Anna Dreber, Drew Fudenberg, and David G. Rand. 2012. Slow to anger and fast to forgive: cooperation in an uncertain world. *American Economic Review* 102(2): 720-749. DOI: 10.1257/aer.102.2.720

Slow to Anger and Fast to Forgive:

Cooperation in an Uncertain World

Drew Fudenberg, David G. Rand, and Anna Dreber*

Forthcoming in American Economic Review

Abstract: We study the experimental play of the repeated prisoner's dilemma when intended actions are implemented with noise. In treatments where cooperation is an equilibrium, subjects cooperate substantially more than in treatments without cooperative equilibria. In all settings there was considerable strategic diversity, indicating that subjects had not fully learned the distribution of play. Furthermore, cooperative strategies yielded higher payoffs than uncooperative strategies in the treatments with cooperative equilibria. In these treatments successful strategies were "lenient" in not retaliating for the first defection, and many were "forgiving" in trying to return to cooperation after inflicting a punishment.

Repeated games with observed actions have a great many equilibrium outcomes when players are patient, as shown by various folk theorems.[1] These theorems show that cooperative play is *possible* when players are concerned about future rewards and punishments, but since repeated play of a static equilibrium is always an equilibrium of the repeated game, the folk theorems do no predict that cooperation will in fact occur. Intuition and evidence (e.g. Robert Axelrod 1984; Pedro Dal Bó 2005; Anna Dreber et al. 2008; Dal Bó and Guillaume Frechette 2009; John Duffy and Jack Ochs 2009) suggest that in repeated games with observed actions

* Fudenberg: Department of Economics, Harvard University, 1805 Cambridge Street, Cambridge MA 02138, USA (email: dfudenberg@harvard.edu); Rand: Program for Evolutionary Dynamics, Harvard University, 1 Brattle Square Suite 6, Cambridge MA 02138, USA (email: drand@fas.harvard.edu); Dreber: SIFR, Drottninggatan 89, 11360 Stockholm, Sweden (email: anna.dreber@sifr.org). Fudenberg and Rand are co-lead authors. We are grateful to Pedro Dal Bó and Guillaume Frechette for sharing their data and code with us, and for taking the time to reply to our numerous queries. We thank Rob Boyd, Armin Falk, Simon Gächter, Edward Glaeser, Stephanie Hurder, Guido Imbens, Magnus Johannesson, Martin Nowak, Parag Pathak, Thomas Pfeiffer and John Scholz for helpful conversations and comments, and three referees for useful reports. Research support was provided by National Science Foundation grants SES-064816 and SES- 0951462, the Dean of the Faculty of Arts and Sciences, and David Rand is supported by a grant from the John Templeton Foundation.

[1] James W. Friedman (1971), Robert J. Aumann and Lloyd S. Shapley (1994) and Drew Fudenberg and Eric Maskin (1986).

players do indeed tend to cooperate when there is a cooperative equilibrium, at least if the gains to cooperation are sufficiently large.[2]

Outside of the laboratory, actions are often observed with noise: someone who claims they worked hard, or that they were too busy or sick to help, may or may not be telling the truth, and an awkward or inconvenient action may have been well-intentioned; similarly, a self-interested action may wind up accidentally benefiting another. In this setting too, there can be equilibria in which players do better than in the one-shot play of the game, as seen for example in the trigger-strategy equilibria constructed by Edward J. Green and Robert H. Porter (1984), and indeed a version of the folk theorem applies to repeated games with imperfect public monitoring provided, as shown by Fudenberg, David K. Levine and Maskin (1994).[3] Moreover, while there are evolutionary arguments for cooperation in repeated games with perfectly observed actions, the evolutionary arguments for cooperative equilibria are even stronger with imperfect observations, as the possibility that punishment may be triggered by "mistake" decreases the viability of unrelenting or grim strategies that respond to a single bad observation by never cooperating again.[4]

