COOPERATION IN THE FINITELY REPEATED PRISONER'S DILEMMA*

Matthew Embrey
Guillaume R. Fréchette
Sevgi Yuksel

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 article 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 article establishes a systematic pattern of behavior: subjects converge to use threshold strategies that conditionally cooperate until a threshold round; 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. *JEL Codes:* C72, C73, C92.

*We 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.

The Quarterly Journal of Economics (2018), 509–551. doi:10.1093/qje/qjx033.
Advance Access publication on August 26, 2017.

I. Introduction

The prisoner's dilemma (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 article, we experimentally study the finitely repeated PD to understand the factors that affect (i) the emergence of cooperative behavior and (ii) 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; 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

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.

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 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 and analyze the data within a unified framework. Of the seven previous studies meeting our criteria, we could obtain the data for five of them.[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, as the horizon increases, cooperation rates increase, and this is attributed, in a loose sense, to the difficulty of reasoning backward through more rounds.[4]

2. For example, Selten and 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 and Miller (1993) find that behavior moves in the opposite direction; namely, they observe that the point of first defection increases with experience.

3. Although we reanalyze the original raw data, rather than collate the results of previous studies, we 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 and 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 and 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

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 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

to multiple papers using the level-k model to explain behavior in the centipede game, which, as we discuss in Section VII and Online Appendix A.1, is closely connected to the finitely repeated PD (see Kawagoe and Takizawa 2012; Ho and Su 2013; Garcia-Pola, Iriberri, and Kovarik 2016). In a recent paper, Alaoui and 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. The finding that cooperation correlates to the size of the basin of attraction in infinitely repeated PDs can be found in Dal Bó and Fréchette (2011). See also Blonski, Ockenfels, and Spagnolo (2011), who provide an axiomatic basis for the role of risk dominance in this context.

	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 I

The Prisoner's Dilemma

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 nonnegligible 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.

II. 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, Panel 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

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

(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, AD is the unique subgame-perfect equilibrium strategy of the finitely repeated game.

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, e.g., Mengel 2014a; 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 article is not to test these theories directly but 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 Lave (1965), Rapoport and Chammah (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,

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 and Sinervo (2016), implies that choosing to cooperate always improves efficiency.

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 cover all seven published papers (that we could find—five of which are included in our metadata) with experiments that include a treatment in which subjects play the finitely repeated PD and in which this is performed more than once.[10]

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 subjects defect earlier with experience. There is evidence

8. Mengel (2014b) presents a meta-study that covers more papers and also supplements the existing literature with new experiments. 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). Thus, the study 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 that 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 have at least one rematching of partners.

10. Online Appendix A.1 also discusses several recent papers (Schneider and Weber 2013; Cox et al. 2015; Kagel and McGee 2016; Kamei and Putterman 2017; Mao et al. 2017) that study heterogeneity in cooperative behavior and the role of reputation building in the finitely repeated PD. 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-player linear VCM is a PD with $g = \ell$. However, few experiments involve repetitions of finitely repeated linear VCMs (with rematching between each supergame); these are Andreoni and Petrie (2004, 2008), Muller et al. (2008), and Lugovskyy et al. (2017).

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).

III. 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.[11] 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 1 and the punishment payoff is 0. The resulting stage-game is shown in Figure 1, Panel B, where $g = \frac{(T-P)}{(R-P)} - 1 > 0$ is the one-shot gain from defecting, compared to the cooperative outcome, and $\ell = -\frac{(S-P)}{(R-P)} > 0$ is the one-shot loss from being defected on, compared to the noncooperative outcome.

III.A. 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 I reports these statistics for each treatment. They are sorted from shortest to longest horizon and from largest to smallest gain from defection.

The first observation that stands out from Table I 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

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

TABLE I

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.

gain.[12] In this sense, there seems to be some consistency across studies.

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 ($H \leqslant 4$).

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 II traces the evolution of these cooperation rates over supergames separated by horizon and payoffs. In most treatments, last-round cooperation rates are close to 0 or reach low levels quickly.[13] The evolution of round-one cooperation rates depends on the horizon. With $H = 2$, cooperation rates in round one are close to 0, 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.

12. 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-analysis 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 unmodelled correlations (see Fréchette 2012, for a discussion of session effects). Two alternative specifications are explored to gauge the robustness of the results in Online Appendix A.4. The different specifications do not change the main findings.

13. 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 and Seutens (2012) and Cabral, Ozbay, and Schotter (2014) use experimental designs to disentangle these two forces and find cooperation to be driven mainly by strategic motives.

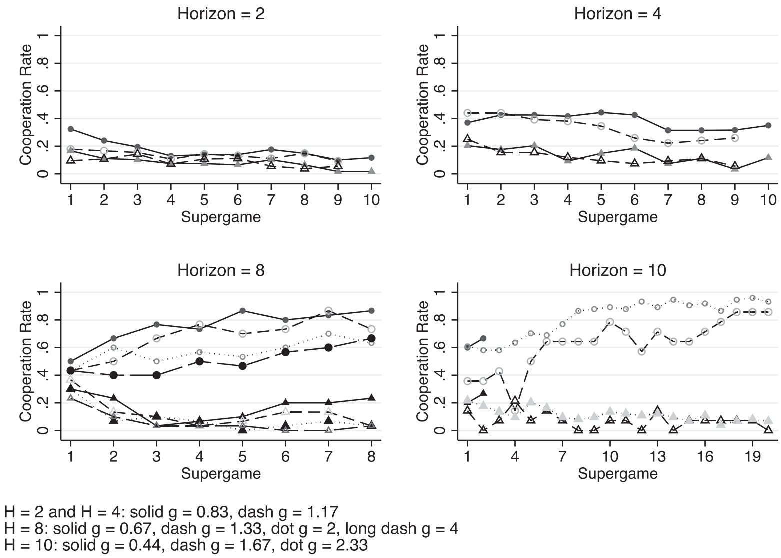


FIGURE II

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

III.B. 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 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 cooperation emerges. The decision to cooperate seems to be better predicted by the size of the basin of attraction of always defecting (*sizeBad*) against the grim trigger strategy (Dal Bó and Fréchette 2011).[14] 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:[15]

