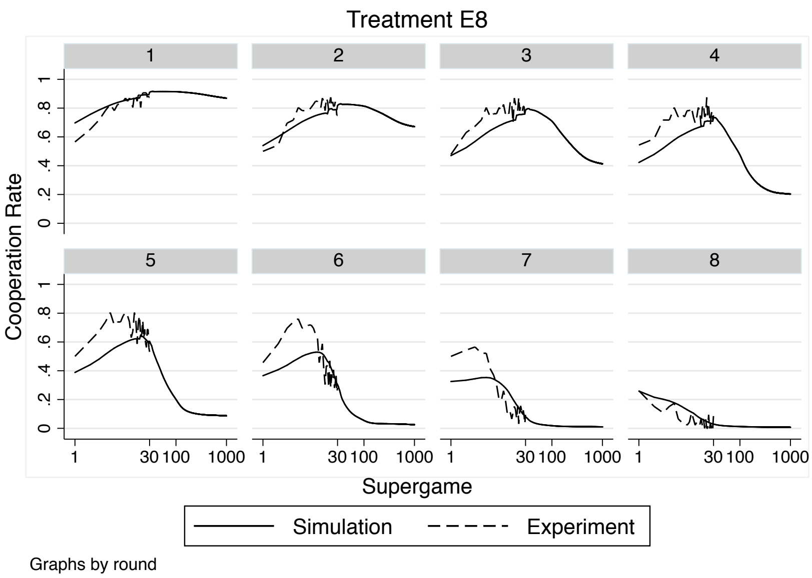


Figure 11: Average Cooperation: Simulation Versus Experimental Data for Each Round In E8

[47]For the simulations, subjects who show limited variation in choice within a session are selected out, and their actions are simulated directly. Specifically, any subject who cooperates for at most two rounds throughout the whole session is labeled an AD-type, and is assumed to continue to play the same action, irrespective of the choices of the subjects she is paired with in future supergames. None of the subjects identified as AD types cooperates in any round of the last ten supergames. This identification gives us 3/5/11/17 subjects to be AD types in treatments E8/D8/E4/D4, respectively.

[48]The composition of each session is obtained by randomly drawing (with replacement) 14 subjects (and their estimated parameters) from the pool of subjects who participated in the corresponding treatment.

[49]Online Appendix A.5 replicates this analysis for other treatments and also includes detailed figures focusing only on the first 30 supergames.

Figure 11 also provides insights into the way cooperation would evolve in the long run. The supergame (number of repeated games) axis is displayed in log scale to facilitate the comparison between evolution of behavior in the short term versus the long term. We observe that in this treatment, which is most conducive to cooperation, there is still cooperation after 1000 supergames. However, this is clearly limited to early rounds. More importantly, cooperation rates, if they are still positive, continue declining in all rounds, even after 1000 supergames.[50] The evolution suggests that there is unraveling of cooperation in all rounds, but that it is so slow that cooperation rates for the first round of a supergame can remain above 80% even after significant experience.[51] In contrast, we show in Online Appendix A.5 that cooperation rates in all other treatments quickly decline to levels below 10% with little experience.

6.3 Counterfactuals

In the remainder of this section, we investigate which factors contribute to the sustained cooperation predicted by the learning model for long run behavior in E8. To do so, we take advantage of the structure of the learning model and study how cooperation evolves in the long-run under different counterfactual specifications.

The Kreps et al. (1982) model shows that sustained cooperation until almost the last round can be a best response to a small fraction of cooperative subjects from a rational agent who understands backward induction. Since our estimations for the learning model are at the subject level, we can directly investigate if there is, indeed, significant heterogeneity in cooperative behavior in the population and whether or not this affects the unraveling of cooperation. In Online Appendix A.5, we compare cooperation rates in simulations where all subjects are included to those where the most cooperative subjects are removed from the sample. The comparative statics suggest that the existence of cooperative types can slow down unraveling, but the effect seems to be limited.

Next, to explore the extent to which stage-game payoffs–through their effect on strategy choice and, consequently, evolution of beliefs–can explain why unraveling is faster in the D8 treatment relative to the E8 treatment, we conduct the following counterfactual simulations: We take the individual-level estimates for the learning model from the E8 treatment and simulate how these subjects would play the D8 stage-game. This exercise enables us to keep the learning dynamics (priors, updating rule, strategy choice, and implementation error) constant while varying only the stage-game parameters. In Figure 12, this is plotted as CF1. The comparison of E8 and CF1 provides a striking depiction of the importance of the stage-game parameters in the evolution of behavior. The gap between the two lines for the first supergame demonstrates the impact of the stage-game

[50]Regressing cooperation on supergame using the last 50 supergames of the simulations by round reveals a negative coefficient for all rounds. The negative coefficient is significant in all rounds except round 4, 5, and 8 where cooperation levels are 20%, 9% and 1% by the 1000th supergame.

[51]While there is evidence of a slow but continued decline in cooperation within the span of our simulations, it does not rule out the possibility that unraveling eventually stagnates at non-zero cooperation levels with further experience.

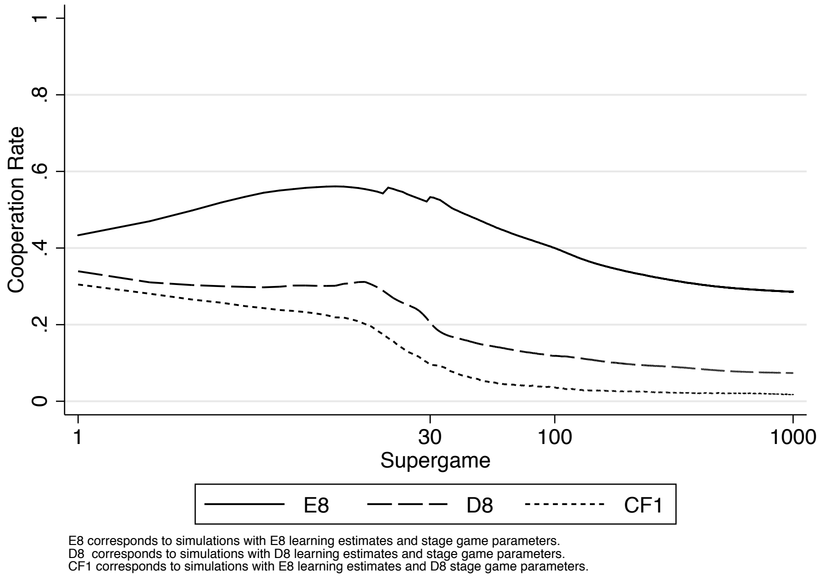


Figure 12: Long Term Evolution of Aggregate Cooperation

parameters on strategy choice when beliefs are kept constant. The gap widens with experience as subjects interact with each other and update their beliefs about others, such that cooperation quickly reaches levels below 10% in less than 50 supergames in CF1.[52]

Estimates from the learning model can also be used to investigate the optimality of strategy choice among subjects. In Online Appendix A.5, we plot the expected payoff associated with each strategy and the frequency with which this strategy is chosen for each treatment. This exercise reveals that expected payoffs are relatively flat in E8.[53] This provides further evidence for why unraveling is slow in this treatment.

Overall, we see that the speed of unraveling is closely connected to how conducive stage-game parameters are to cooperation, closely mirroring our results on the size of the basin of attraction of AD as a determinant of initial cooperation.

[52]We can also study the opposite counterfactual (as plotted in Online Appendix A.5). That is, we can keep the E8 stage-game parameters constant, but use learning estimates for the subjects who participated in the D8 treatment. Limited unraveling of cooperation with this counterfactual further highlight that this behavior is driven by stage-game parameters rather than treatment specific learning dynamics.

[53]For example, we see that in the first supergame, the optimal strategy is using Threshold 7 or 8, while in the last supergame of the session, it is Threshold 5 or 6. For the frequency of choice, TFT is the most popular strategy early on in the session, but it is replaced by late threshold strategies by the end of the session. In both cases, some of the most popular strategies are suboptimal, but the expected loss associated with using them is small. In comparison, expected payoffs and frequency of choice associated with the strategies are quite different in D8. AD (Threshold 1) is the optimal strategy at both the beginning and the end of the session. While TFT and STFT are common choices in the first supergame, AD is the most frequent by the last.

7. Discussion

Despite the wealth of experimental research on the finitely repeated PD, prior evidence provides a limited understanding of the factors that contribute to the emergence of cooperation and its possible unraveling with experience.

In this paper, to understand how cooperative behavior and its evolution with experience vary with the environment in this canonical game, we re-analyze the data from prior experimental studies and supplement these results with a new experiment. In doing so, we are able to reconcile many of the contradictory results in the prior literature, which, we argue, are driven by two behavioral regularities: the role of the value of cooperation and the emergence of threshold strategies.

Our paper makes several further contributions to the literature. First, we show that the parameters of the supergame–the horizon in particular–have a significant impact on initial cooperation. Our analysis reveals that a longer horizon increases initial cooperation because it increases the value of using conditionally cooperative strategies, which can be captured by a simple statistic: the size of the basin of attraction of the Always Defect strategy. This value-of-cooperation result relates to recent studies on continuous-time PD games (Friedman & Oprea (2012); Bigoni et al. (2015); Calford & Oprea (2017)). Friedman & Oprea (2012) conclude that the unraveling argument of backward induction loses its force when players can react quickly. Treatment differences in our experiment are driven by similar forces. The decision to cooperate depends on how the temptation to become the first defector compares to the potential loss from defecting too early. The size of the basin of attraction captures this trade-off precisely and, in doing so, highlights the role of strategic uncertainty in determining cooperative behavior. The predictive power of the size of the basin of attraction can also be understood from an evolutionary game theory perspective. The size of the basin of attraction can be interpreted to capture the *robustness* of Always Defect as an evolutionary stable strategy in a finitely repeated prisoners dilemma.[54] It has been argued that, while defection should dominate in short-horizon finitely repeated PD games, as the horizon increases, the emergence of conditionally cooperative strategies should become more likely (for instance, see Fudenberg & Imhof (2008); Imhof et al. (2005)). This is highly intuitive. The presence of a small share of conditionally cooperative players can make it worthwhile to initiate cooperative play, especially in long-horizon games conducive to cooperation. This also fits nicely with our results on long-term dynamics using the learning model. Noise in strategy choice or implementation of actions can be interpreted as stochastic invasions by alternative strategies that consequently slow down, or even could possibly prevent, the unraveling of cooperation (as we observe in the E8 treatment).

Second, the paper identifies a crucial regularity–namely, that threshold strategies always emerge over time. That is, in *every* study of the finitely repeated PD in which the

[54] It is linked to the size of the invasion (share of the population following the alternative strategy) needed to take over Always Defect.

game is played more than once, threshold strategies are substantially more common by the end of the experiment. While the role of threshold strategies has been noted in the previous literature (for instance theoretically in Radner (1986) and recently empirically investigated in Friedman & Oprea (2012)), we find convergence to using threshold strategies to be a critical and systematic feature of the evolution of behavior in this game. Hence, we identify the interaction of two opposing forces–learning to cooperate in early rounds by convergence to using threshold strategies and learning to defect in later rounds due to the unraveling argument of backward induction–to be fundamental in explaining the variation across papers and treatments in the evolution of behavior. This result also highlights an essential difference between the finitely repeated PD and the centipede game, which, by construction, constrains players to conditionally cooperative threshold strategies. While both games have been extensively used to study backward induction, our results suggest that, (at least) short-term dynamics in these games are governed by potentially different forces.

Finally, although our study is not explicitly designed to test alternative theories that predict cooperation in the finitely repeated PD, we can relate our results to these theories. Analysis using the learning model indicates that there is some heterogeneity across subjects in terms of responsiveness to past experiences and willingness to follow cooperative strategies. This observation suggests that the reputation-building forces identified in the model of Kreps et al. (1982) may play a role in generating cooperation and slowing down the unraveling of cooperation in the finitely repeated PD. Although, in contrast to the static nature of the Kreps et al. (1982) model, the behavior we observe suggests that beliefs change significantly across supergames in response to past experiences.[55] On the other hand, as discussed earlier, the value-of-cooperation result supports the approximate best-responses approach of the epsilon-equilibrium model in Radner (1986), as suggested by Friedman & Oprea (2012). Differences in cooperative behavior across our treatments appear to be driven primarily by the corresponding differences in the trade-off between initiating cooperation versus defection when there is uncertainty about the strategy followed by one's opponent. While our analysis suggests that the unraveling of cooperation is still happening towards the end in all of our treatments, especially in environments with potentially high returns to cooperation, we cannot rule out that cooperation would stabilize at positive levels with further experience. In such treatments where unraveling is particularly slow, we estimate that a portion of the population follows more cooperative strategies than the optimal best-response to the population, but the relative cost of adopting these strategies is quite small.

[55]Note that we do not see any evidence of subjects following unconditionally cooperative strategies. This confines the space of behavioral types that can be meaningfully considered in this setting.

References

Alaoui, L. & Penta, A. (2016), 'Endogenous depth of reasoning', *Review of Economic Studies* **83**(4), 297–1333.

Andreoni, J. & Miller, J. H. (1993), 'Rational cooperation in the finitely repeated prisoner's dilemma: Experimental evidence', *The Economic Journal* **103**(418).

Andreoni, J. & Petrie, R. (2004), 'Public goods experiments without confidentiality: a glimpse into fund-raising', *Journal of public Economics* **88**(7), 1605–1623.

Andreoni, J. & Petrie, R. (2008), 'Beauty, gender and stereotypes: Evidence from laboratory experiments', *Journal of Economic Psychology* **29**(1), 73–93.

Bereby-Meyer, Y. & Roth, A. E. (2006), 'The speed of learning in noisy games: Partial reinforcement and the sustainability of cooperation', *The American Economic Review* **96**(4), 1029–1042.

Bertrand, M., Duflo, E. & Mullainathan, S. (2004), 'How much should we trust differences-in-differences estimates?', *The Quarterly Journal of Economics* pp. 249–275.

Bigoni, M., Casari, M., Skrzypacz, A. & Spagnolo, G. (2015), 'Time horizon and cooperation in continuous time', *Econometrica* **83**(2), 587616.

Blonski, M., Ockenfels, P. & Spagnolo, G. (2011), 'Equilibrium selection in the repeated prisoner's dilemma: Axiomatic approach and experimental evidence', *The American Economic Journal: Microeconomics* **3**(3), 164–192.

Bornstein, G., Kugler, T. & Ziegelmeyer, A. (2004), 'Individual and group decisions in the centipede game: Are groups more "rational" players?', *Journal of Experimental Social Psychology* **40**(5), 599–605.

Burton-Chellew, M. N., El Mouden, C. & West, S. A. (2016), 'Conditional cooperation and confusion in public-goods experiments', *Proceedings of the National Academy of Sciences* **113**(5), 1291–1296.

Cabral, L. M., Ozbay, E. & Schotter, A. (2014), 'Revisiting strategic versus non-strategic cooperation', *Games and Economic Behavior* **87**, 100–121.