This paper studies experimental play of the repeated prisoner's dilemma when intended actions are implemented with noise. Our main goals are to understand whether and when subjects play cooperatively, and also to get a sense of the sorts of strategies that they use. We present evidence from four different payoff specifications for a repeated prisoner's dilemma, with stage game actions "Cooperate" ("C") and "Defect" ("D") (neutral language was used in the experiment itself); the difference in specifications was the benefit that playing "C" gives to the other player. We consider these four payoff specifications with a continuation probability of 7/8

[2] Dal Bó (2005) and Dal Bó and Frechette (2011) find that there need not be cooperation when the gain from cooperation is small. Earlier papers had found less cooperative play than these studies, but as Dal Bó (2005) discusses these papers had various methodological flaws, such as subjects playing vs. an experimenter instead of each other, or extremely low payments.

[3] Public monitoring means that all players observe a public signal whose distribution depends on the actions played, as in the work of Dilip Abreu, David Pearce and Ennio Stachetti (1990). The noisy games we study have public monitoring, and satisfy the "identifiability" conditions that are needed for the folk theorem.

[4] Axelrod and William D. Hamilton (1981) applied the ESS solution concept to a small subset of the repeated game strategies. Fudenberg and Maskin (1990, 1994) and Ken Binmore and Larry Samuelson (1992) and show that ESS predicts cooperation in the prisoner's dilemma with noisy observations even when all finite-complexity strategies are allowed. Robert Boyd (1989) notes that noise permits cooperative strategies to be evolutionarily stable, but in his framework non-cooperative strategies are stable as well. A parallel literature considers explicit evolutionary dynamics on a restricted strategy space; this includes Marcus W. Feldman and Ewart A.C. Thomas (1987), Martin A. Nowak and Karl Sigmund (1990), and Nowak et al. (2004) for the case of observed actions, and Nowak and Sigmund (1993) and Lorens A. Imhof, Fudenberg and Nowak (2007) for repeated games with noise.

and an error rate of 1/8. As controls we also consider error rates of 1/16 and 0 (no exogenously imposed error at all) under the "most cooperative" payoff specification, i.e. that with highest benefit to cost ratio. We find that there is much more cooperation in specifications where there is a "cooperative equilibrium" (an equilibrium in which players initially cooperate, and continue to do so at least until a D is observed): players cooperate 49-61% of the time in treatments with cooperative equilibria, compared to 32% in the specification without cooperative equilibria. In these specifications, we also find that cooperative strategies yielded higher payoffs than uncooperative ones. Conversely, in the one treatment where "Always Defect" is the only equilibrium, this strategy was the most prevalent and had the highest payoff.

As compared to experiments on the prisoner's dilemma without noise, which we review in Section II, subjects were markedly more lenient (slower to resort to punishment). For example, in the payoff treatments that support cooperation, we find that subjects played C in 63-77% of the rounds immediately following their partner's first defection, compared to only 13-42% in the cooperative treatments of Dal Bó and Frechette (2011), Dreber et al. (2008) and our no-error control, which are the no-error games we will be using for comparison throughout.[5] Subjects also showed a considerable level of forgiveness (willingness to give cooperation a second chance): in the noisy specifications with cooperative equilibria, our subjects returned to playing C in 18-33% of rounds immediately following the breakdown of cooperation, as compared to 6-28% in the games without error.[6] Consistent with these findings, subjects are more likely to condition on play more than one round ago in the noisy treatments than in the no-error games; 65% use strategies with longer memories in presence of error, compared to 28% in the games without error.

In addition to such descriptive statistics, we more explicitly explore which strategies our subjects used, using the techniques of Dal Bó and Frechette (2011).[7] Relatively few subjects used

[5] These statistics describe cooperation after one's partner defects for the first time, rather than overall cooperation. We find similar levels of overall cooperation without noise (78%) and with low noise of 1/16 (82%), but substantially less overall cooperation with high noise 1/18 (59%).

[6] This describes histories in which (i) at least one subject chose C in the first round, (ii) in at least one previous round, the initially cooperative subject chose C while the other subject chose D and (iii) in the immediately previous round the formerly cooperative subject played D.