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

Values close to 1 suggest that the environment is not conducive to supporting (nonequilibrium) 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 0. As can be seen, *sizeBAD* is increasing in g and ℓ but decreasing in H.

Table II 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.[16] The first specification controls for the normalized stage-game parameters, g and ℓ, and H.[17] As can be seen, both g and H have a significant effect on round-one choices, and in the

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

15. 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: defecting in the last round would always achieve a higher payoff.

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

17. We report this specification, as it makes the effect of the regressors of interest easy to read. However, a more complete 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.

TABLE II

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	5,398	5,398

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.

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 nonlinear 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.[18]

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.

18. Another way to assess 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*.

	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 III

Stage-Games in the Experiment

IV. 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 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 III.[19] 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%.

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 III, the (g, ℓ) combination is (3, 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.

19. Payoffs are in experimental currency units (ECU) converted to dollars at the end of the experiment.

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, eight times. Combining the two treatment variables gives the four treatments: D4, D8, E4, and E8, where D and E refer to the stage-game, and 4 and 8 to the horizon.[20]

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 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.[21]

IV.A. Procedures

The experiments were conducted at NYU's Center for Experimental Social Science using undergraduate students from all majors, recruited via email.[22] 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

20. 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.

21. Other indexes to correlate with cooperation in the finitely repeated PD have been considered, but they only depend on stage-game payoffs (Murnighan and Roth 1983; Mengel 2014b).

22. Instructions were read aloud (see Online Appendix A.6). The computer interface was implemented using zTree (Fischbacher 2007).

TABLE III

Cooperation Rates: Early Supergames (1–15) versus 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	∼	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%.

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.[23] 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.[24] Throughout, a subject experienced just one set of treatment parameters: the stage-game payoffs and the supergame horizon.

IV.B. The Standard Perspective in the Experiment

Table III 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

23. 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.

24. More details about each session are provided in Online Appendix A.3.

TABLE IV

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%.

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.

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 IV 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 nonmonotonicity 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.