Calford, E. & Oprea, R. (2017), 'Continuity, inertia and strategic uncertainty: A test of the theory of continuous time games', *Econometrica* . Forthcoming.

Camerer, C. F. & Ho, T.-H. (1999), 'Experience-weighted attraction learning in normal form games', *Econometrica* **67**, 827–874.

Cameron, A. C., Gelbach, J. B. & Miller, D. L. (2008), 'Bootstrap-based improvements for inference with clustered errors', *The Review of Economics and Statistics* **90**(3), 414–427.

Cameron, A. C. & Miller, D. L. (2015), 'A practitioners guide to cluster-robust inference', *Journal of Human Resources* **50**(2), 317–372.

Carter, A. V., Schnepel, K. T. & Steigerwald, D. G. (2017), 'Asymptotic behavior of at test robust to cluster heterogeneity', *Review of Economics and Statistics* (0).

Cheung, Y.-W. & Friedman, D. (1997), 'Individual learning in normal form games: Some laboratory results', *Games and Economic Behavior* **19**(1), 46–76.

Cooper, D. J., Garvin, S. & Kagel, J. H. (1997), 'Signalling and adaptive learning in an entry limit pricing game', *The RAND Journal of Economics* **28**(4), 662–683.

Cooper, R., DeJong, D. V., Fosythe, R. & Ross, T. W. (1996), 'Cooperation without reputation: experimental evidence from prisoner's dilemma games', *Games and Economic Behavior* **12**(2), 187–218.

Cox, C. A., Jones, M. T., Pflum, K. E. & Healy, P. J. (2015), 'Revealed reputations in the finitely repeated prisoners? dilemma', *Economic Theory* **58**(3), 441–484.

Crawford, V. P. (1995), 'Adaptive dynamics in coordination games', *Econometrica* **63**(1), 103–43.

Dal Bó, P. (2005), 'Cooperation under the shadow of the future: experimental evidence from infinitely repeated games', *The American Economic Review* pp. 1591–1604.

Dal Bó, P. & Fréchette, G. R. (2011), 'The evolution of cooperation in infinitely repeated games: Experimental evidence', *The American Economic Review* **101**(1), 411–429.

Fey, M., McKelvey, R. D. & Palfrey, T. R. (1996), 'An experimental study of constant-sum centipede games', *International Journal of Game Theory* **25**(3), 269–287.

Fischbacher, U. (2007), 'z-tree: Zurich toolbox for ready-made economic experiments', *Experimental Economics* **10**(2), 171–178.

Flood, M. M. (1952), 'Some experimental games', *Research Memorandum RM-789, The Rand Corporation, Santa Monica* .

Fréchette, G. R. (2009), 'Learning in a multilateral bargaining experiments', *Journal of Econometrics* **153**(2), 183–195.

Fréchette, G. R. (2012), 'Session-effects in the laboratory', *Experimental Economics* **15**(3), 485–498.

Friedman, D. & Oprea, R. (2012), 'A continuous dilemma', *The American Economic Review* **102**(1).

Friedman, D. & Sinervo, B. (2016), *Evolutionary Games in Natural, Social, and Virtual Worlds*, Oxford University Press.

Fudenberg, D. (1998), *The theory of learning in games*, Vol. 2, MIT press.

Fudenberg, D. & Imhof, L. A. (2008), 'Monotone imitation dynamics in large populations', *Journal of Economic Theory* **140**(1), 229–245.

Garcia-Pola, B., Iriberri, N. & Kovarik, J. (2016), Non-equilibrium play in centipede games. Working paper.

Hauk, E. & Nagel, R. (2001), 'Choice of partners in multiple two-person prisoner's dilemma games an experimental study', *Journal of Conflict Resolution* **45**(6), 770–793.

Ho, T.-H. & Su, X. (2013), 'A dynamic level-k model in sequential games', *Management Science* **59**(2), 452–469.

Imbens, G. W. & Kolesar, M. (2016), 'Robust standard errors in small samples: Some practical advice', *Review of Economics and Statistics* **98**(4), 701–712.

Imhof, L. A., Fudenberg, D. & Nowak, M. A. (2005), 'Evolutionary cycles of cooperation and defection', *Proceedings of the National Academy of Sciences of the United States of America* **102**(31), 10797–10800.

Kagel, J. H. & McGee, P. (2016), 'Team versus individual play in finitely repeated prisoner dilemma games', *AEJ: Microeconomics* **8**(2), 253–276.

Kahn, L. M. & Murnighan, J. K. (1993), 'A general experiment on bargaining in demand games with outside options', *The American Economic Review* **88**(5), 1260–1280.

Kamei, K. & Putterman, L. (2015), 'Play it again: Partner choice, reputation building and learning in restarting, finitely-repeated dilemma games', *Economic Journal* .

Kawagoe, T. & Takizawa, H. (2012), 'Level-k analysis of experimental centipede games', *Journal Of Economic Behavior & Organization* **82**(2), 548–566.

Kline, P., Santos, A. et al. (2012), 'A score based approach to wild bootstrap inference', *Journal of Econometric Methods* **1**(1), 23–41.

Kreps, D. M., Milgrom, P., Roberts, J. & Wilson, R. (1982), 'Rational cooperation in the finitely repeated prisoners' dilemma', *Journal of Economic Theory* **27**(2), 245–252.

Lave, L. B. (1965), 'Factors affecting cooperation in the prisoner's dilemma', *Behavioral Science* **10**(1), 26–38.

Levitt, S. D., List, J. A. & Sadoff, S. E. (2011), 'Checkmate: Exploring backward induction among chess players', *The American Economic Review* **101**(975-990).

Lugovskyy, V., Puzzello, D., Sorensen, A., Walker, J. & Williams, A. (2017), 'An experimental study of finitely and infinitely repeated linear public goods games', *Games and Economic Behavior* .

MacKinnon, J. G. & Webb, M. D. (2017), 'Wild bootstrap inference for wildly different cluster sizes', *Journal of Applied Econometrics* **32**(2), 233254.

Mantovani, M. (2016), Limited foresight in sequential games: an experiment. Working Paper.

Mao, A., Dworkin, L., Suri, S. & Watts, D. J. (2017), 'Resilient cooperators stabilize long-run cooperation in the finitely repeated prisoner's dilemma', *Nature Communications* **8**.

McKelvey, R. D. & Palfrey, T. R. (1992), 'An experimental study of the centipede game', *Econometrica* pp. 803–836.

Mengel, F. (2014*a*), 'Learning by (limited) forward looking players', *Journal of Economic Behavior Organization* **108**, 5977.

Mengel, F. (2014*b*), Risk and temptation: A meta-study on social dilemma games. Working Paper.

Moinas, S. & Pouget, S. (2013), 'The bubble game: An experimental study of speculation', *Econometrica* **81**(4), 1507–1539.

Morehous, L. G. (1966), 'One-play, two-play, five-play, and ten-play runs of prisoner's dilemma', *Journal of Conflict Resolution* pp. 354–362.

Muller, L., Sefton, M., Steinberg, R. & Vesterlund, L. (2008), 'Strategic behavior and learning in repeated voluntary contribution experiments', *Journal of Economic Behavior & Organization* **67**(3), 782–793.

Murnighan, J. K. & Roth, A. E. (1983), 'Expecting continued play in prisoner's dilemma games', *Journal of Conflict Resolution* **27**(2), 279–300.

Nagel, R. & Tang, F. F. (1998), 'Experimental results on the centipede game in normal form: an investigation on learning', *Journal of Mathematical Psychology* **42**(2), 356–384.

Normann, H.-T. & Wallace, B. (2012), 'The impact of the termination rule on cooperation in a prisoner's dilemma experiment', *International Journal of Game Theory* **41**(3), 707–718.

Palacios-Huerta, I. & Volij, O. (2009), 'Field centipedes', *The American Economic Review* **99**(4), 1619–1635.

Radner, R. (1986), 'Can bounded rationality resolve at the prisoner's can bounded rationality resolve at the prisoner's dilemma?', *Conributions to Mathematical Economics, ed. by A. Mas- Colell and W. Hildenbrand, North-Holland, Amsterdam* pp. 387–399.

Rapoport, A. & Chammah, A. M. (1965), 'Sex differences in factors contributing to the level of cooperation in the prisoner's dilemma game', *Journal of Personality and Social Psychology* **2**(6), 831.

Rapoport, A., Stein, W. E., Parco, J. E. & Nicholas, T. E. (2003), 'Equilibrium play and adaptive learning in a three-person centipede game', *Games and Economic Behavior* **43**(2), 239–265.

Reuben, E. & Seutens, S. (2012), 'Intrinsic and instrumental reciprocity: An experimental study', *Experimental Economics* **15**(1), 24–43.

Roth, A. E. (1988), 'Laboratory experimentation in economics: A methodological overview', *Economic Journal* **93**(393), 974–1031.

Roth, A. E. & Erev, I. (1995), 'Learning in extensive-form games: Experimental data and simple dynamic models in the intermediate term', *Games and Economic Behavior* **8**(1), 164–212.

Schneider, F. & Weber, R. A. (2013), Long-term commitment and cooperation. Working paper.

Selten, R. & Stoecker, R. (1986), 'End behavior in sequences of finite prisoner's dilemma supergames a learning theory approach', *Journal of Economic Behavior & Organization* **7**(1), 47–70.

Webb, M. D. (2014), Reworking wild bootstrap based inference for clustered errors, Technical report, Queen's Economics Department Working Paper.

Zauner, K. G. (1999), 'A payoff uncertainty explanation of results in experimental centipede games', *Games and Economic Behavior* **26**(1), 157–185.

Online Appendix for

Cooperation in the Finitely Repeated Prisoner's Dilemma

Matthew Embrey, U. of Sussex
Guillaume R. Fréchette, NYU
Sevgi Yuksel, UCSB

Contents

List of Figures

List of Tables

A.1. Literature Review

Selten & Stoecker (1986) study behavior in a finitely repeated PD with a horizon of 10 rounds. Subjects play 25 supergames where they are rematched between every supergame. They observe that behavior converges to a specific pattern with experience: joint cooperation in early rounds followed by joint defection in subsequent rounds once defection is initiated by either player.[56] Importantly, they state that the point at which subjects intend to first deviate moves earlier with experience.[57] Roth (1988) summarizes these observations as follows: "in the initial [supergames] players learned to cooperate [...]. In the later [supergames], players learned about the dangers of not defecting first, and cooperation began to unravel."[58] The impression at the time is that unraveling comes about with experience. We should point out that Selten and Stoecker in their paper do not take a position on whether, with more experience, unraveling would lead to complete defection in this game. They acknowledge that unraveling might slow down such that cooperation could stabilize at some level. Furthermore, their analysis is based on results from a single set of parameters, a point noted by Selten and Stoecker, as well as Roth. Hence, to what extent these results would be robust to variations is not clear. In addition, the observation about the evolution of intended deviation round is not directly linked to the pattern of play observed in the game, as it is in part based on inferences about how players *expected* to play.[59]

Andreoni & Miller (1993) and Kahn & Murnighan (1993) directly investigate whether cooperation in the finitely repeated PD is consistent with the incomplete information model of Kreps et al. (1982). Both papers involve varying the probability that subjects interact with a pre-programmed opponent to affect the subjects' beliefs over the value of building a reputation. Because we use their data in our meta-study, we focus on Andreoni & Miller (1993).

Andreoni & Miller (1993) conducted four treatments all involving 200 choices in total. In the *Partners* treatment, these were 20 finitely repeated PDs with a horizon of 10 rounds. In the *Strangers* treatment, these were 200 one-shot PDs. The two additional treatments are variations on the *Partners* treatment, where subjects are probabilistically matched

[56] In the last five supergames, 95.6% of the supergames are consistent with that pattern, although only 17.8% of the data fits that requirement in the first five supergames.

[57] It is on average round 9.2 in supergame 13, and steadily moves down to 7.4 in supergame 25. The *intended deviation period* is computed for a subset of the data which changes by supergame, but includes almost all of the data by the end of the experiment. If there was no prior defection by either player, it is taken to be the first period at which a player defects; otherwise it is either obtained from the written comments of the subject, or inferred from reported expectations about the opponents combined in an unspecified way with past behavior and past written comments. If no defection happens, the deviation period is recorded as 11.

[58] p.1000. Roth, as Selten and Stoecker, uses the word round, to which we substituted supergame for clarity.

[59] Moreover, it is calculated on a subsample of the subject population that changes in every supergame. Consequently, it is not clear if the diminishing average is a result of subjects defecting in earlier rounds as required for unraveling, or a by-product of the changing subsample.

to play against a computer that follows the Tit-For-Tat (TFT) strategy.[60] Cooperation rates are highest in the treatment where subjects are most likely to be playing against the computer, and lowest in the *Strangers* treatment.[61] By the end of the session, in all treatments except *Strangers*, cooperation rates are above 60% in round one, stay above 50% for at least 6 rounds, then fall under 10% in the last round. In the *Strangers* treatment, cooperation rates fall below 30% in the last 10 rounds. These facts, which are consistent with the findings of Kahn & Murnighan (1993), imply that subjects' choices depend on their beliefs about the type of opponent they are faced with. This is interpreted by both papers as evidence consistent with the reputation building hypothesis of Kreps et al. (1982). Andreoni & Miller (1993) go on to note that in two of the treatments where subjects play the 10-round finitely repeated PD,[62] the mean round at which the first defection is observed in a pair increases over the course of the experiment, starting below two in the first supergame and ending above 5 in the last.[63] This observation is inconsistent with unraveling, and in contrast to the result of Selten & Stoecker (1986). Again, both papers considered a single set of payoffs and a single horizon. Hence, the contrasting results could be due to the payoffs or the different ways in which each research group constructed the relevant statistic.

Cooper et al. (1996) design an experiment to separate the reputation building and altruism hypotheses. They compare a treatment with 20 one-shot PDs to a treatment where subjects play two finitely repeated PDs with an horizon of 10 rounds.[64] They observe higher cooperation rates in the finitely repeated PD than in the one-shot treatment. Cooperation rates start above 50% in the finitely repeated game and end below, but are always lower for the one-shot game. However cooperation is significantly above zero in both treatments. Due to the limited number of repetitions, they cannot analyze the evolution of behavior. They conclude that there is evidence of both reputation building and altruism; and that neither model can explain all the features of the data on its own. As with previous studies, a single set of payoffs and horizon was considered.

Hauk & Nagel (2001) study the effect of entry-choice on cooperation levels in the finitely repeated PD with an horizon of 10 rounds.[65] A control *lock-in* treatment (with

[60] In the *Computer50* treatment this probability is 50%; in *Computer0* it is 0.1%. The TFT strategy starts by cooperating and from then on matches the opponent's previous choice.

[61] In almost all rounds cooperation rates averaged over all supergames are ordered as *Computer50* > *Partners* > *Computer0* > *Strangers*. The exception are the final two rounds where it is more or less equal in most treatments and round one where *Computer50* and *Partners* are inverted. Cooperation rates are not statistically different between the *Computer0* and *Partners* treatments, but in both cases they are significantly higher than for *Strangers* and significantly less than for *Computer50*.