[7] In section III we summarize the findings of Claus Wedekind and Manfred Milinski (1996) and Masaki Aoyagi and Frechette (2009), who also analyzed the strategies used in a repeated prisoner's dilemma, and of Jim Engle-Warnick and Robert L. Slonim (2006), who examined the strategies used in a trust game. With the exception of Aoyagi and Frechette, these experiments considered environments without noise; the introduction of noise leads more information sets to be reached and so makes it easier to distinguish between strategies.

the strategy "Tit-for-Tat" (TFT) which was prevalent in the Dal Bó and Frechette (2011) experiment, or the strategy "Perfect Tit-for-Tat" (PTFT, also called "Win-Stay, Lose-Shift") which is favored by evolutionary game theory in treatments where it is an equilibrium *if* players are restricted to strategies that depend only on the previous round's outcome, as is commonly assumed in that literature.[8] Instead, the most prevalent cooperative strategies were "Tit-for-2-Tats" (TF2T, punish once after two defections), "2-Tits-for-2-Tats" (2TF2T, punish two times after two defections), both of which are forgiving, and modified, lenient versions of the Grim strategy, which wait for two or three defections before abandoning cooperation. These results, and other descriptive measures of subjects' play, show that subjects can and do use strategies that look back more than one round, at least in games with noise.

We find considerable strategic diversity in all settings: No strategy has probability of greater than 30% in any treatment, and typically three or four strategies seem prevalent. Moreover, in every treatment a substantial number of players seem to use ALLD ("Always Defect"). That strategy does quite poorly in treatments where most of the subjects are playing strategies that are conditionally cooperative, but it is a best response to the belief that most other subjects play ALLD. Similarly, a substantial fraction of agents cooperate in the treatment where ALLD earns the highest payoff. We take this as evidence that learning was incomplete, and that it is difficult to learn the optimal response to the prevailing distribution of play.[9]

An alternative explanation for the diversity of play is that it reflects a distribution of social preferences, with some subjects preferring to cooperate even if it does not maximize their own payoffs, and others playing to maximize the difference between their partner's payoff and their own. To test this alternative hypothesis, we had subjects play a dictator game at the end of the session, with payoffs going to recipients recruited at a later experimental session, and we also asked subjects to fill out a post-experiment survey on attitudes and motivations. In another paper (Dreber, Fudenberg and David G. Rand 2010) we explore this possibility; our main conclusion is that when the selfish payoffs strongly support cooperation, social preferences do not seem to be a

[8] The explicit analysis of evolutionary dynamics becomes quite difficult when longer memories are possible.
[9] We used a continuation probability of 7/8, instead of the ½ and ¾ in Dal Bó and Frechette (2011), to investigate the extent to which players condition on observations before the previous round. When the continuation probability is ½, many interactions will last three or fewer rounds, which makes it hard to study how far back players look in choosing their actions.

key factor in explaining who cooperates and what strategies they use. Leniency and forgiveness seem to be motivated by strategic concerns rather than social preferences.

I. Experimental Design

The purpose of the experimental design is to test what happens when subjects play an infinitely repeated prisoner's dilemma with error. The infinitely repeated game is induced by having a known constant probability that the interaction will continue between two players following each round. We let the continuation probability be $\delta=7/8$. With probability $1-\delta$, the interaction ends and subjects are informed that they have been re-matched with a new partner. There is also a known constant error probability that an intended move is changed to the opposite move. Our main conditions use E=1/8; we also ran control conditions with E=1/16 and E=0. Subjects were informed when their own move has been changed (i.e. when they make an error), but not when the other player's move has been changed; they were only notified of the other player's actual move, not the other's intended move. Subjects were informed of all of the above in the experimental instructions, which are included in the online appendix.