Figure IV 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

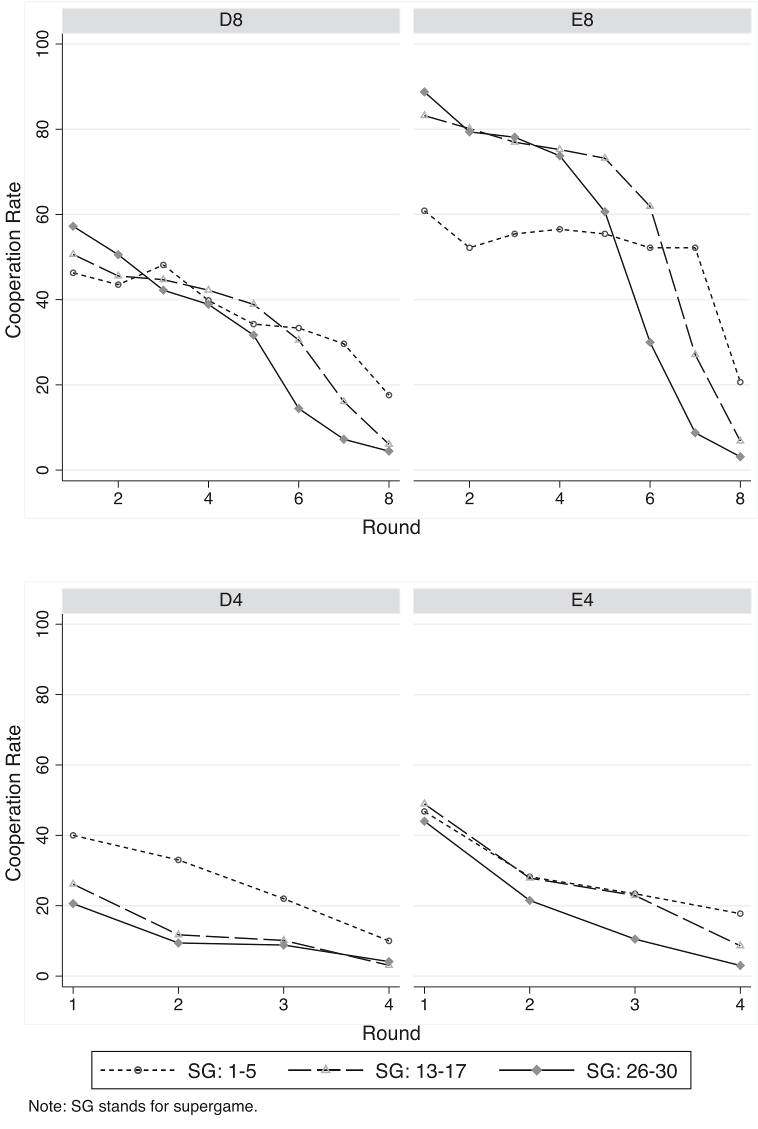


FIGURE IV

Cooperation Rate by Round Separated in Groups of Five Supergames

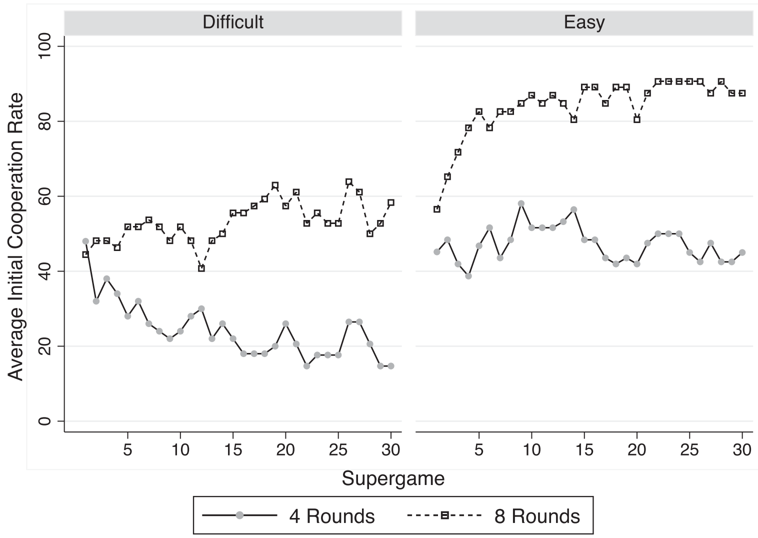


FIGURE V

Average Cooperation Rates in the First Round

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.[25]

IV.C. Determinants of Initial Cooperation

Figure V 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 shows decreasing initial cooperation

25. Online Appendix A.3 includes pairwise comparisons of the cooperation measures by treatment.

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 1 and 30 are statistically indistinguishable between the two treatments.[26]

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.

V. 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 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.

V.A. Behavior in the E8 treatment

Figure VI 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.[27] In 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,

26. 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% levels, respectively) and D8 is higher in supergames 18, 19, and 21 (at the 10%, 5%, and 5% levels, 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 Appendixes A.3 and A.4 for details).

27. Online Appendix A.3 replicates Figure VI, including all rounds.

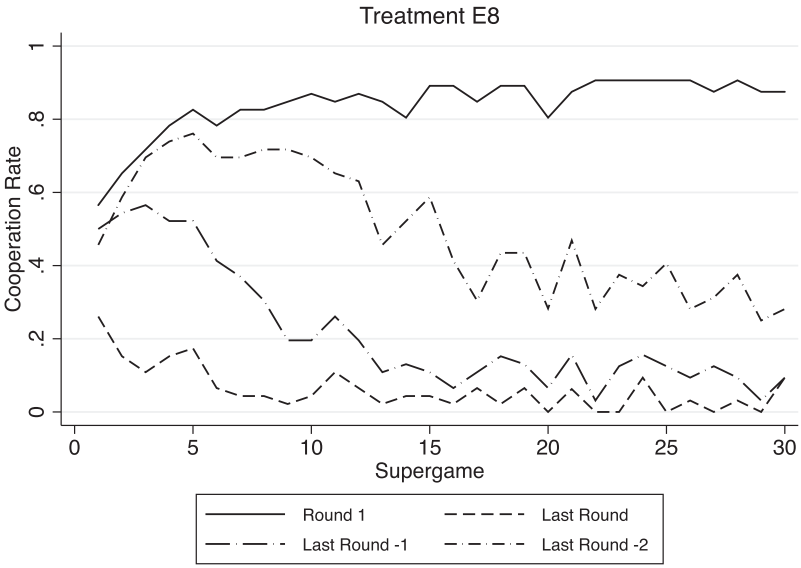


FIGURE VI

Mean Cooperation Rate by Round

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 toward 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, nonmonotonicity 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 30th supergame, cooperation is not decreasing in the first 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 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.[28]

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.[29] These different statistics are presented in the same graph in the left panel of Figure VII. 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

28. 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 $m = 1$).

29. 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.

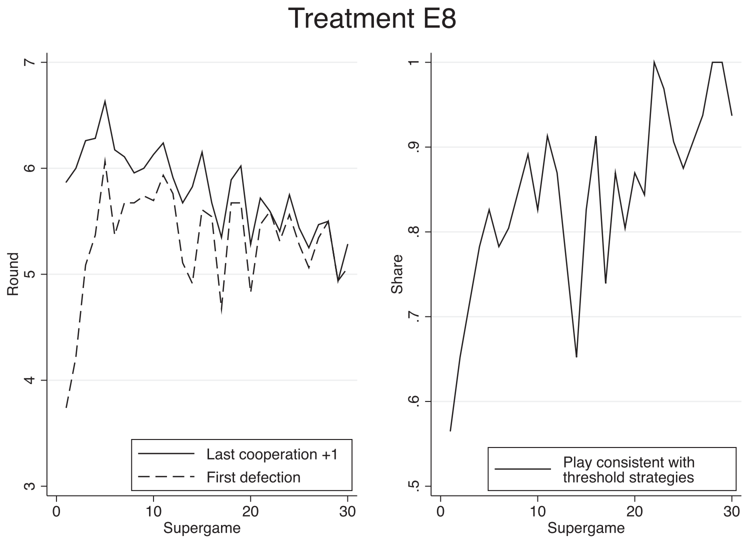


FIGURE VII

Evolution of Threshold Strategies

while mean round to first defection is increasing (at least in the early parts of a session).