[62] The *Partners* and *Computer50* treatments.

[63] The first defection round is set to 11 for a subject that never defects, otherwise it is simply the first round in which a subject defects.

[64] They use a turnpike protocol to avoid potential contagion effects (McKelvey & Palfrey 1992).

[65] Certain design choices for this paper differ significantly from the other papers discussed. Each session had seven subjects; and each subject played 10 supergames simultaneously against the remaining 6 players. ID numbers for partners were used to separate the different partners, and were randomly reassigned in the following supergame.

no ability to choose partners) is compared to two *choice* treatments where subjects are unilaterally and multilaterally given an exit opportunity with a sure payoff instead of playing the PD game. The exit option yields higher payoffs than mutual defection. Hence, an entry decision reveals intentions on how to play the game, and beliefs about how other subjects might play. Results show that entry-choice can have ambiguous effects on welfare: Conditional on entering, cooperation levels are much higher in the choice treatments. However, when the entry-choice is taken into account, overall cooperation levels are indistinguishable (unilateral choice), or significantly lower (mutual choice). The treatment differences in this paper suggest that a subject's decision to cooperate changes with beliefs about what type of opponent he is facing.

Bereby-Meyer & Roth (2006) compare play in the one-shot PD to play in the finitely repeated PD with either deterministic or stochastic payoffs.[66] The one-shot condition involves 200 rounds with random rematching whereas the finitely repeated PD has 20 supergames with an horizon of 10 rounds. They report more cooperation in round one of the repeated games than in the one-shot games. They also find that in the repeated games, with experience, subjects learn to cooperate more in the early rounds and less towards the end of the supergame. This effect is dampened with stochastic payoffs. They interpret these observations to be consistent with models of reinforcement learning: adding randomness to the link between an action and its consequences, while holding expected payoffs constant, slows learning.

Dal Bó (2005) and Friedman & Oprea (2012) both conduct finitely repeated PD experiments as controls for their respective studies, the first on infinitely repeated games and the second on continuous time games. Dal Bó (2005) looks at two stage-game payoffs with horizon of one, two or four rounds. The main focus of the paper is to compare behavior in finitely repeated games to behavior in randomly terminated repeated games of the same expected length. The results establish that cooperations rates in the first round are much higher when the game is indefinitely repeated. In the finitely repeated games, aggregate cooperation rates decline with experience. Within a supergame, there is a sharp decline in cooperation in the final rounds. However, consistent with previous findings, first round cooperation rates are higher in games with a longer horizon, and the cooperation rates in the four-horizon game is at 20% even after 10 supergames.

Friedman & Oprea (2012) study four stage-game payoffs with an horizon of 8 rounds. They find cooperation rates to increase with experience when payoffs of the stage-game are conducive to cooperation (low temptation to defect, and high efficiency gains from cooperation), but to decrease otherwise. They conclude that "even with ample opportunity to learn, the unraveling process seems at best incomplete in the laboratory data". When behavior in these treatments is compared to the continuous time version with flow payoffs, they find cooperation rates to dramatically increase. They conclude that the unraveling argument of backward induction loses its force when players can react quickly.

[66]In addition, Bereby-Meyer & Roth (2006) also vary the feedback in the stochastic condition.

They formalize this idea in terms of -equilibrium (Radner 1986). Agents determine their optimal first defection point in a supergame by balancing two opposing forces: incentives to become the first defector, and potential losses from preempting one's opponent to start defecting early. The capacity to respond rapidly weakens the first incentive and stabilizes cooperation. Both Friedman & Oprea (2012) and Dal Bó (2005) use a within subjects design, making it difficult to isolate the effect of experience.

In addition to repeated PD experiments, backward induction has been extensively studied in the centipede game.[67] In the many experimental studies on the game, subjects consistently behave in stark contrast to the predictions of backward induction.[68] The pattern of behavior observed in this game share many features to the experimental findings on the finitely repeated PD. First, round 1 behavior diverges from the predictions of subgame-perfection. In the seminal paper on the game, McKelvey & Palfrey (1992) find that even after 10 supergames, less than 10% of subjects choose to stop the game in the first round. Second, the horizon of the centipede game has a significant impact on initial behavior: the stopping rate in the first round is significantly lower in the longer horizon games (less than 2% after 10 supergames.) Third, there is heterogeneity in the subject pool with respect to how behavior changes in response to past experience. While most subjects learn to stop earlier with experience, at the individual level, some subjects never choose to stop despite many opportunities to do so. Motivated by this observation, McKelvey & Palfrey (1992) show that an incomplete information game that assumes the existence of a small proportion of altruists in the population can account for many of the salient features of their data.[69]

Several recent papers study heterogeneity in cooperative behavior and the role of reputation building in the finitely repeated PD. Schneider & Weber (2013) allow players to select the interaction length (horizon of each supergame). They find commitment to long-term relationships to work as a screening device. Conditionally cooperative types are more likely to commit to long term relationships relative to uncooperative types. While longer interactions facilitate more cooperation even when the interaction length is exogenously imposed, endogenously chosen long-term commitment yields even higher cooperation rates.

Kagel & McGee (2016) compare individual play and team play in the finitely-repeated

[67]The standard centipede game consists of two players moving sequentially for a finite number of rounds, deciding on whether to stop or continue the game. In every round, when it is one's turn to make a decision, the payoff from stopping the game is greater than the payoff associated with continuing and letting the opponent stop in the next round, but lower than the payoff associated with stopping the game in two rounds if the game continues that far. Applying backward induction gives the unique subgame perfect Nash equilibrium for the game which dictates the first player to stop in the first round.

[68]McKelvey & Palfrey (1992), Nagel & Tang (1998); Fey et al. (1996); Zauner (1999); Rapoport, Stein, Parco & Nicholas (2003); Bornstein et al. (2004).

[69]Subsequent experimental papers on the centipede game have focused on identifying how beliefs about one's opponent affects play to provide evidence for the reputation hypothesis (Palacios-Huerta & Volij (2009), Levitt, List & Sadoff (2011)).

PD.[70] Although under team play defection occurs earlier and unraveling is faster, cooperation persists in all treatments. Subjects attempt to anticipate when their opponents might defect and try to defect one period earlier, without accounting for the possibility of their opponents thinking similarly. This is interpreted to be consistent with a strong status quo bias in when to defect across super-games. The authors interpret these results as a failure of common knowledge of rationality. Analysis of team dialogues reveal beliefs regarding the strategies of the others to change significantly across supergames. This observation is in contrast to standard models of cooperation in the finitely repeated PD,

Finally, Cox et al. (2015) test the reputation building hypothesis in a sequential-move finitely-repeated PD. Cooperation can be sustained in this setting if the first-mover has uncertainty about the second mover's type. To eliminate this channel, they reveal second-mover histories from an earlier finitely repeated PD experiment to the first-mover. In contradiction to standard reputation-building explanations of cooperation in finitely repeated PDs, they find higher cooperation rates when histories are revealed. They provide a model of semi-rational behavior that is consistent with the pattern of behavior observed in the experiment. According to the model, players use strategies that follow TFT until a predetermined round and then switch to AD. Players decide how long to conditionally cooperate in each supergame based only on *naive* prior beliefs about what strategy their opponent is playing. Similar to the Kagel & McGee (2016) findings, the model does not assume any higher-level reflection about the rationality or best-response of the opponent.[71]

Mao et al. (2017) study long-term behavior in the finitely repeated prisoner's dilemma by running a virtual lab experiment using Amazon's Mechanical Turk in which 94 subjects play up to 400 supergames of a 10-round prisoner's dilemma (with random matching) over the course of twenty consecutive weekdays. While the first defection round moves earlier with experience, partial cooperation mostly stabilizes by the end of the first week. Cooperation is sustained by about 40% of the population who behave as conditional cooperators never preempting defection even when following this strategy comes with significant payoff costs.

[70]In the team play treatments each role is played by two subjects who choose their common action together after free form communication.

[71]Recently, Kamei & Putterman (2015) investigate reputation building in a finitely repeated PD where there is endogenous partner choice, and the parameters of the game allow for substantial gains from cooperation. While subjects repeatedly observe end-game effects, under the right information conditions (how much is revealed about subject's past history of play), learning to invest in building a cooperative reputation becomes the dominant force. This leads to higher cooperation rates with experience.

A.2. Further Analysis of the Meta Data

Henceforth, Andreoni & Miller (1993) will be identified as AM1993, Cooper, DeJong, Fosythe & Ross (1996) as CDFR1996, Dal Bó (2005) as DB2005, Bereby-Meyer & Roth (2006) as BMR2006, and Friedman & Oprea (2012) as FO2012.

Table A1: Summary of Experiments and Sessions Included in the Meta-Study

Experiment	Sessions	Subjects	Supergames	Horizon	g	ℓ	Within-subject variation
DB2005	4	192					horizon
	2	108	8-10	2	0.83	1.17	
	2	84	5-9	2	1.17	0.83	
	2	108	8-10	4	0.83	1.17	
	2	84	5-9	4	1.17	0.83	
FO2005	3	30					stage-game
	3	30	8	8	0.67	0.67	
	3	30	8	8	1.33	0.67	
	3	30	8	8	2.00	4.00	
	3	30	8	8	4.00	4.00	
BMR2006	4	74	20	10	2.33	2.33	
AM1993	1	14	20	10	1.67	1.33	
CDFR1996	3	30	2	10	0.44	0.78	
Total	15	340					

Just over a quarter of the sessions come from BMR2006, which implemented a stage-game with both larger gain and loss parameters. The sessions that implemented a shorter horizon–just over a quarter of the sessions–come from DB2005, which also varied horizon within subject. By varying the stage-game within-subjects, the study of FO2012 includes most of the extreme points of the set of normalized parameter combinations that have been studied.

Table A2: Marginal Effects of Correlated Random Effects Regressions for the *Standard Perspective*. (See last paragraph of page 13.)

	Round 1	Cooperation Rate Last Round	Average	Mean Round to First Defection
g	−0.04*** (0.008)	−0.03*** (0.003)	−0.01 (0.013)	−0.43*** (0.041)
l	−0.02*** (0.005)	−0.01*** (0.002)	−0.05*** (0.004)	−0.16*** (0.032)
Horizon	0.03*** (0.004)	0.01** (0.002)	0.05*** (0.008)	0.37*** (0.058)
Supergame $\times \{H = 2\}$	−0.02*** (0.001)	−0.02*** (0.001)	−0.03*** (0.003)	−0.00 (0.009)
Supergame $\times \{H = 4\}$	−0.00*** (0.001)	−0.01*** (0.001)	−0.01*** (0.001)	−0.05*** (0.009)
Supergame $\times \{H = 8\}$	0.03*** (0.002)	−0.01*** (0.001)	0.00 (0.003)	0.25*** (0.015)
Supergame $\times \{H = 10\}$	0.02*** (0.001)	−0.01*** (0.001)	0.02*** (0.004)	0.21*** (0.011)
Initial Coop. in Supergame 1	0.23*** (0.042)	0.04*** (0.004)	0.16*** (0.025)	0.63*** (0.138)

Notes: For the cooperation rates, the regression model is a probit; for the mean round to first defection, it is linear. Standard errors clustered (at the study level) in parentheses. ***1%, **5%, *10% significance.
The Supergame $\times$ $\mathbf{1}\{\cdot\}$ variable takes the value of the supergame number only for those observations with the relevant horizon.
The total number of supergames varies between 5 to 10 for sessions with $H = 2$ and $H = 4$, is 8 for sessions with $H = 8$, and is either 2 or 20 when $H = 10$.

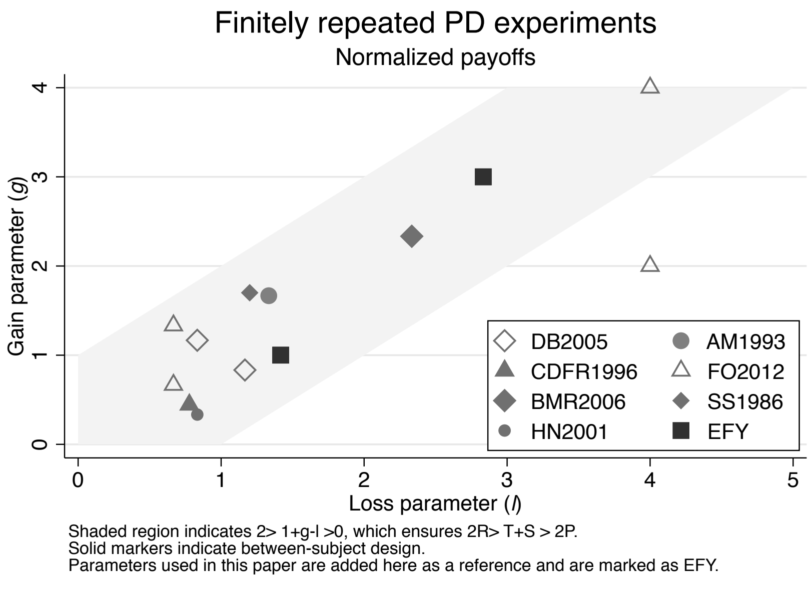


Figure A1: Normalized Game Parameters

The shaded region indicates the set of parameters for which (1) The mutual cooperation payoff is larger than the average of the sucker and temptation payoffs, thus ensuring cooperation is more efficient in the repeated game than any alternating behavior; (2) The mutual defection payoff is lower than the average of the sucker and temptation payoffs, thus ensuring that the average payoff always increases with cooperation. The sixth set of sessions included in the diagram are from our own experiment, labelled EFY.

Table A3: Marginal Effects of Correlated Random Effects Probit Regression of the Probability of Cooperating in Round One. (See Table 2.)

	(1)	(2)
g	-0.04^{***} (0.009)	-0.03^{***} (0.006)
ℓ	-0.02^{***} (0.005)	0.00 (0.005)
Horizon	0.03^{***} (0.004)	0.01 (0.005)
$sizeBAD$		-0.24^{***} (0.025)
Supergame $\times \mathbf{1}\{H=2\}$	-0.02^{***} (0.001)	-0.01^{***} (0.001)
Supergame $\times \mathbf{1}\{H=4\}$	-0.00^{***} (0.001)	-0.01^{***} (0.000)
Supergame $\times \mathbf{1}\{H=8\}$	0.03^{***} (0.002)	0.03^{***} (0.002)
Supergame $\times \mathbf{1}\{H=10\}$	0.02^{***} (0.001)	0.02^{***} (0.001)
Other Initial Coop. in Supergame - 1	0.04^{***} (0.007)	0.04^{***} (0.007)
Initial Coop. in Supergame 1	0.16^{**} (0.049)	0.15^{**} (0.049)
Observations	5398	5398

Notes: Standard errors clustered (at the study level) in parentheses. ***1%, **5%, *10% significance.
The Supergame $\times$ $\mathbf{1}\{\cdot\}$ variable takes the value of the supergame number only for those observations with the relevant horizon.
The total number of supergames varies between 5 to 10 for sessions with $H = 2$ and $H = 4$, is 8 for sessions with $H = 8$, and is either 2 or 20 when $H = 10$.