The stage game is the prisoner's dilemma in Figure 1 where the payoffs are denoted in points. Cooperation and defection take the "benefit/cost" (b/c) form, where cooperation means paying a cost c to give a benefit b to the other player, while defection gives 0 to each party; b/c took the values 1.5, 2, 2.5, and 4 in our four different treatments.[10] Subjects were presented with both the b/c representation and the resulting pre-error payoff matrix, in neutral language (the choices were labeled A and B as opposed to the "C vs. D" choice that is standard in the prisoner's dilemma). We used the exchange rate of 30 units = $1. Subjects were given a show-up fee of $10 plus their winnings from the repeated prisoner's dilemma and an end-of session dictator game. To allow for negative stage-game payoffs, subjects began the session with an "endowment" of 50

[10] This payoff specification gives us a simple one-parameter ordering of the treatments; we do not think it is essential for our results. Note that the specification implies that the short-run gain to playing D instead of C is independent of the other player's action. The prisoner's dilemma is more general than this; its defining characteristics are that D is a dominant strategy and that both playing C yields the highest payoff - in particular both playing C should be more efficient than alternating between (C,D) and (D,C).

units (in addition to the show–up fee).[11] On average subjects made $22 per session, with a range from $14 to $36. Sessions lasted approximately 90 minutes.[12]

A total of 384 subjects voluntarily participated at the Harvard Decision Science Laboratory in Cambridge, MA. In each session, 12-32 subjects interacted anonymously via computer using the software Z-Tree (Urs Fischbacher 2007) in a sequence of infinitely repeated prisoner's dilemmas (see Table 1 for summary statistics on the different treatments). We conducted a total of 18 sessions between September 2009 and October 2010.[13] We only implemented one treatment during a given session, so each subject participated in only one treatment. To rematch subjects after the end of each repeated game, we used the turnpike protocol as in Dal Bó (2005). Subjects were divided into two equally-sized groups, A and B. A-subjects only interacted with B-subjects and vice versa, so that no subject ever played twice with another subject, or with a subject who has played with a subject they have played with, so that subjects could not influence the play of subjects they interacted with in the future. [14] Subjects were informed about this setup. To implement random game lengths we pre-generated a sequence of integers $t_1, t_2, ..$ according to the specified geometric distribution to use in all sessions, such that in each session every first interaction lasted t_1 rounds, every second interaction lasted t_2 etc.[15]

[11] No subject ever had fewer than 19 units, and only 4 out of 278 subjects ever dropped below 40 units.

[12] Subjects were given at most 30 seconds to make their decision, and informed that after 30 seconds a random choice would be made. The average decision time was 1.3 seconds, much less than the 30 second limit, and the frequency of random decisions was very low, 0.0055.

[13] All sessions were conducted during the academic year, and all subjects were recruited through the CLER lab at Harvard Business School using the same recruitment procedure. Subject demographics varied relatively little across sessions and treatments. See online appendix for demographic summary statistics by session.

[14] Thus the maximum number of interactions in a session with N subjects was N/2, and in each session we ran the maximum number of interactions. This explains why the average number of interactions differs between the different treatments. The average numbers of interactions per subject in Table 1 are not integers because there were multiple sessions per treatment.

[15] The starting place in the sequence of random game lengths that was used in the experiment was picked by the programmer, and the sequence following the chosen starting place had an unusually low number of short games. As a result the overall distribution of game lengths differed from what would be expected from a geometric distribution, which raises the concern that the subjects could have noticed this and adjusted their play. However, analysis of the data shows that subjects did not become less likely to cooperate in later rounds over the course of the session; see the online appendix for details.

Figure 1. *Payoff matrices for each specification. Payoffs are denoted in points.*