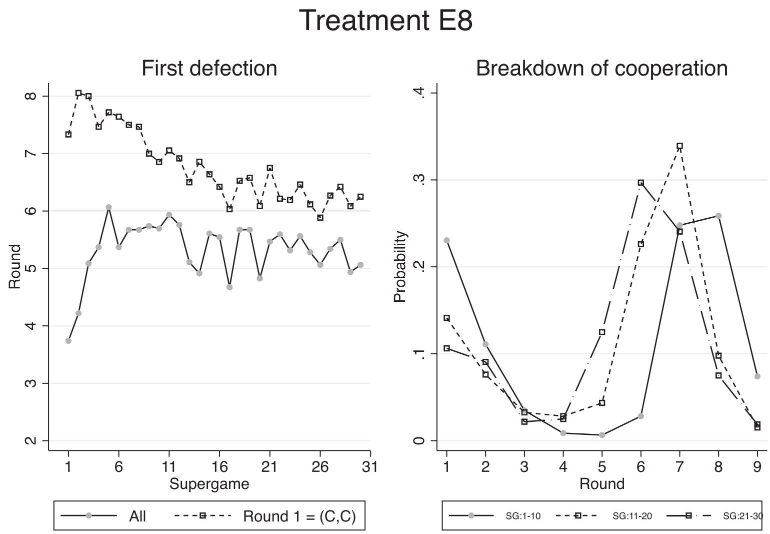
If, instead of mean round to first defection, one considers mean round to last cooperation, then it appears as if subjects move in the direction suggested by backward induction in all treatments, including E8. The gap between the two lines suggests that the use of threshold strategies becomes dominant over the course of the experiment. This suggestion is confirmed in the right panel of Figure VII, 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 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.[30]

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 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 VIII 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 one. As expected, the line conditional on round-one cooperation is higher, but the gap between the two lines

30. Burton-Chellew, El Mouden, and West (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.



Note: SG stands for supergame.

FIGURE VIII

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

shrinks as round-one cooperation rates increase over time. Most important, conditional on achieving cooperation in the first round, mean round to first defection actually decreases over time. 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 10 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 toward 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 10 supergames.

V.B. Breakdown of Cooperation in Other Treatments

Figure IX illustrates, for the three other treatments, the evolution of cooperation for the first and last three rounds. D8 has a

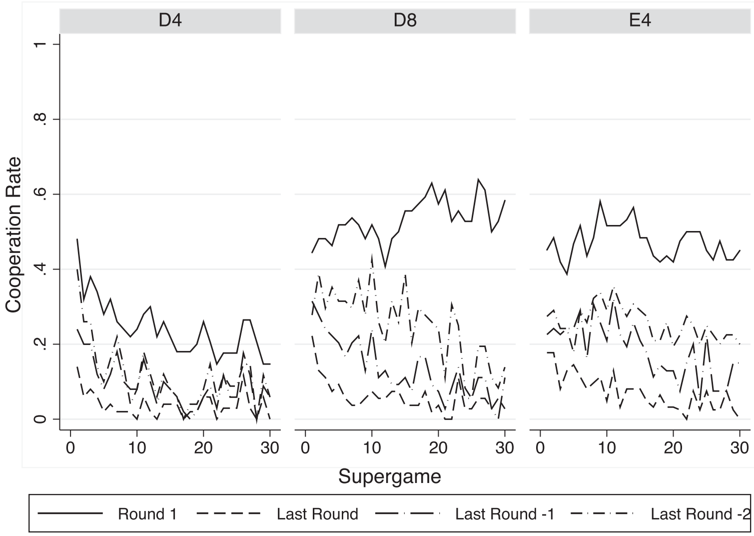


FIGURE IX

Cooperation Rates in Selected Rounds across Supergames

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

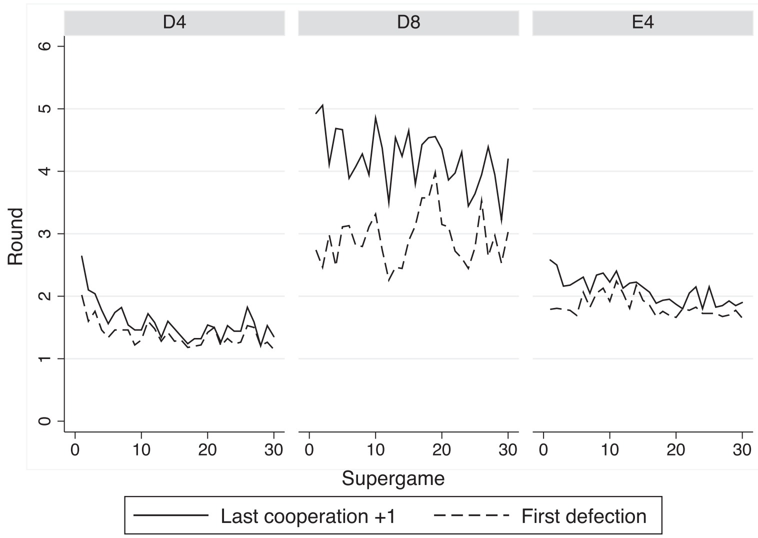


FIGURE X

Evolution of First Defection versus Last Cooperation across Supergames

decreasing 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 X 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.[31] These suggest that

31. Indeed, H is negatively correlated with play consistent with a threshold strategy in the first supergame. This does not reach statistical significance if only

convergence to threshold strategies would happen faster in shorter horizon games.