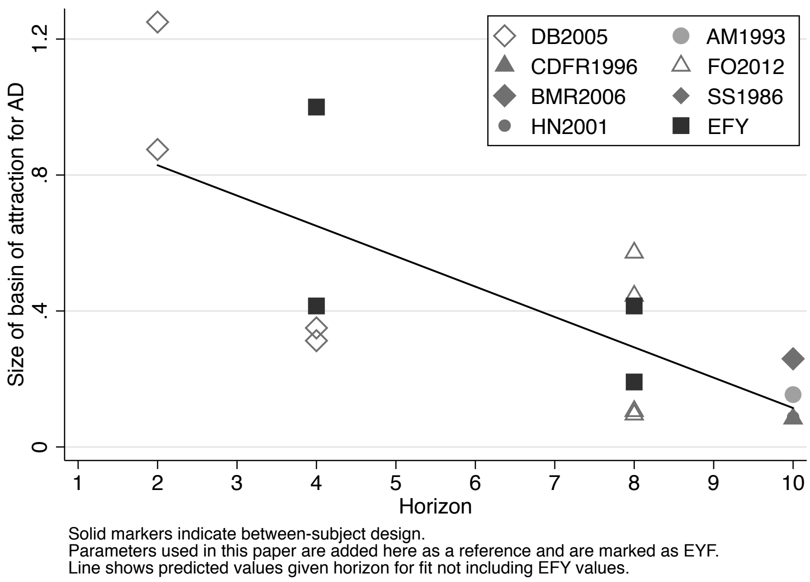
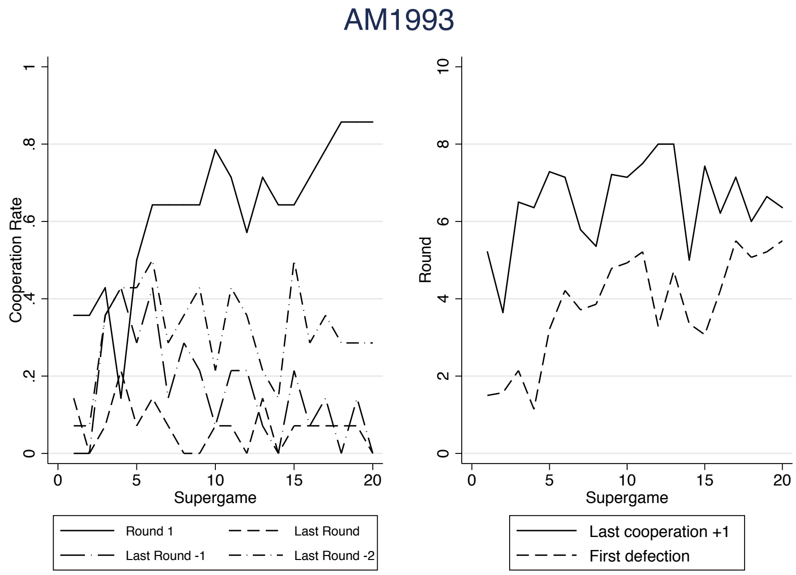


Figure A2: Comparison of the Size of the Basin of Attraction of AD and the Horizon

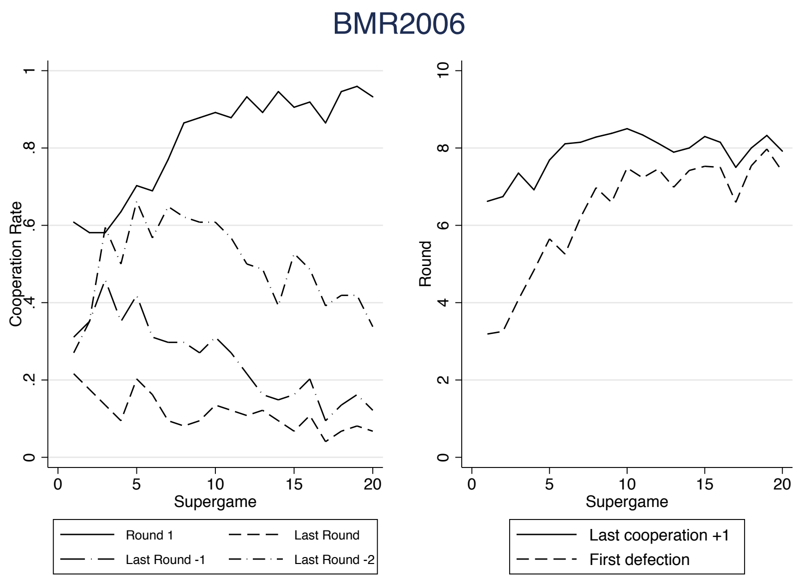
Table A4: Marginal Effects of Correlated Random Effects Probit Regression of the Probability of Cooperating in Round One. (Alternative Specification for Table A3.)

	(1)	(2)
g	-0.10^{***} (0.026)	-0.05^{***} (0.012)
l	0.02^{*} (0.010)	0.04^{***} (0.009)
Horizon	0.03^{***} (0.008)	-0.00 (0.009)
sizebad		-0.35^{***} (0.039)
Other Initial Coop. in Supergame - 1	0.04^{***} (0.008)	0.04^{***} (0.008)
Initial Coop. in Supergame 1	0.16^{***} (0.049)	0.16^{***} (0.051)
Supergame $\times \mathbf{1}\{g = 0.83,\ \ell = 1.17,\ H = 2\}$	-0.02^{***} (0.001)	-0.01^{***} (0.001)
Supergame $\times \mathbf{1}\{g = 1.17,\ \ell = 0.83,\ H = 2\}$	-0.02^{***} (0.003)	-0.00^{***} (0.001)
Supergame $\times \mathbf{1}\{g = 0.83,\ \ell = 1.17,\ H = 4\}$	-0.00^{*} (0.001)	-0.01^{***} (0.001)
Supergame $\times \mathbf{1}\{g = 1.17,\ \ell = 0.83,\ H = 4\}$	-0.01^{***} (0.001)	-0.02^{***} (0.001)
Supergame $\times \mathbf{1}\{g = 0.67,\ \ell = 0.67,\ H = 8\}$	0.03^{***} (0.008)	0.04^{***} (0.005)
Supergame $\times \mathbf{1}\{g = 1.33,\ \ell = 0.67,\ H = 8\}$	0.03^{***} (0.006)	0.03^{***} (0.004)
Supergame $\times \mathbf{1}\{g = 2,\ \ell = 4,\ H = 8\}$	0.01^{***} (0.002)	0.01^{***} (0.002)
Supergame $\times \mathbf{1}\{g = 4,\ \ell = 4,\ H = 8\}$	0.04^{***} (0.008)	0.03^{***} (0.005)
Supergame $\times \mathbf{1}\{g = 0.44,\ \ell = 0.78,\ H = 10\}$	-0.03 (0.041)	0.02 (0.022)
Supergame $\times \mathbf{1}\{g = 1.67,\ \ell = 1.33,\ H = 10\}$	0.02^{***} (0.002)	0.02^{***} (0.001)
Supergame $\times \mathbf{1}\{g = 2.33,\ \ell = 2.33,\ H = 10\}$	0.02^{***} (0.002)	0.03^{***} (0.002)
Observations	5398	5398

Notes: Standard errors clustered (at the study level) in parentheses. ***1%, **5%, *10% significance.
The Supergame $\times \mathbf{1}\{\cdot\}$ variable takes the value of the supergame number only for those observations with the relevant with the relevant parameters.
The total number of supergames varies between 5 to 10 for sessions with $H = 2$ and $H = 4$, is 8 for sessions with $H = 8$, and is either 2 or 20 when $H = 10$.

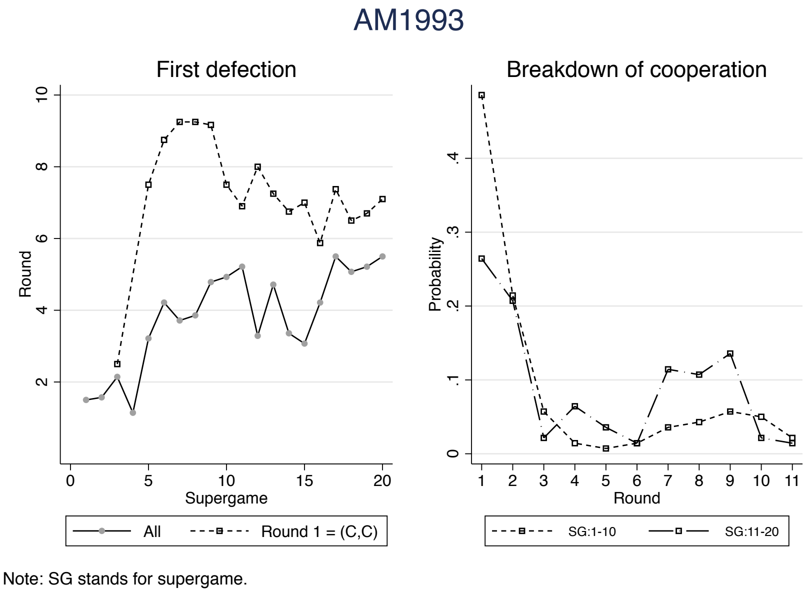


(a)

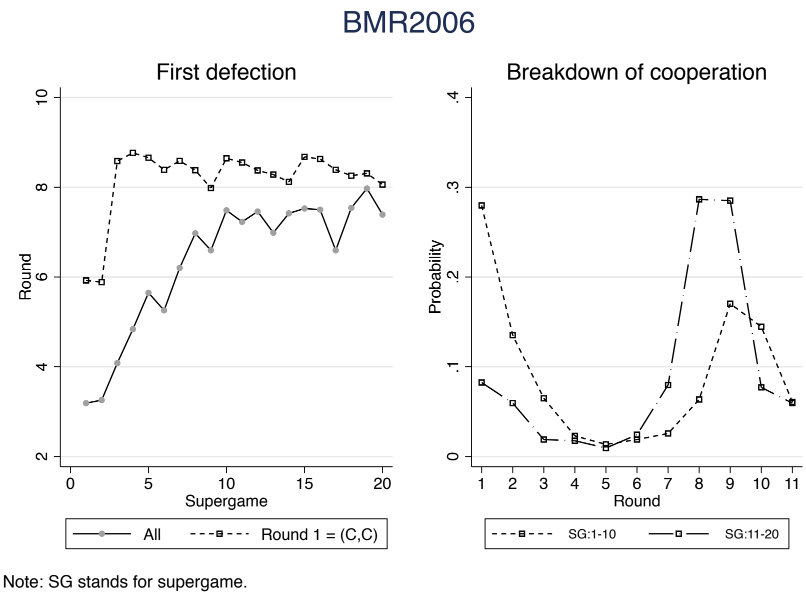


(b)

Figure A3: Evolution of Cooperation by Round and First Defection



(a)



(b)

Figure A4: (LHS) Mean Round to First Defection: All Pairs Versus Those That Cooperated in Round 1; (RHS) Probability of Breakdown in Cooperation

Table A5: Consistency of Play with Threshold Strategies. (See Table 5.)

Experiment	Horizon	g	ℓ	Play Consistent With Threshold Strategy		
				First Supergame		Last Supergame
DB2005	2	1.17	0.83	–		–
	2	0.83	1.17	–		–
	4	1.17	0.83	0.68	$<^{***}$	0.80
	4	0.83	1.17	0.68	$<^{***}$	0.78
FO2012	8	4.00	4.00	0.43	$<^{***}$	0.90
	8	2.00	4.00	0.43	$<^{***}$	0.90
	8	1.33	0.67	0.37	$<^{***}$	0.77
	8	0.67	0.67	0.47	$<^{***}$	0.87
BMR2006	10	2.33	2.33	0.42	$<^{***}$	0.81
AM1993	10	1.67	1.33	0.29	$<^{***}$	0.79
CDFR1996	10	0.44	0.78	0.30	$<^{***}$	0.50
Meta All	.	.	.	0.52	$<^{***}$	0.79
EFY (D4)	4	3.00	2.83	0.66	$<^{***}$	0.94
EFY (E4)	4	1.00	1.42	0.66	$<^{***}$	0.94
EFY (D8)	8	3.00	2.83	0.50	$<^{***}$	0.65
EFY (E8)	8	1.00	1.42	0.57	$<^{***}$	0.89
EFY All	.	.	.	0.60	$<^{***}$	0.85

Notes: Supergame refers to supergame within a set of payoff and horizon parameters. Significance reported using subject random effects with standard errors clustered at the study level. In the meta study, the total number of supergames varies between 5 to 10 for sessions with $H = 2$ and $H = 4$, is 8 for sessions with $H = 8$, and is either 2 or 20 when $H = 10$. For the EFY experiments, the total number of supergames is either 20 or 30 for all parameter combinations.

A.3. Further Analysis of the Experiment

Table A6: Session Characteristics

Treatment	Number of Sessions	Number of Subjects	Earnings Avg (\$)	Earnings Min (\$)	Earnings Max (\$)
D4	3	50	14.67	12.29	17.04
D8	3	54	31.10	27.41	34.46
E4	3	62	14.92	13.34	16.28
E8	3	46	32.83	30.40	34.70

Table A7: Cooperation Rates and Mean Round to First Defection

					Cooperation Rate (%)						Mean Round to First Defection	
					Average		Round 1		Last Round			
Treatment	Supergames	H	g	ℓ	1	L	1	L	1	L	1	L
D4	30	4	3	2.83	0.32	0.07	0.48	0.15	0.14	0.00	2.0	1.1
D8	30	8	3	2.83	0.36	0.33	0.44	0.58	0.22	0.03	2.7	3.0
E4	30	4	3	1.42	0.28	0.20	0.45	0.45	0.18	0.00	1.8	1.7
E8	30	8	1	1.42	0.48	0.52	0.57	0.88	0.26	0.09	3.7	5.1

Notes: First defection is set to Horizon + 1 if there is no defection. 1: First Supergame; L: Supergame 30.

Table A8: Pair-Wise Comparison of Measures of Cooperation Across Treatments.

	All rounds				Round 1				First defect			
	D4	D8	E4	E8	D4	D8	E4	E8	D4	D8	E4	E8
Supergames 1–15												
D4	15.4	$<^{**}$	$<^{*}$	$<^{***}$	29.1	$<^{**}$	$<^{*}$	$<^{***}$	1.5	$<^{***}$	$<^{**}$	$<^{***}$
D8		34.6	$>$	$<^{***}$		49.3	$>$	$<^{***}$		2.8	$>^{***}$	$<^{***}$
E4			28.0	$<^{***}$			49.0	$<^{***}$			1.9	$<^{***}$
E8				60.1				79.7				5.3
Supergames 16–30												
D4	9.0	$<^{***}$	$<^{**}$	$<^{***}$	19.5	$<^{***}$	$<$	$<^{***}$	1.3	$<^{***}$	$<^{**}$	$<^{***}$
D8		33.2	$>^{***}$	$<^{***}$		57.1	$>$	$<^{***}$		3.1	$>^{***}$	$<^{***}$
E4			21.2	$<^{***}$			45.2	$<^{***}$			1.7	$<^{***}$
E8				55.2				88.2				5.3

Notes: The symbol indicates how the cooperation rate of the row treatment compares (statistically) to the column treatment. Significance reported using subject random effects and clustered (session level) standard errors.
***1%, **5%, *10%.