Realized payoffs | **Expected payoffs**

b/c = 1.5

	C	D
C	1,1	-2,3
D	3,-2	0,0

b/c = 1.5, E = 1/8

	C	D
C	0.875, 0.875	-1.375, 2.375
D	2.375, -1.375	0.125, 0.125

b/c = 2

	C	D
C	2,2	-2,4
D	4,-2	0,0

b/c = 2, E = 1/8

	C	D
C	1.75, 1.75	-1.25, 3.25
D	3.25, -1.25	0.25, 0.25

b/c = 2.5

	C	D
C	3,3	-2,5
D	5,-2	0,0

b/c = 2.5, E = 1/8

	C	D
C	2.625, 2.625	-1.125, 4.125
D	4.125, -1.125	0.375, 0.375

b/c = 4

	C	D
C	6,6	-2,8
D	8,-2	0,0

b/c = 4, E = 1/8

	C	D
C	5.25, 5.25	-0.75, 6.75
D	6.75, -0.75	0.75, 0.75

b/c = 4, E = 1/16

	C	D
C	5.625, 5.625	-1.375, 7.375
D	7.375, -1.375	0.375, 0.375

Table 1. *Summary statistics for each treatment.*

	b/c=1.5 E=1/8	**b/c=2 E=1/8**	**b/c=2.5 E=1/8**	**b/c=4 E=1/8**	**b/c=4 E=1/16**	**b/c=4 E=0**
Sessions per treatment	3	2	3	4	3	3
Subjects per treatment	72	52	64	90	58	48
Average number of interactions	11	11.5	10.7	11.3	9.9	7.8
Average number of rounds per interaction	8.4	8.3	8.3	8.1	8.0	8.2

Following the end of the series of repeated prisoner's dilemmas, subjects played a dictator game and answered survey questions related to prosocial behavior, motivation, strategies and demographics (see Dreber, Fudenberg and Rand (2010) for more information).

II. Theoretical and Experimental Background

We begin by analyzing the set of equilibria of the various specifications. In all of the treatments, the only static equilibrium is to defect. In the treatment with b/c=1.5, the only Nash equilibrium is ALLD, while the other treatments all allow cooperative equilibria.[16] As there are no explicit characterization theorems for the entire set of equilibrium outcomes for noisy repeated games with fixed discount factors, our initial analysis focused on a few repeated game strategies that have previously received attention.

In particular, we chose the payoffs so that when b/c=4, the memory-1 strategy PTFT- "Play C if yesterday's outcome was (C,C) or (D,D) and otherwise play D" -is an equilibrium. This strategy has received a great deal of attention in the literature on evolutionary game theory, where it is called Win-Stay Loose-Shift (Nowak and Sigmund 1993; Wedekind and Milinski 1996; Martin Posch 1999; Imhof, Fudenberg and Nowak 2007). When both players use PTFT, one play of D (either intentionally or by mistake) leads to one round of (D,D) followed by a return to the equilibrium path; the strategy is called "perfect" because this error-correcting property allows it to be subgame-perfect, in contrast to TFT which typically is not.[17] PTFT has theoretical appeal because it is error-correcting and has only memory 1, but we conjecture that most subjects will view it as counterintuitive to cooperate after mutual defection, which raises the question of how widely the strategy is actually used.

Standard equilibrium analysis predicts no cooperation when b/c=1.5, but offers little guidance when b/c is large enough that there are cooperative equilibria. The evolutionary game theory models of Nowak and Sigmund (1993) and Imhof, Fudenberg and Nowak (2007) predict cooperation when b/c=4, and moreover predict that subjects will play PTFT. Since these analyses

[16] Because the error term is strictly positive regardless of the actions played, every information set is reached with positive probability, and Nash equilibrium implies sequential rationality. Thus in the games with errors every Nash equilibrium is a sequential equilibrium, and every pure-strategy Nash equilibrium is equivalent to a perfect public equilibrium.

[17] In the game without errors, PTFT is a subgame-perfect equilibrium if $c < \delta(b-c)$ or $\delta > 1/(b/c-1)$ which is the case when b/c=2.5 or 4. Analysis of the game with errors shows that PTFT is not an equilibrium when b/c = 2 or 2.5, essentially because the errors lower the expected value of cooperation, but PTFT is an equilibrium of the game with errors when b/c=4. See the online appendix for details.

restrict attention to strategies that base their decisions only on the last round's outcome, however, they predict defection when b/c=2 or 2.5. The Fudenberg and Maskin (1990, 1994) evolutionary analysis of repeated games with vanishingly rare errors predicts cooperation in all three treatments with cooperative equilibria, but does not provide a precise prediction of what strategies will be played.