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 VII and VIII 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 compare the beginning of an experiment to the end (see Table V). In the first supergame of all studies with $H \geqslant 8$, 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 nonnegligible amount of experimentation or confusion at the beginning of a session, followed by a universal convergence to using threshold strategies.[32]

VI. 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

considering our experiment (p = .11) but it is significant at the 1% level when considering the entire meta-data.

32. Statistical significance is established in the regressions reported in Online Appendix A.2.

TABLE V
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

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

for many more supergames than can be observed during a typical lab session. This can help us gain 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.

VI.A. Model

The general structure of the learning model we adopt is motivated by the following observations documented in the previous sections: (i) cooperation rates in the first round of a supergame are decreasing in the size of the basin of attraction of AD; (ii) choices respond to experiences with other players in previous supergames; and (iii) 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 ii 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 i above). We specify a simple belief-based learning model that can capture these key features.[33]

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).[34] Note that, contrary to the threshold strategies, TFT and STFT allow for cooperation to reemerge after a period of defection within a supergame. We have included all strategies for which there is evidence of systematic use in the data.[35]

Beliefs evolve over time, given a subject's experience within a supergame. After every supergame, a subject updates her beliefs as follows:

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

33. 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 and Fréchette (2011). The model is in the style of Cheung and 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 and Erev (1995), Cooper, Garvin, and Kagel (1997), and Camerer and Ho (1999).

34. 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.

35. 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 data set defected at least 34 times throughout the session, and at least 15 times in the last 10 supergames.

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.[36] 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.[37]

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:

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

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

36. 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 priors: with θ_i it captures the importance of new experiences.

37. 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 toward the end of the sessions in our data.

the strategies adopted by others in the population and, consequently, describes learning about the optimality of different strategies.

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.

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 $0 \leqslant \sigma_{it} \leqslant 0.5$. 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.[38]

In summary, for each subject, we estimate β_{i0}, λ_i, σ_i, which describe initial beliefs, noise in strategy, and action choice implementation, and θ_i, κ_i, which describe how beliefs are updated with experience and how execution noise changes over time.[39] The estimates are obtained via maximum likelihood estimation for each

38. 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.

39. 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.

subject separately.[40] 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.

VI.B. 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 simulations with more repeated games to understand how cooperation would evolve in the long run.[41] These simulations consist of 100,000 sessions by treatment.[42]

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 XI, which compares the average simulated cooperation rate for each round of the repeated game with the experimental results.[43] 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 five and six, there is nonmonotonicity in cooperation rates, as they first

40. An alternative would be to pool the data. However, for the purpose of this article 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.

41. 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.

42. 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.

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

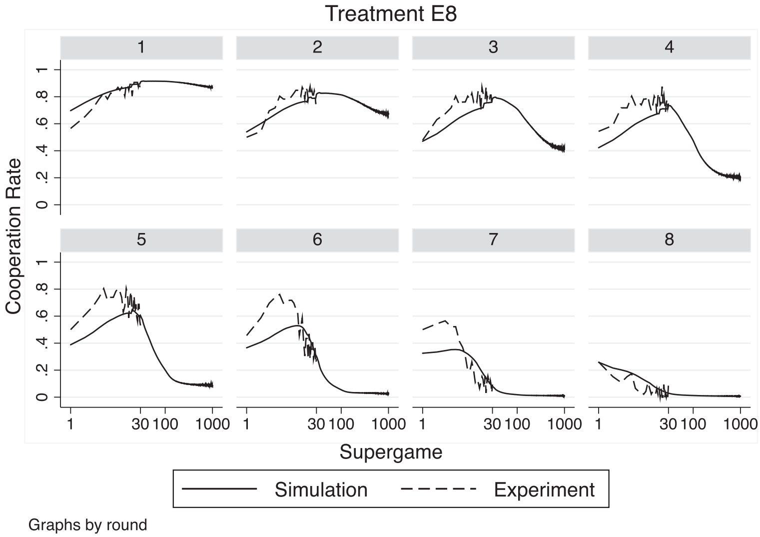


FIGURE XI

Average Cooperation: Simulation versus Experimental Data for Each Round in E8

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 XI 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 1,000 supergames. However, this is clearly limited to early rounds. More important, cooperation rates, if they are still positive, continue declining in all rounds, even after 1,000 supergames.[44] The evolution suggests that there is unraveling of cooperation in all rounds, but that it is

44. 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 1,000th supergame.

so slow that cooperation rates for the first round of a supergame can remain above 80% even after significant experience.[45] 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.