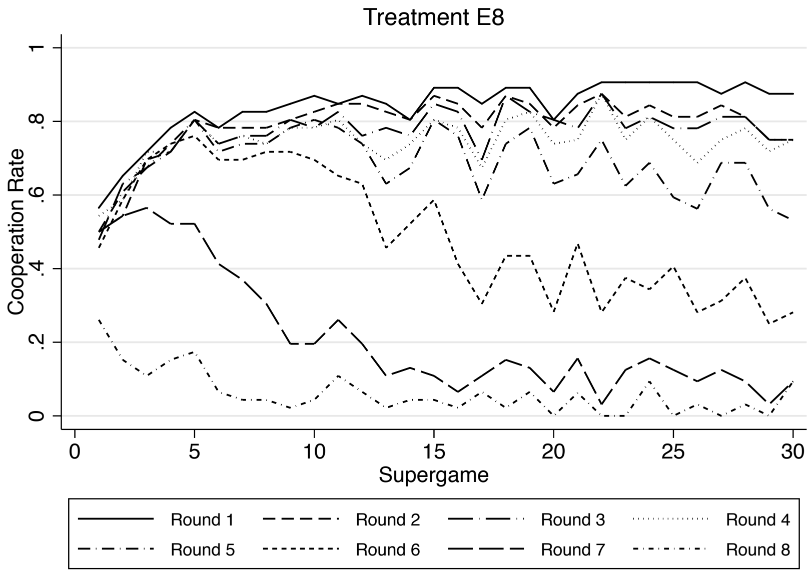
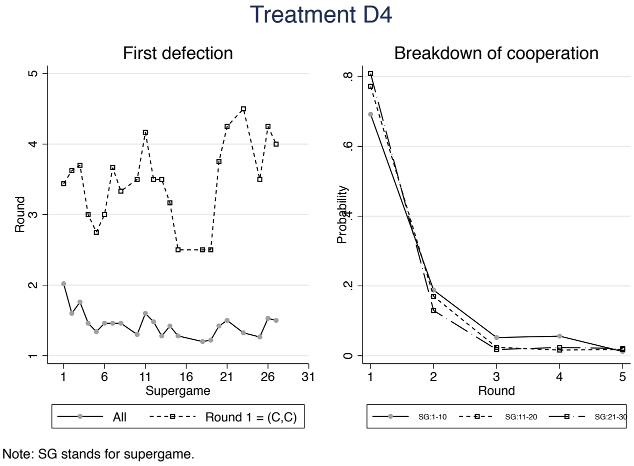
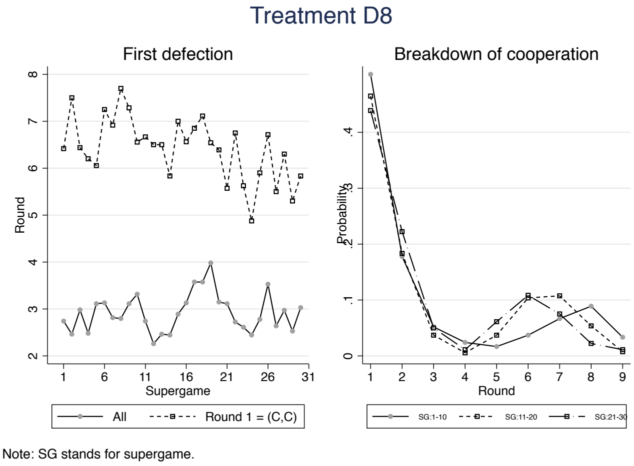


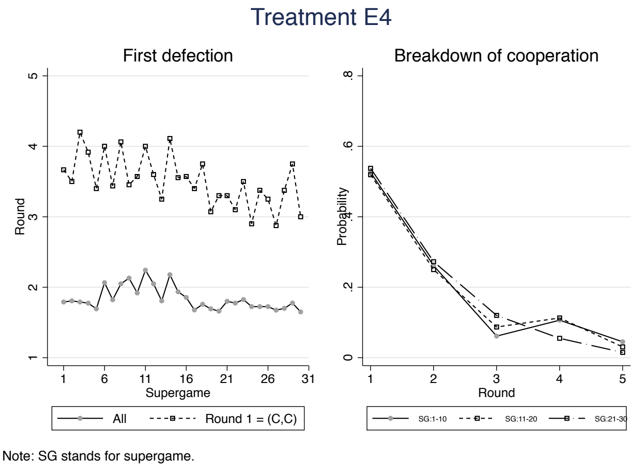
Figure A5: Mean Cooperation Rate by Round. (See Figure 6.)



(a)



(b)



(c)

Figure A6: (LHS) Mean round to first defection: all pairs versus those that cooperated in round 1; (RHS) Probability of breakdown in cooperation

A.4. Robustness Checks: Alternative Specifications to Evaluate Statistical Significance

The data analysis reported in the main body of the text uses two main specifications: probit with subject-level random effects and variance-covariance clustered at the level of the paper for the meta data, and probit with subject-level random effects and variance-covariance clustered at the level of the session for analysis of the data from our own experiment (or for paper specific tests from the meta).[72] These specifications are meant to account for heterogeneity across subjects as well as potential, unmodeled correlations that emerge due to the interactions of subjects within a session, or to study-specific idiosyncrasies (see Fréchette 2012, for a discussion of session-effects). One potential concern with this approach is that having a low number of clusters can lead to corrected standard-errors that do not have the correct coverage probability in finite samples–see, for example, Cameron & Miller (2015) for a recent survey.

Papers that establish the extent of the problem and the effectiveness of various alternatives mostly rely on simulation studies (see, for example, Bertrand et al. 2004, Cameron et al. 2008). These simulations, however, are not geared towards data typically arising from laboratory experiments. For example, the extent of heterogeneity across clusters in the number of observations, the realisation of covariates and the error variance-covariance matrix are all important factors for understanding the potential for over-rejection when using cluster robust standard errors (see, amongst others, Imbens & Kolesar 2016, MacKinnon & Webb 2017, Carter et al. 2017). These are all dimensions on which data from laboratory studies can be expected to differ substantially from the data for which these simulation studies were designed for–indeed, these dimensions are likely to vary between laboratory studies given that details such as matching group size and number, re-matching protocol, and feedback are all experimental design choices.

Nonetheless, this a potential concern, and this appendix explores alternative specifications for the results reported in the paper. One approach is to model within cluster dependency more explicitly. We do this by estimating specifications with paper, session, and subject random effects, or session and subject random effects, as the case may be. Another approach is to remain agnostic about the form of the dependency between observations at the highest level (paper or session), while using bootstrap methods that are designed to provide proper coverage in cases with a small number of clusters. For this, we use a score-based wild bootstrap procedure (Kline et al. 2012) with a six point random weight distribution (Webb 2014).[73] To our knowledge, this is the only bootstrap-based

[72]A few specifications involve the equivalent linear version of these two when the dependent variable is not binary, such as the first round of defection.

[73]For the specifications that use a linear model, a wild bootstrap t-testing procedure (Cameron et al. 2008) is used, again with a six point random weight distribution (Webb 2014). For both the score and wild bootstrap-t procedures, the null hypothesis is imposed before applying random weights to residuals or scores.

procedure developed so far to deal with a small number of clusters that can be used when estimating a probit (see also Cameron & Miller 2015).[74] However, for these specifications we have to drop the subject-level random effect, thus ignoring a main feature of the panel structure of the data. We note that we find a great deal of evidence for the importance of subject-level random effects in our data, which are typically more important than session or paper level effects in the models that we estimate with multiple levels. By not explicitly taking into account an important source of within cluster error correlation, this potentially magnifies the small cluster problem. Nonetheless, this agnostic specification provides a useful benchmark as the non-panel estimator of the coefficients is necessarily less efficient than the panel estimator under the usual exogeneity assumption of the random effects model.

The tables in this appendix reproduce all of the main statistical tests. All tables report the p-value for the t-test of the approach in the main text (labeled CR-t for cluster robust), the p-value for the t-test of the multi-level random effects model (labeled RE-t for random effect model for the highest cluster level), and the p-value for the score-based wild bootstrap t-test of the probit specification (labeled Bt-t for bootstrap). In the few cases where the dependent variable is not dichotomous, the linear version of these is reported. In cases where estimates of the regression are of interest, we also report the respective marginal effects, to show how the magnitude of the estimated effects vary with the specification. Note that the p-values are not of the marginal effects, but of the actual coefficients from the underlying model estimated.

Although the p-values vary with the estimation method, the main results of the paper remain. For instance, here are some of the important results: The fact that most of the impact of the horizon on round one cooperation rates in the meta is absorbed by *sizeBAD* remains true in all estimations (see Table A10). The finding that round one cooperation rates are not statistically different when comparing treatments D8 and E4 from our experiment is true in all specifications (see the D8 vs E4 rows for the Round 1 block in Table A14, which shows this separately for early and late supergames; the result also holds combining all supergames, with p-values for the CR-t, Bt-t and RE-t tests of 0.45, 0.54 and 0.61, respectively). The observation that the play of threshold strategies increases between the first and last supergame of a session is true in all specifications for the data of our experiment, as well as the meta data.

[74]An alternative bootstrap method that is generally applicable for a wide variety of estimators is the pairs cluster bootstrap, which resamples with replacement from the sample of clusters. However, with very few clusters this method can run into a number of implementation problems. See, for example, Cameron & Miller (2015) for details. Another alternative is to use the linear probability model instead of the probit, and then use the more commonly applied wild bootstrap t-testing procedure (Cameron et al. 2008), again with a six point random weight distribution (Webb 2014). Given this approach did not produce any notable differences in robustness of the main results–as well as the potential problems for the linear probability model when the regressors are no longer just a complete set of treatment indicator variables, as is the case with regressions including g, ℓ and $Horizon$–we only report the results of a bootstrap method that keeps the functional form of the limited dependent variable model fixed throughout.

A.4.1 Meta Data

In addition to what is described above, Tables A9 and A10 also report the p-value for the t-test on the estimated coefficients of the non-panel probit model using the standard cluster robust variance-covariance estimator (in addition to the bootstrapped version). This is to give a sense of what drives the changes between the random effects probit CR-t in the text and the probit Bt-t: part of it is from the bootstrapping, but part of it is simply the result of dropping the subject random effects. Table A9 shows variations across specifications. In particular, none of g, , or $Horizon$ are statistically significant when using bootstrapped standard errors for any of round one, the last round, all rounds, or the round of first defection. On the other hand, the multi-level random effects almost exclusively finds statistically significant effects. Importantly, note that the lack of significance when bootstrapping does not mean that g, , and $Horizon$ do not matter as the next table makes it clear.

Indeed, Table A10 revisits the estimation of the determinants of round one cooperation controlling for experience. The main result is the significance of $sizeBAD$ in all specifications. This confirms that the effect of g, , and $Horizon$ can be summarized by how it affects the value of cooperation. Clearly there could be additional effects of these parameters that $sizeBAD$ does not fully capture, but it accounts for an important part of the variation.

Table A9: Alternative Specifications for Table A2: Marginal Effects of Correlated Random Effects Regressions for the *Standard Perspective*

	RE Probit		Multiple REs		Probit		
	ME	CR-t	ME	RE-t	ME	CR-t	Bt-t
Round 1							
g	−0.04	0.00	−0.05	0.00	−0.04	0.00	0.31
ℓ	−0.02	0.00	−0.02	0.07	−0.01	0.77	0.92
Horizon	0.03	0.00	0.03	0.00	0.02	0.00	0.12
Last Round							
g	−0.03	0.00	−0.02	0.04	−0.02	0.00	0.17
ℓ	−0.01	0.00	−0.01	0.36	−0.01	0.30	0.66
Horizon	0.01	0.01	0.00	0.10	0.00	0.13	0.23
All Rounds							
g	−0.04	0.00	−0.04	0.00	−0.03	0.00	0.39
ℓ	−0.03	0.00	−0.03	0.00	−0.02	0.47	0.83
Horizon	0.04	0.00	−0.02	0.00	0.03	0.00	0.17
First Defect							
g	−0.43	0.00	−0.46	0.00	−0.32	0.06	0.37
ℓ	−0.16	0.00	−0.17	0.01	−0.06	0.80	0.78
Horizon	0.37	0.00	0.29	0.00	0.33	0.02	0.12

Notes: Additional controls include experience variables (supergame interacted with horizon) and an indicator variable for whether the player cooperated initially in the first supergame. The ME columns give the average marginal effect of each explanatory variable.

Table A10: Alternative Specifications for Table A4: Marginal Effects of Correlated Random Effects Probit Regression of the Probability of Cooperating in Round One

	RE Probit		Multiple REs		Probit		
	ME	CR-t	ME	RE-t	ME	CR-t	Bt-t
Independent Variable Specification (1)							
g	−0.10	0.00	−0.11	0.00	−0.09	0.00	0.32
ℓ	0.02	0.09	0.02	0.48	0.03	0.17	0.35
Horizon	0.03	0.00	0.04	0.00	0.03	0.00	0.14
Independent Variable Specification (2)							
g	−0.05	0.00	−0.05	0.16	−0.04	0.13	0.80
ℓ	0.04	0.00	0.05	0.10	0.04	0.01	0.16
Horizon	−0.00	0.60	−0.01	0.65	−0.01	0.48	0.48
$sizeBAD$	−0.35	0.00	−0.39	0.00	−0.36	0.00	0.10

Notes: Additional controls include experience variables (supergame interacted with each combination of stage-game and horizon parameters) and choice history variables (whether the player cooperated in the first supergame and whether the player they were matched with cooperated in the round one of the last supergame). The ME columns give the average marginal effect of each explanatory variable.

Table A11: Alternative Specifications for Table A5: Consistency of Play with Threshold Strategies

				Play Consistent With Threshold Strategy			
Experiment	Horizon	g	ℓ	1 v L Difference	CR-t	Bt-t	RE-t
DB2005	4	1.17	0.83	0.12	0.00	0.86	0.05
DB2005	4	0.83	1.17	0.10	0.00	0.34	0.06
FO2012	8	4.00	4.00	0.47	0.00	0.16	0.00
FO2012	8	2.00	4.00	0.47	0.00	0.33	0.00
FO2012	8	1.33	0.67	0.40	0.00	0.84	0.00
FO2012	8	0.67	0.67	0.40	0.00	0.33	0.00
BMR2006	10	2.33	2.33	0.39	0.00	0.50	0.00
AM1993	10	1.67	1.33	0.50	0.00	0.66	0.00
CDFR1996	10	0.44	0.78	0.20	0.00	0.34	0.08
Meta All	.	.	.	0.27	0.00	0.05	0.00
EFY	4	3.00	2.83	0.28	0.00	0.09	0.00
EFY	4	1.00	1.42	0.27	0.00	0.12	0.00
EFY	8	3.00	2.83	0.15	0.00	0.12	0.11
EFY	8	1.00	1.42	0.33	0.01	0.12	0.00
EFY All	.	.	.	0.25	0.00	0.00	0.00

Notes: The 1 v L Difference column gives the difference between the first and last supergames. Supergame refers to supergame within a set of payoff and horizon parameters. Where possible, the CR-t and Bt-t columns use standard errors clustered at the session level (for AM1993 row, standard errors are clustered at subject level since there is only one session; for the Meta All row, standard errors are clustered at the study level). In the meta study, the total number of supergames varies between 5 to 10 for sessions with $H = 2$ and $H = 4$, is 8 for sessions with $H = 8$, and is either 2 or 20 when $H = 10$. For the EFY experiments, the total number of supergames is either 20 or 30 for all parameter combinations.