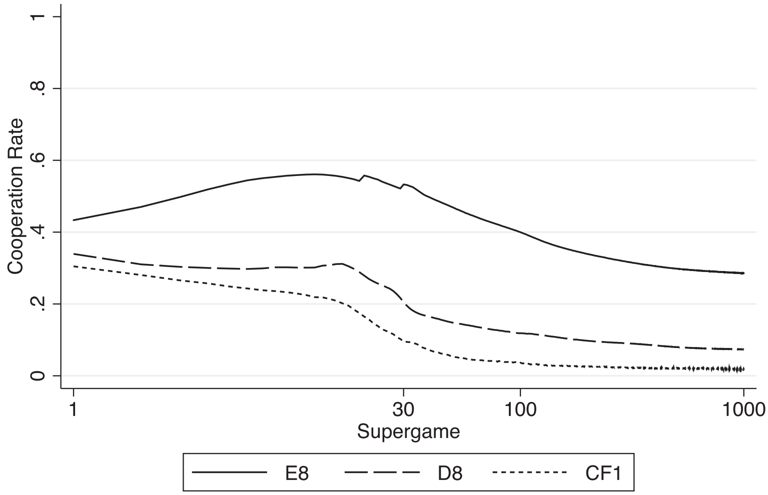
VI.C. Counterfactuals

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 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 XII, 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

45. 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 nonzero cooperation levels with further experience.



E8 corresponds to simulations with E8 learning estimates and stage game parameters.
D8 corresponds to simulations with D8 learning estimates and stage game parameters.
CF1 corresponds to simulations with E8 learning estimates and D8 stage game parameters.

FIGURE XII

Long-Term Evolution of Aggregate Cooperation

demonstrates the impact of the stage-game 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 fewer than 50 supergames in CF1.[46]

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. 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

46. 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 highlights that this behavior is driven by stage-game parameters rather than treatment-specific learning dynamics.

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. 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.

VII. 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 article, to understand how cooperative behavior and its evolution with experience vary with the environment in this canonical game, we reanalyze 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 article 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 AD strategy. This value-of-cooperation result relates to recent studies on continuous-time PD games (Friedman and Oprea 2012; Bigoni et al. 2015; Calford and Oprea 2017). Friedman and 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 AD as an evolutionary stable strategy in a finitely repeated prisoner's dilemma.[47] 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 (e.g., see Fudenberg and Imhof 2008; Imhof, Fudenberg, and Nowak 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 article 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 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 (e.g., theoretically in Radner 1986 and recently empirically investigated in Friedman and 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

47. It is linked to the size of the invasion (share of the population following the alternative strategy) needed to take over AD.

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. However, 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.[48] 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 and Oprea (2012). Differences in cooperative behavior across our treatments appear to be driven primarily by the corresponding differences in the tradeoff 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 toward 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.

48. 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.

University of Sussex
New York University
University of California, Santa Barbara

Supplementary Material

An Online Appendix for this article can be found at *The Quarterly Journal of Economics* online. Data and code replicating the tables and figures in this article can be found in Embrey, Fréchette, and Yuksel (2017), in the Harvard Dataverse, doi:10.7910/DVN/WCHA2J.

References

Alaoui, Larbi, and Antonio Penta, "Endogenous Depth of Reasoning," *Review of Economic Studies*, 83 (2016), 297–1333.

Andreoni, James, and John H. Miller, "Rational Cooperation in the Finitely Repeated Prisoner's Dilemma: Experimental Evidence," *The Economic Journal*, 103 (1993), 570–585.

Andreoni, James, and Ragan Petrie, "Public Goods Experiments without Confidentiality: a Glimpse into Fund-Raising," *Journal of Public Economics*, 88 (2004), 1605–1623.

———, "Beauty, Gender and Stereotypes: Evidence from Laboratory Experiments," *Journal of Economic Psychology*, 29 (2008), 73–93.

Bereby-Meyer, Yoella, and Alvin E. Roth, "The Speed of Learning in Noisy Games: Partial Reinforcement and the Sustainability of Cooperation," *American Economic Review*, 96 (2006), 1029–1042.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan, "How Much Should We Trust Differences-in-Differences Estimates?," *Quarterly Journal of Economics*, 119 (2014), 249–275.

Bigoni, Maria, Marco Casari, Andrzej Skrzypacz, and Giancarlo Spagnolo, "Time Horizon and Cooperation in Continuous Time," *Econometrica*, 83 (2015), 587–616.

Blonski, Matthias, Peter Ockenfels, and Giancarlo Spagnolo, "Equilibrium Selection in the Repeated Prisoner's Dilemma: Axiomatic Approach and Experimental Evidence," *American Economic Journal: Microeconomics*, 3 (2011), 164–192.

Bornstein, G., T. Kugler, and A. Ziegelmeyer, "Individual and Group Decisions in the Centipede Game: Are Groups More 'Rational' Players?," *Journal of Experimental Social Psychology*, 40 (2004), 599–605.

Burton-Chellew, Maxwell N., Claire El. Mouden, and Stuart A. West, "Conditional Cooperation and Confusion in Public-Goods Experiments," *Proceedings of the National Academy of Sciences*, 113 (2016), 1291–1296.

Cabral, Luis M. B., Erkut Ozbay, and Andrew Schotter. "Revisiting Strategic versus Non-Strategic Cooperation," *Games and Economic Behavior*, 87 (2014), 100–121.

Calford, Evan, and Ryan Oprea, "Continuity, Inertia and Strategic Uncertainty: A Test of the Theory of Continuous Time Games," *Econometrica*, 85 (2017), 915–935.

Camerer, Colin F., and Teck-Hua Ho, "Experience-Weighted Attraction Learning in Normal Form Games," *Econometrica*, 67 (1999), 827–874.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller, "Bootstrap-Based Improvements for Inference with Clustered Errors," *Review of Economics and Statistics*, 90 (2008), 414–427.

Cameron, A. Colin, and Douglas L. Miller, "A Practitioner's Guide to Cluster-Robust Inference," *Journal of Human Resources*, 50 (2015), 317–372.
Carter, Andrew V., Kevin T. Schnepel, and Douglas G. Steigerwald, "Asymptotic Behavior of a t-Test Robust to Cluster Heterogeneity," *Review of Economics and Statistics*, 99 (2017), 698–709
Cheung, Yin-Wong, and Daniel Friedman, "Individual Learning in Normal Form Games: Some Laboratory Results," *Games and Economic Behavior*, 19 (1997), 46–76.
Cooper, Russell, Douglas V. DeJong, Robert Fosythe, and Thomas W. Ross, "Cooperation without Reputation: Experimental Evidence from Prisoner's Dilemma Games," *Games and Economic Behavior*, 12 (1996), 187–218.
Cooper, David J., Susan Garvin, and John H. Kagel, "Signalling and Adaptive Learning in an Entry Limit Pricing Game," *RAND Journal of Economics*, 28 (1997), 662–683.
Cox, Caleb A., Matthew T. Jones, Kevin E. Pflum, and Paul J. Healy, "Revealed Reputations in the Finitely Repeated Prisoners Dilemma," *Economic Theory*, 58 (2015), 441–484.
Crawford, Vincent P., "Adaptive Dynamics in Coordination Games," *Econometrica*, 63 (1995), 103–43.
Dal Bó, Pedro, "Cooperation under the Shadow of the Future: Experimental Evidence from Infinitely Repeated Games," *American Economic Review*, 95 (2005), 1591–1604.
Dal Bó, Pedro, and Guillaume R. Fréchette, "The Evolution of Cooperation in Infinitely Repeated Games: Experimental Evidence," *American Economic Review*, 101 (2011), 411–429.
Embrey, Matthew, Guillaume R. Fréchette, and S. Yuksel, "Replication Data for: 'Cooperation in the Finitely Repeated Prisoner's Dilemma,'" *Harvard Dataverse* (2017), doi:10.7910/DVN/WCHA2J.
Fey, Mark, Richard D. McKelvey, and Thomas R. Palfrey, "An Experimental Study of Constant-Sum Centipede Games," *International Journal of Game Theory*, 25 (1996), 269–287.
Fischbacher, Urs, "z-tree: Zurich Toolbox for Ready-made Economic Experiments," *Experimental Economics*, 10 (2007), 171–178.
Flood, Merrill M., "Some Experimental Games," Research Memorandum RM-789, Rand Corporation, Santa Monica, (1952).
Fréchette, Guillaume R., "Learning in a Multilateral Bargaining Experiments," *Journal of Econometrics*, 153 (2009), 183–195.
———, "Session-Effects in the Laboratory," *Experimental Economics*, 15 (2012), 485–498.
Friedman, Daniel, and Ryan Oprea, "A Continuous Dilemma," *American Economic Review*, 102 (2012).
Friedman, Daniel, and Barry Sinervo, *Evolutionary Games in Natural, Social, and Virtual Worlds* (Oxford: Oxford University Press, 2016).
Fudenberg, Drew, and David K. Levine, *The Theory of Learning in Games* (Cambridge, MA: MIT Press, 1998).
Fudenberg, Drew, and Lorens A. Imhof, "Monotone Imitation Dynamics in Large Populations," *Journal of Economic Theory*, 140 (2008), 229–245.
Garcia-Pola, Bernardo, Nagore Iriberri, and Jaromir Kovarik, "Non-Equilibrium Play in Centipede Games," CEPR Discussion Paper No. DP11477, 2016.
Hauk, Esther, and Rosemarie Nagel, "Choice of Partners in Multiple Two-Person Prisoner's Dilemma Games: An Experimental Study," *Journal of Conflict Resolution*, 45 (2001), 770–793.
Ho, Teck-Hua, and Xuanming Su, "A Dynamic Level-k Model in Sequential Games," *Management Science*, 59 (2013), 452–469.
Imbens, Guido W., and Michal Kolesar, "Robust Standard Errors in Small Samples: Some Practical Advice," *Review of Economics and Statistics*, 98 (2016), 701–712.