A.4.2 Experiment Data

Table A12: Alternative Specifications for Table 3: Cooperation Rates: Early Supergames (1–15) vs Late Supergames (16–30)

	Round 1				Last Round			
Treatment	Diff	CR-t	Bt-t	RE-t	Diff	CR-t	Bt-t	RE-t
D4	−9.6	0.20	0.33	0.00	−0.9	0.05	0.13	0.25
D8	7.9	0.01	0.09	0.00	−3.9	0.00	0.34	0.00
E4	−3.8	0.26	0.43	0.03	−6.6	0.00	0.13	0.00
E8	8.5	0.00	0.08	0.00	−5.7	0.00	0.09	0.00
4	−6.6	0.09	0.18	0.00	−4.1	0.00	0.02	0.00
8	8.3	0.00	0.10	0.00	−4.8	0.00	0.03	0.00
All	0.6	0.67	0.86	0.23	−4.4	0.00	0.00	0.00
	All Rounds				First defect			
	Diff	CR-t	Bt-t	RE-t	Diff	CR-t	Bt-t	RE-t
D4	−6.3	0.03	0.13	0.00	−0.2	0.24	0.36	0.16
D8	−1.4	0.26	0.85	0.00	0.3	0.21	0.36	0.00
E4	−6.8	0.00	0.09	0.00	−0.2	0.02	0.12	0.01
E8	−4.9	0.01	0.31	0.00	−0.0	0.74	0.84	0.56
4	−6.7	0.00	0.02	0.00	−0.2	0.02	0.05	0.01
8	−2.9	0.02	0.45	0.00	0.2	0.40	0.62	0.05
All	−4.1	0.00	0.12	0.00	−0.0	0.78	0.97	0.55

Notes: For the cooperation measures, the regression model is a random effects probit on an indicator variable for late supergames, with standard errors clustered at the session level; for first defect, the regression model is a linear equivalent. The Diff column gives the difference in the measure between early and late supergames.

Table A13: Alternative Specifications for Table 4: Cooperation Rate for All Rounds in Supergames 1, 2, 8, 20 and 30

	SG 1 vs. SG 2				SG 1 vs. SG 8			
Treatment	Diff	CR-t	Bt-t	RE-t	Diff	CR-t	Bt-t	RE-t
D4	−10.5	0.01	0.12	0.01	−19.0	0.00	0.11	0.00
D8	0.5	0.92	0.89	0.91	−0.7	0.93	0.96	0.77
E4	1.6	0.72	0.64	0.75	2.0	0.79	0.87	0.74
E8	6.2	0.04	0.13	0.06	13.9	0.00	0.08	0.00
4	−3.8	0.25	0.35	0.14	−7.4	0.15	0.21	0.00
8	3.1	0.23	0.28	0.16	6.0	0.35	0.38	0.01
	SG 1 vs. SG 20				SG 1 vs. SG 30			
Treatment	Diff	CR-t	Bt-t	RE-t	Diff	CR-t	Bt-t	RE-t
D4	−20.0	0.01	0.07	0.00	−24.9	0.00	0.12	0.00
D8	−1.2	0.86	0.88	0.71	−3.7	0.01	0.62	0.00
E4	−8.9	0.20	0.38	0.01	−8.2	0.07	0.36	0.03
E8	3.8	0.64	0.73	0.26	4.0	0.18	0.64	0.16
4	−13.8	0.01	0.07	0.00	−15.8	0.00	0.04	0.00
8	1.1	0.84	0.85	0.62	0.0	0.59	1.00	0.26

Notes: The regression model is a random effects probit on an indicator variable for the later supergame, with standard errors clustered at the session level. The Diff column gives the difference in the all-rounds cooperation rate in supergame 1 versus the later supergame.

Table A14: Alternative Specifications for Table A8: Pair-Wise Comparison of Measures of Cooperation Across Treatments

	Round 1				Last Round			
	Diff	CR-t	Bt-t	RE-t	Diff	CR-t	Bt-t	RE-t
Supergames 1–15								
D4 vs D8	−20.2	0.05	0.10	0.01	−3.8	0.08	0.19	0.05
D4 vs E4	−20.0	0.07	0.13	0.01	−6.3	0.00	0.01	0.00
D4 vs E8	−50.6	0.00	0.01	0.00	−4.9	0.01	0.05	0.02
D8 vs E4	0.2	0.91	0.97	0.94	−2.5	0.13	0.36	0.18
D8 vs E8	−30.5	0.00	0.01	0.00	−1.1	0.62	0.68	0.69
E4 vs E8	−30.7	0.00	0.01	0.00	1.4	0.07	0.35	0.38
Supergames 16–30								
D4 vs D8	−37.7	0.01	0.02	0.00	−0.7	0.65	0.72	0.60
D4 vs E4	−25.7	0.11	0.02	0.01	−0.6	0.58	0.78	0.56
D4 vs E8	−68.7	0.00	0.01	0.00	−0.1	0.68	0.97	0.73
D8 vs E4	11.9	0.17	0.26	0.12	0.2	0.97	0.94	0.97
D8 vs E8	−31.0	0.00	0.01	0.00	0.7	0.85	0.75	0.87
E4 vs E8	−43.0	0.00	0.01	0.00	0.5	0.76	0.75	0.84
	All Rounds				First Defect			
	Diff	CR-t	Bt-t	RE-t	Diff	CR-t	Bt-t	RE-t
Supergames 1–15								
D4 vs D8	−19.2	0.01	0.03	0.00	−1.3	0.00	0.01	0.00
D4 vs E4	−12.6	0.07	0.09	0.02	−0.5	0.02	0.11	0.06
D4 vs E8	−44.7	0.00	0.01	0.00	−3.9	0.00	0.01	0.00
D8 vs E4	6.6	0.16	0.26	0.26	0.8	0.00	0.01	0.00
D8 vs E8	−25.5	0.00	0.01	0.00	−2.5	0.00	0.01	0.00
E4 vs E8	−32.1	0.00	0.01	0.00	−3.4	0.00	0.01	0.00
Supergames 16–30								
D4 vs D8	−24.1	0.00	0.02	0.00	−1.8	0.00	0.01	0.00
D4 vs E4	−12.2	0.02	0.09	0.00	−0.4	0.03	0.16	0.16
D4 vs E8	−46.2	0.00	0.02	0.00	−4.0	0.00	0.01	0.00
D8 vs E4	12.0	0.00	0.06	0.01	1.4	0.00	0.01	0.00
D8 vs E8	−22.0	0.00	0.02	0.00	−2.2	0.00	0.02	0.00
E4 vs E8	−34.0	0.00	0.01	0.00	−3.6	0.00	0.01	0.00

Notes: For all cooperation measures, the regression model is a random effects probit on a complete set of treatment dummies, with standard errors clustered at the session level; for first defect, the model is the linear equivalent. The Diff column gives the difference in the measure between the measures for the comparison treatments.

A.5. Further Details and Analysis of the Learning Model

A.5.1 Estimates

Tables A15 and A16 report summary statistics for the estimates of the learning model for each treatment. To facilitate comparison, the parameters representing initial beliefs in supergame 1 are normalized. $\bar{\ }$ denotes $\sum_k {}_{k0}$, as defined in the learning model. Using this, $\tilde{\ }_k = \frac{\beta_{k0}}{\beta}$ so that $\sum_k \tilde{\ }_k = 1$.

E8

Variable	Mean	Std. Dev.
	0.83	0.86
	0.83	0.22
σ	0.16	0.17
κ	4.22	2.76
$\bar{\ }$	4.45	2.92
$\tilde{\ }_1$	0.14	0.27
$\tilde{\ }_2$	0.03	0.07
$\tilde{\ }_3$	0.03	0.06
$\tilde{\ }_4$	0.06	0.11
$\tilde{\ }_5$	0.08	0.19
$\tilde{\ }_6$	0.04	0.06
$\tilde{\ }_7$	0.06	0.1
$\tilde{\ }_8$	0.1	0.12
$\tilde{\ }_9$	0.07	0.11
$\tilde{\ }_{10}$	0.16	0.24
$\tilde{\ }_{11}$	0.21	0.17
ll	54.77	28.77

D8

Variable	Mean	Std. Dev.
	2.68	4.9
	0.62	0.34
σ	0.22	0.18
κ	$33.\times 10^{12}$	$233.\times 10^{12}$
$\bar{\ }$	9.04	5.36
$\tilde{\ }_1$	0.31	0.39
$\tilde{\ }_2$	0.07	0.19
$\tilde{\ }_3$	0.02	0.09
$\tilde{\ }_4$	0.02	0.05
$\tilde{\ }_5$	0	0.01
$\tilde{\ }_6$	0.02	0.04
$\tilde{\ }_7$	0.03	0.06
$\tilde{\ }_8$	0.05	0.14
$\tilde{\ }_9$	0.09	0.18
$\tilde{\ }_{10}$	0.12	0.14
$\tilde{\ }_{11}$	0.26	0.3
ll	91.12	43.92

Table A15: Summary statistics for long horizon treatments

D4

Variable	Mean	Std. Dev.
	2.2	4.83
	0.71	0.32
σ	0.16	0.18
κ	1.12	7.88
$\bar{}$	$4.\times10^{14}$	3.24
$\tilde{}_{1}$	0.32	0.35
$\tilde{}_{2}$	0.07	0.19
$\tilde{}_{3}$	0.05	0.16
$\tilde{}_{4}$	0.05	0.09
$\tilde{}_{5}$	0.05	0.08
$\tilde{}_{6}$	0.12	0.19
$\tilde{}_{7}$	0.34	0.32
ll	29.39	19.02

E4

Variable	Mean	Std. Dev.
	8.17	11.9
	0.46	0.31
σ	0.22	0.19
κ	$31.\times10^{12}$	$17.\times10^{13}$
$\bar{}$	1.98	1.01
$\tilde{}_{1}$	0.15	0.32
$\tilde{}_{2}$	0.18	0.33
$\tilde{}_{3}$	0.02	0.04
$\tilde{}_{4}$	0.09	0.19
$\tilde{}_{5}$	0.11	0.25
$\tilde{}_{6}$	0.16	0.31
$\tilde{}_{7}$	0.29	0.37
ll	27.64	18.76