Imhof, Lorens A., Drew Fudenberg, and Martin A. Nowak, “Evolutionary Cycles of Cooperation and Defection,” *Proceedings of the National Academy of Sciences of the United States of America*, 102 (2005), 10797–10800.
Kagel, John H., and Peter McGee, “Team versus Individual Play in Finitely Repeated Prisoner Dilemma Games,” *American Economics Journal: Microeconomics*, 8 (2016), 253–276.
Kahn, Lawrence M., and J. Keith Murnighan, “A General Experiment on Bargaining in Demand Games with Outside Options,” *American Economic Review*, 88 (1993), 1260–1280.
Kamei, Kenju, and Louis Putterman, “Play It Again: Partner Choice, Reputation Building and Learning in Restarting, Finitely Repeated Dilemma Games,” *The Economic Journal*, 127 (2017), 1069–1095.
Kawagoe, Toshiji, and Hirokazu Takizawa, “Level-k Analysis of Experimental Centipede Games,” *Journal of Economic Behavior and Organization*, 82 (2012), 548–566.
Kline, Patrick, and Andres Santos, “A Score Based Approach to Wild Bootstrap Inference,” *Journal of Econometric Methods*, 1 (2012), 23–41.
Kreps, David M., Paul Milgrom, John Roberts, and Robert Wilson, “Rational Cooperation in the Finitely Repeated Prisoners’ Dilemma,” *Journal of Economic Theory*, 27 (1982), 245–252.
Lave, Lester B., “Factors Affecting Cooperation in the Prisoner’s Dilemma, ” *Behavioral Science*, 10 (1965), 26–38.
Levitt, Steven D., John A. List, and Sally E. Sadoff, “Checkmate: Exploring Backward Induction among Chess Players,” *American Economic Review*, 101 (2011), 975–990.
Lugovskyy, Volodymyr, Daniela Puzzello, Andrea Sorensen, James Walker, and Arlington Williams, “An Experimental Study of Finitely and Infinitely Repeated Linear Public Goods Games,” *Games and Economic Behavior*, 102 (2017), 286–302
MacKinnon, James G., and Matthew D. Webb, “Wild Bootstrap Inference for Wildly Different Cluster Sizes,” *Journal of Applied Econometrics*, 32 (2017), 233–254.
Mantovani, Marco, “Limited Foresight in Sequential Games: An Experiment,” Working paper, University of Milan – Bicocca, 2016.
Mao, Andrew, Lili Dworkin, Siddharth Suri, and Duncan J. Watts, “Resilient Cooperators Stabilize Long-Run Cooperation in the Finitely Repeated Prisoner’s Dilemma,” *Nature Communications*, 8 (2017), 13800
McKelvey, R. D., and T. R. Palfrey, “An Experimental Study of the Centipede Game,” *Econometrica*, 60 (1992), 803–836.
Mengel, Friederike, “Learning by (Limited) Forward Looking Players,” *Journal of Economic Behavior and Organization*, 108 (2014a), 59–77.
———, “Risk and Temptation: A Meta-study on Prisoner’s Dilemma Games,” *The Economic Journal*, forthcoming, doi:10.1111/ecoj.12548.
Moinas, Sophie, and Sebastien Pouget, “The Bubble Game: An Experimental Study of Speculation,” *Econometrica*, 81 (2013), 1507–1539.
Morehous, L. G., “One-Play, Two-Play, Five-Play, and Ten-Play Runs of Prisoner’s Dilemma,” *Journal of Conflict Resolution*, 10 (1966), 354–362.
Muller, Laurent, Martin Sefton, Richard Steinberg, and Lise Vesterlund, “Strategic Behavior and Learning in Repeated Voluntary Contribution Experiments,” *Journal of Economic Behavior and Organization*, 67 (2008), 782–793.
Murnighan, J. Keith, and Alvin E. Roth, “Expecting Continued Play in Prisoner’s Dilemma Games,” *Journal of Conflict Resolution*, 27 (1983), 279–300.
Nagel, Rosemarie, and Fang Fang Tang, “Experimental Results on the Centipede Game in Normal Form: An Investigation on Learning,” *Journal of Mathematical Psychology*, 42 (1998), 356–384.
Normann, Hans-Theo, and Brian Wallace, “The Impact of the Termination Rule on Cooperation in a Prisoner’s Dilemma Experiment,” *International Journal of Game Theory*, 41 (2012), 707–718.

Palacios-Huerta, Ignacio, and Oscar Volij, "Field Centipedes," *American Economic Review*, 99 (2009), 1619–1635.
Radner, Roy, "Can Bounded Rationality Resolve at the Prisoner's Dilemma?," in *Contributions to Mathematical Economics*, A. Mas-Colell and W. Hildenbrand, ed. (Amsterdam: North-Holland, 1986), 387–399.
Rapoport, Anatol, and Albert M. Chammah, "Sex Differences in Factors Contributing to the Level of Cooperation in the Prisoner's Dilemma Game," *Journal of Personality and Social Psychology*, 2 (1965), 831.
Rapoport, Amnon, William E. Stein, James E. Parco, and Thomas E. Nicholas, "Equilibrium Play and Adaptive Learning in a Three-Person Centipede Game," *Games and Economic Behavior*, 43 (2005), 239–265.
Reuben, Ernesto, and Sigrid Seutens, "Intrinsic and Instrumental Reciprocity: An Experimental Study," *Experimental Economics*, 15 (2012), 24–43.
Roth, Alvin E., "Laboratory Experimentation in Economics: A Methodological Overview," *The Economic Journal*, 98 (1988), 974–1031.
Roth, Alvin E., and Ido Erev, "Learning in Extensive-Form Games: Experimental Data and Simple Dynamic Models in the Intermediate Term," *Games and Economic Behavior*, 8 (1995), 164–212.
Schneider, Frederic, and Roberto A. Weber, "Long-Term Commitment and Cooperation," Working paper, University of Zurich Department of Economics Working Paper No. 130, 2013.
Selten, Reinhard, and Rolf Stoecker, "End Behavior in Sequences of Finite Prisoner's Dilemma Supergames: A Learning Theory Approach," *Journal of Economic Behavior and Organization*, 7 (1986), 47–70.
Webb, Matthew D., "Reworking Wild Bootstrap Based Inference for Clustered Errors," Working paper, Queen's Economics Department Working Paper, 2014.
Zauner, Klaus G., "A Payoff Uncertainty Explanation of Results in Experimental Centipede Games," *Games and Economic Behavior*, 26 (1999), 157–185.