Table A16: Summary statistics for short horizon treatments

A.5.2 Figures

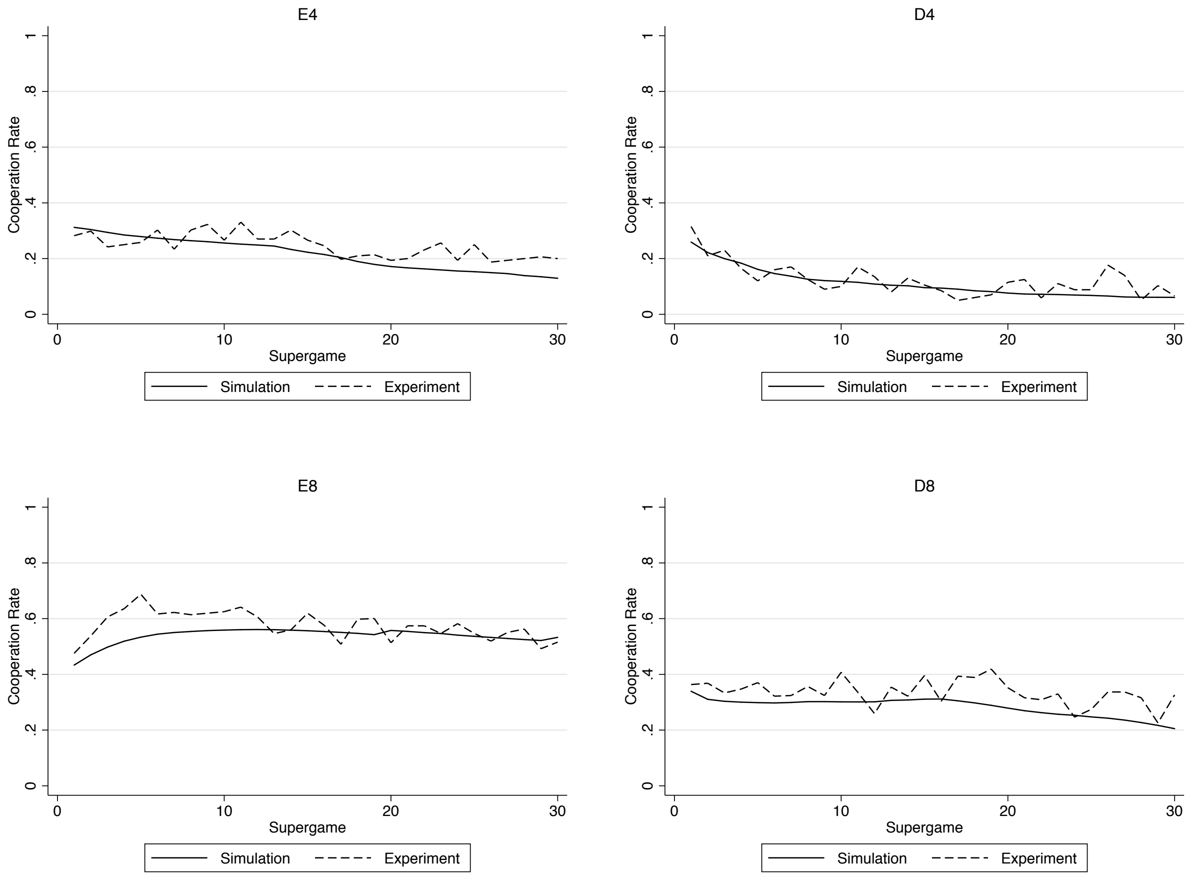


Figure A7: Average Cooperation: Simulation Versus Experimental data

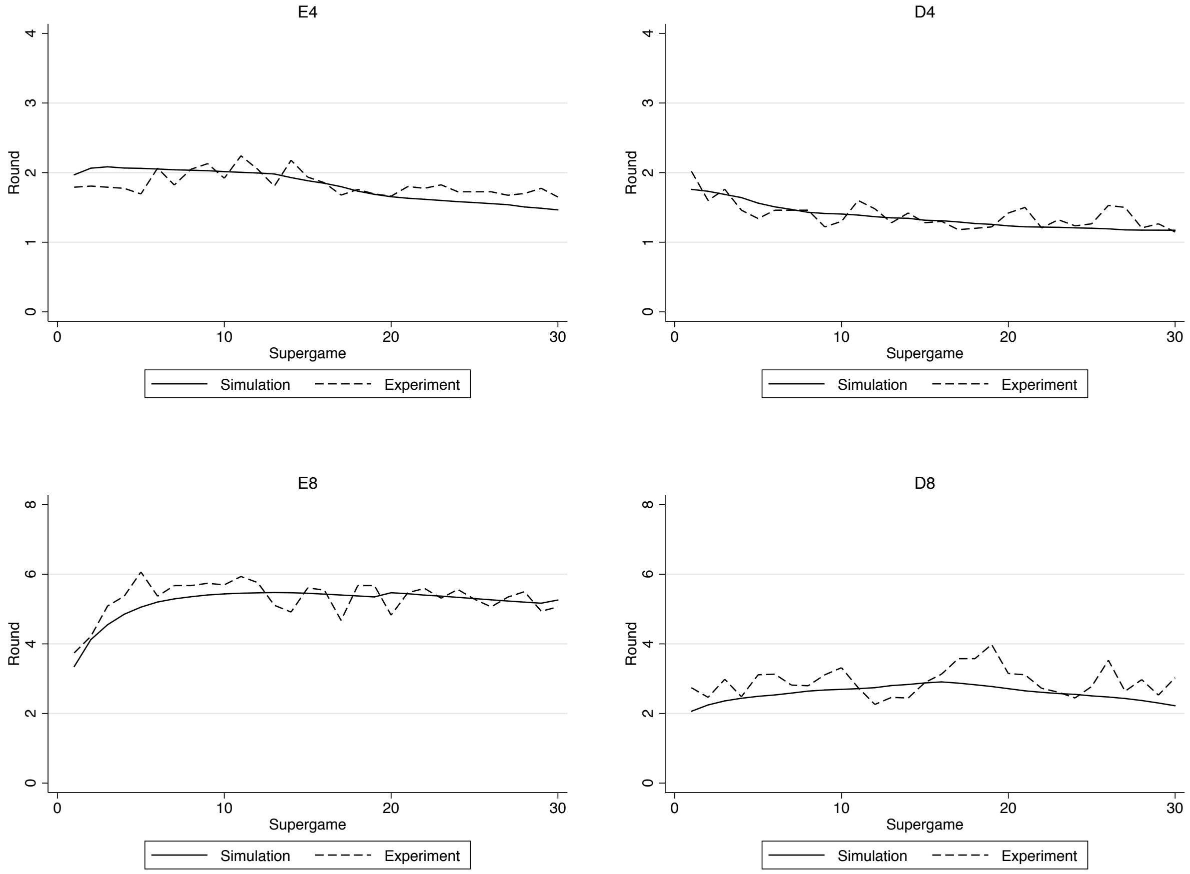


Figure A8: Mean Round to First Defection by Supergame: Simulation versus Experimental Data

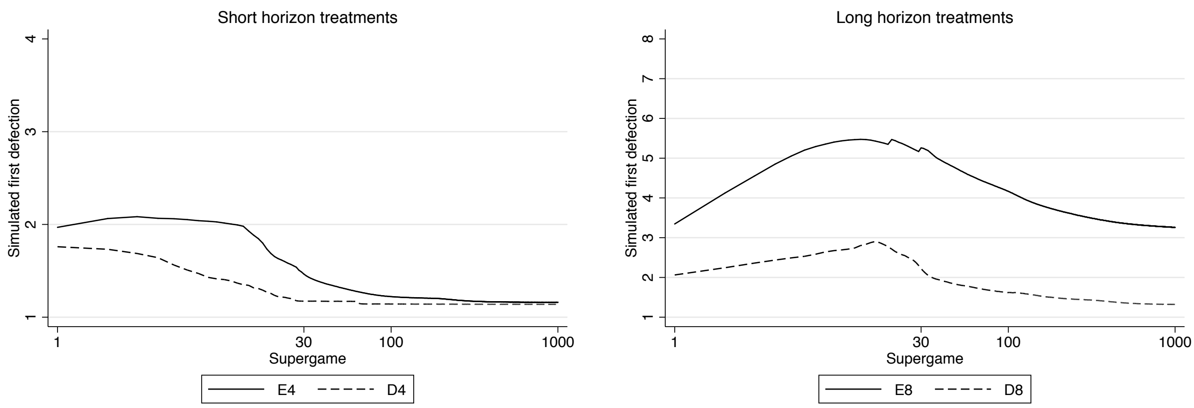
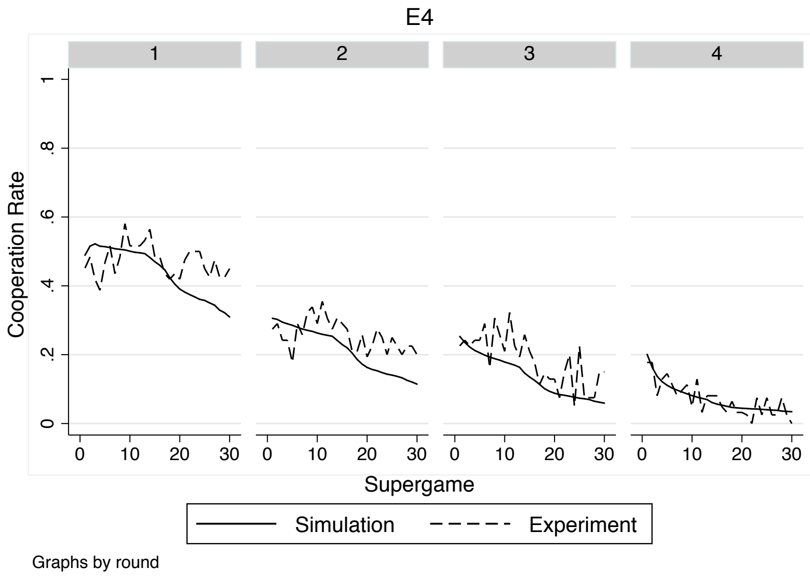


Figure A9: Long Term Evolution of Mean Round to First Defection by Supergame



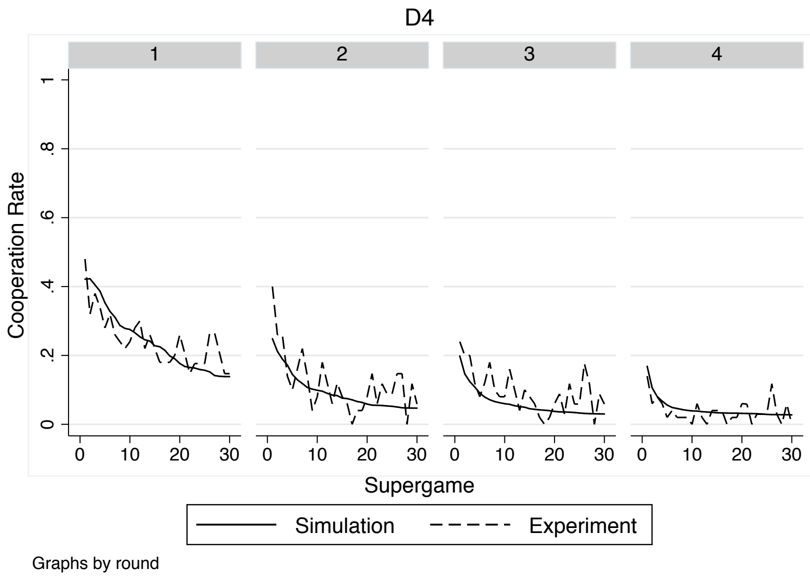
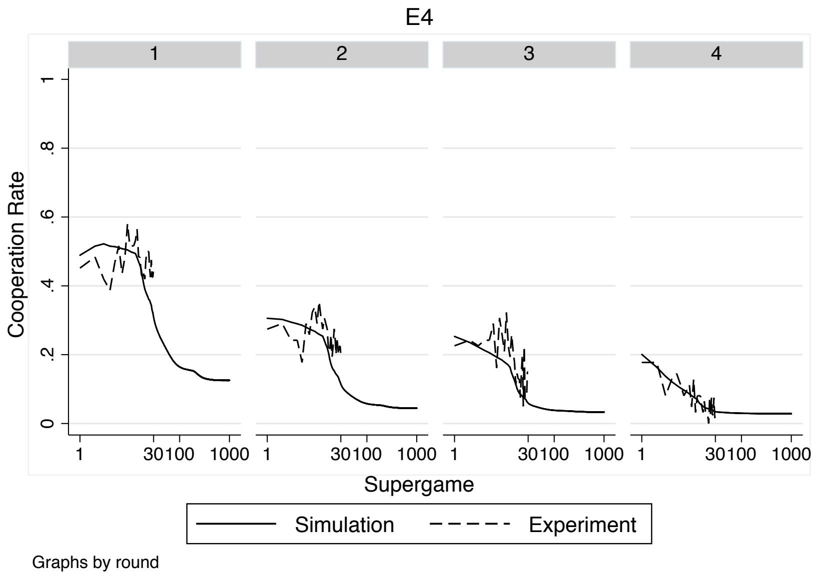


Figure A10: Average Cooperation Rate by Supergame: Simulation versus Experimental Data for Each Round in the Short Horizon Treatments



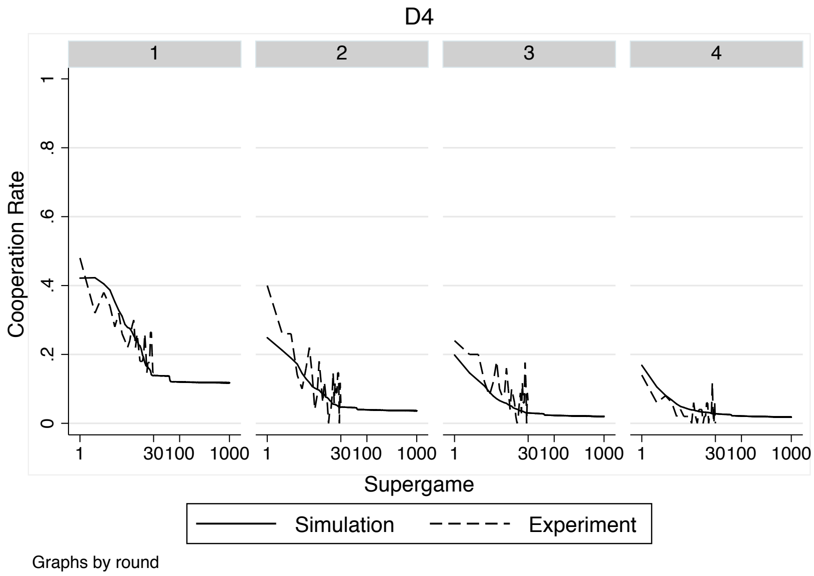


Figure A11: Long Term Evolution of Cooperation Rate for Each Round of the Short Horizon Treatments

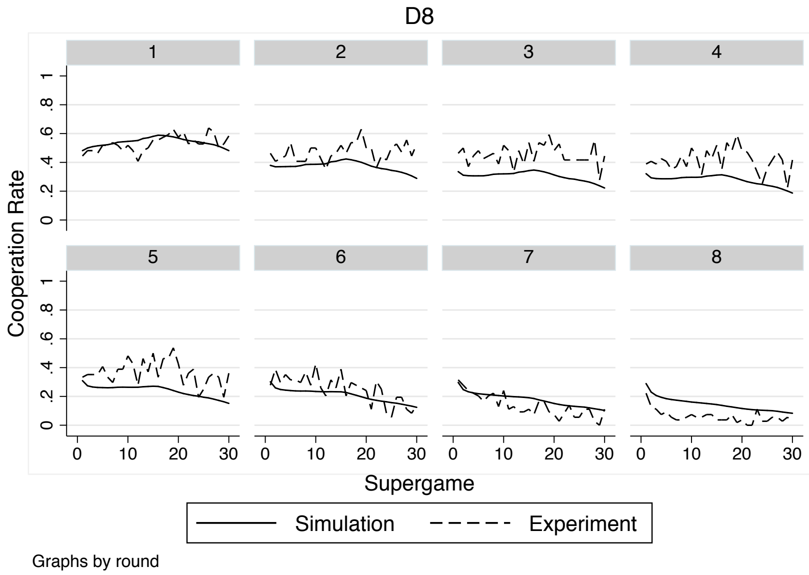


Figure A12: Average Cooperation Rate by Supergame: Simulation versus Experimental Data for Each Round in D8

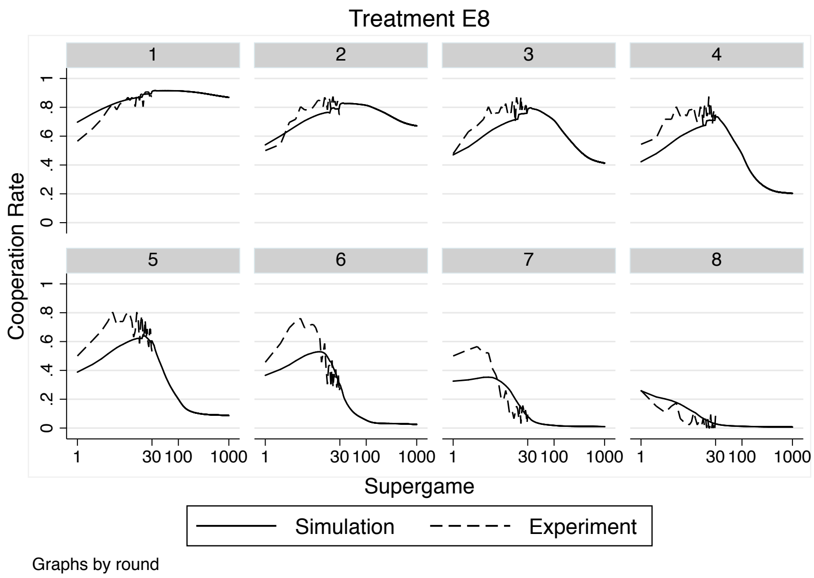


Figure A13: Long Term Evolution of Aggregate cooperation For Each Round In E8

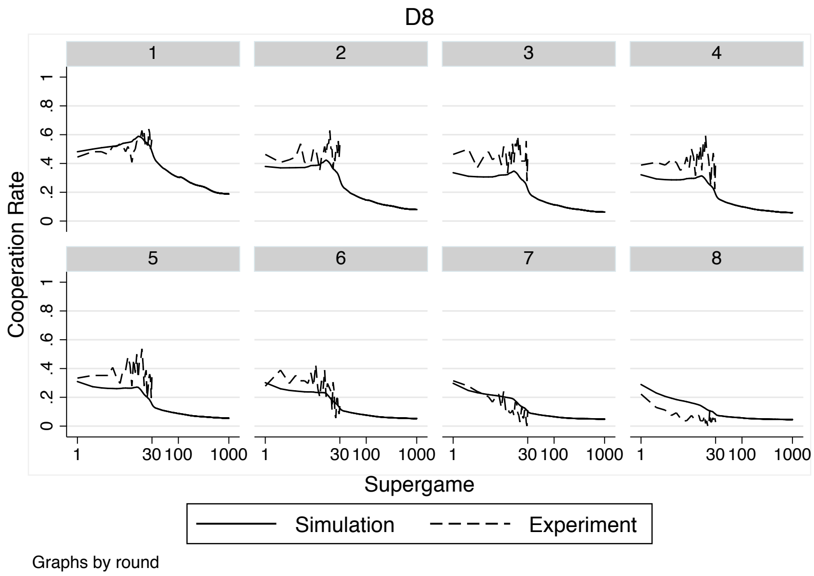


Figure A14: Long Term Evolution of Aggregate cooperation For Each Round In D8

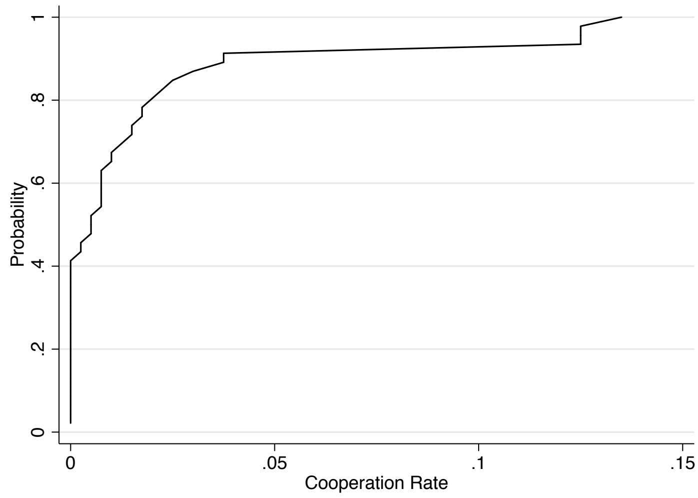


Figure A15: Cumulative Distribution of Cooperation Against an AD type In E8 (Supergames 250-300)

Each subject is simulated to play against an AD type–someone who defects in all rounds of a supergame regardless of past experience–for 300 supergames. The average cooperation rate for the subject from supergames 250-300 is taken as a measure of that subject's cooperativeness. Such a measure of cooperativeness combines the effects of the parameters estimated in the model in an intuitive way. It effectively captures how well a subject is able to learn to defect against a defector.[75,76]

Figure A15 plots the cumulative distribution of simulated cooperation rates after 250 supergames against a player who is following the AD strategy. The distribution has a mass point around 0 implying that about 40% of the subjects learn to defect perfectly with sufficient experience in this environment. There is limited but positive levels of cooperation for the remaining subjects. Note that this corresponds to subjects making cooperative choices after observing their partners defecting in every single round of 250 supergames; hence, this suggests the existence of cooperative types. The model allows for multiple kinds of cooperative types: some forces that can drive cooperative actions in such an extreme environment are strong priors, limited learning from past experiences, and noise in strategy choice and implementation.

[75]An horizon of 250-300 is chosen to correspond to the time frame we are analyzing in what follows, but the exercise can easily be repeated for a different range of supergames. Looking at cooperation rates in supergames 900-1000 gives very similar results.

[76]Focusing on cooperation in later supergames also dampens the effect of a strong prior and execution noise in early supergames. This exercise can be repeated by constructing a measure of cooperativeness by focusing on behavior in early supergames. As expected, removing subjects based on such a measure has a bigger impact on cooperation in earlier supergames, but the effect quickly disappears with experience.

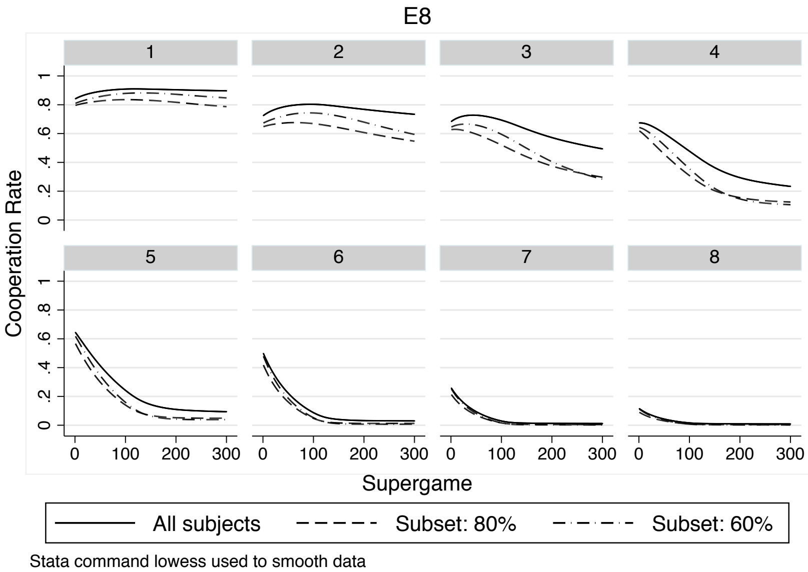


Figure A16: Long term Evolution of Aggregate Cooperation For Each Round In E8 By Subset

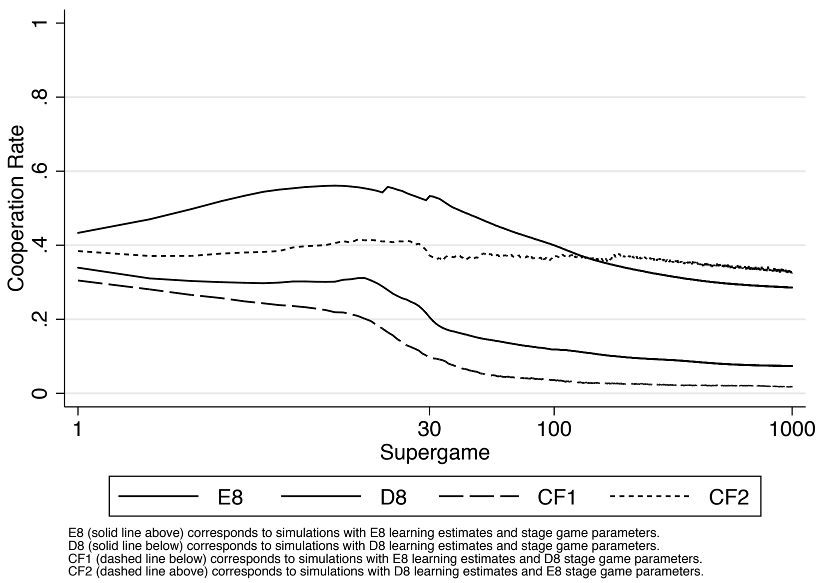


Figure A17: Long term Evolution of Aggregate Cooperation

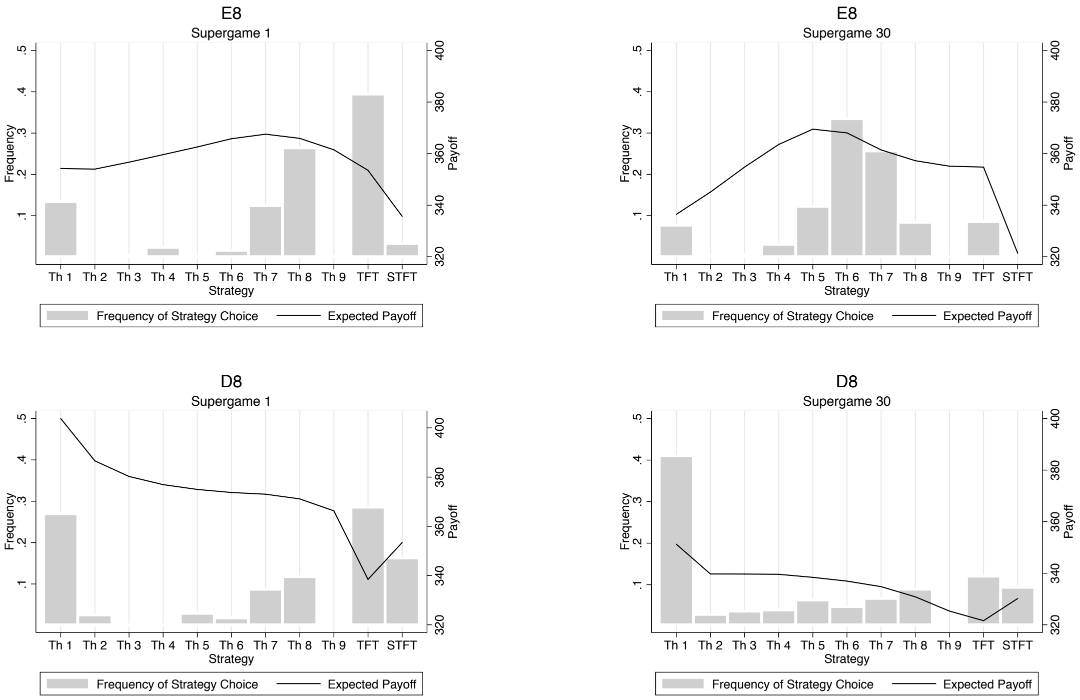


Figure A18: Frequency and Expected Payoff of Each Strategy

These values are estimated by simulating behavior in 1000 sessions composed of 14 randomly drawn subjects. The frequency of choice for each strategy is recorded, along with how well each strategy performs when played against each subject of the session.

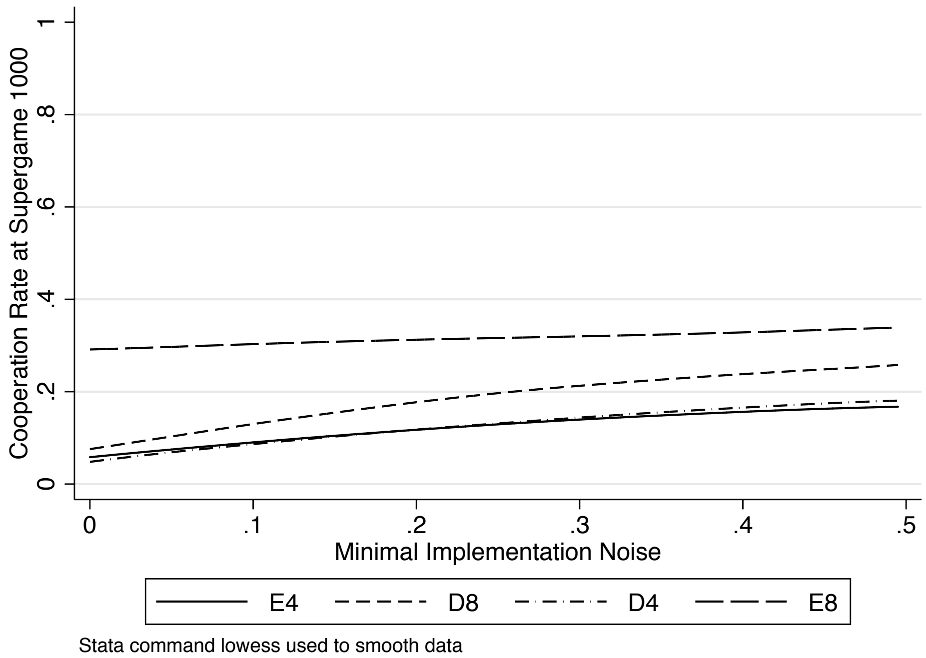


Figure A19: Effects of Constraining the Decline (with Experience) of Implementation Noise

One additional concern may be the robustness of the results to specific parameters. In particular, one may wonder what happens in the long run, if implementation error is not allowed to completely disappear with experience. To explore this possibility, we conduct additional simulations constraining how much the implementation error can decline as as a result of learning (through the κ parameter). Formally, if the constraint is set to σ_{min}, implementation noise in supergame t for subject i is calculated to be $\max\{\min(\sigma_{min}, \sigma_i), \sigma_i^{t^{\kappa_i}}\}$. According to this specification, σ_{min} does not constrain initial implementation noise σ_i, but limits how much it can decline over time with experience through the κ parameter. We recover our original simulation results when the constraint is never binding (set to 0), and we see what long term cooperation results would be like if the implementation noise never changed (corresponding to the case where the constraint is set to 0.5, which is equivalent to setting $\kappa_i = 0$ for all subjects).

The results show this constraint to have little effect on long term cooperation rates in the E8 treatment. In the case of treatments E4 and D4, looking directly at the experimental data reveals over 95% of play to be consistent with threshold strategies by supergame 30. Thus, persistent implementation error seems less of a concern in these treatments. D8 is the treatment where persistent implementation noise has the most effect. In this treatment, cooperation rates in the 1000th supergame are significantly affected by the constraint (although still remain below 30%) and our experimental data cannot inform us of the extent to which implementation errors may persist.

A.6. Sample Instructions: D8 Treatment

Welcome

You are about to participate in an experiment on decision-making. What you earn depends partly on your decisions, partly on the decisions of others, and partly on chance. Please turn off cell phones and similar devices now. Please do not talk or in any way try to communicate with other participants.

We will start with a brief instruction period in which you will be given a description of the main features of the experiment. If you have any questions during this period, raise your hand and your question will be answered so everyone can hear.

General Instructions

1. You will be asked to make decisions in several rounds. You will be randomly paired with another person in the room for a sequence of rounds. Each sequence of rounds is referred to as a *match*.

2. Each match will last for **8** rounds.

3. Once a match ends, you will be randomly paired with someone for a new match. You will not be able to identify who you've interacted with in previous or future matches.

Description of a Match

4. The choices and the payoffs in each round of a match are as follows:

	1	2
1	$51, 51$	$5, 87$
2	$87, 5$	$39, 39$

The first entry in each cell represents your payoff for that round, while the second entry represents the payoff of the person you are matched with.

(a) The table shows the payoffs associated with each combination of your choice and choice of the person you are paired with.

(b) That is, in each round of a match, if:

- $(1, 1)$: You select 1 and the other selects 1, you each make 51.
- $(1, 2)$: You select 1 and the other selects 2, you make 5 while the other makes 87.
- $(2, 1)$: You select 2 and the other selects 1, you make 87 while the other makes 5.
- $(2, 2)$: You select 2 and the other selects 2, you each make 39.

To make a choice, click on one of the rows on the table. Once a row is selected, it will change color and a red submit button will appear. Your choice will be finalized once you click on the submit button.

Once you and the person you are paired with have made your choices, those choices will be highlighted and your payoff for the round will appear.

End of the Session

5. The experiment will end after 30 matches have been played.

6. Total payoffs for each match will be the sum of payoffs obtained from each round of that match. Total payoffs for the experiment will be the sum of payoffs for all matches played. Your total payoffs will be converted to dollars at the rate of 0.003$ for every point earned.

Are there any questions?

Before we start, let me remind you that:

- Each match will last for 8 rounds. Payoffs in each round of a match, as given in the table above, depend on your choice and the choice of the person youre paired with.

- After a match is finished, you will be randomly paired with someone for a new match.