Experimental work on repeated games without errors also suggests that cooperation is more likely when it is more beneficial, and in particular that the existence of an cooperative equilibrium is necessary but not sufficient for there to be a substantial amount of cooperation (e.g., Alvin E. Roth and J. Keith Murnighan 1978; Murnighan and Roth 1983; Robert A. Feinberg and Thomas A. Husted 1993; Duffy and Ochs 2004; Dal Bó 2005; Dreber et al. 2008; Dal Bó and Frechette 2011; Matthias Blonski, Peter Ockenfels and Giancarlo Spagnolo 2011). Blonski and Spagnolo (2004) proposed that the key to whether cooperation occurs is whether TFT is a best response to a ½-½ probability distribution over TFT and ALLD, i.e. whether TFT risk-dominates ALLD in the game with only those two strategies.[18] Dal Bó and Frechette (2011) and Blonski, Ockenfels and Spagnolo (2011) find empirical support for this risk-dominance criterion in games without noise; and in a re-analysis of the no-noise experiments of Dreber et al. (2008), we find that the same criterion successfully predicts when cooperation is stable.

The success of the "risk dominance by TFT" criterion in experiments without noise raises the question of whether a similar criterion will explain cooperation in games with noise. One complication is that in our experiment, noise lowers the payoff of TFT against itself sufficiently that TFT is not an equilibrium of the overall game.[19] It is however an equilibrium of the 2x2 game where players are restricted to play either TFT or ALLD; in this 2x2 game TFT is only risk dominant if b/c=2.5 or 4. Thus to the extent that the risk-dominance criterion extends to games with noise, it predicts substantially more cooperation when b/c=2.5 than when b/c=2.

This combination of observations about equilibria of the game and insights from past experiments leads to our first set of experimental questions:

[18] This is the case when the present value of cooperation is sufficiently high compared to the loss caused by one period of (C,D). Note that Blonski, Ockenfels and Spagnolo (2011) offer an alternative theoretical justification for this equilibrium selection criterion.

[19] See the online appendix for TFT equilibrium calculations.

QUESTION 1: Is cooperation more frequent in treatments where there are cooperative equilibria? Is risk dominance of TFT over ALLD a good predictor of cooperation?

QUESTION 2: Is there substantially more cooperation when cooperation is the outcome of an equilibrium in strategies which base their play on outcomes in the previous round?

The answers to questions 1 and 2 provide some indirect evidence on the strategies that subjects use. To get a more precise understanding of the particular strategies being employed by subjects, we examine more direct evidence on their play. This leads to our four additional experimental questions:

QUESTION 3: What strategies do subjects use in the noisy prisoner's dilemma? Do subjects use PTFT when it is an equilibrium strategy?

QUESTION 4: How do the strategies used vary with the gains to cooperation?

QUESTION 5: How do the strategies used vary with the level of noise?

III. Methodology

The general theory of repeated games, like that of extensive form games, views strategies as complete contingent plans, which specify how the player will act in every possible information state. In practice, cognitive constraints may lead subjects to use relatively simple strategies, corresponding to automata with a small number of internal states. However, it is unclear what a priori restrictions one should impose on subjects' strategies, and one of the goals of our experiment is to let the data reveal what sorts of strategies are actually used. For this reason we did not want to use the "strategy method," where subjects are asked to pick a strategy that is implemented for them: The full set of strategies is infinite, so forcing subjects to choose a strategy that depends only on the previous round's outcome (memory-1) is much too restrictive while allowing for all strategies that depend on the last two periods (memory-2) gives too large a strategy set to explicitly present. In addition, we would like to consider some simple strategies such as Grim, which can be viewed as a two-state automata but has arbitrarily long memory, as its current play depends on whether "D" was played in any previous round. As the data cannot discriminate between all possible repeated game strategies, we used a combination of prior intuition, survey responses and data analysis to identify a small set of strategies that seem to best describe actual play.

There has been comparatively little past work on identifying the strategies subjects use in repeated game experiments. In the repeated prisoner's dilemma, Wedekind and Milinski (1996) note that subjects rarely play C in the round after they played D and the opponent played C, and take this as evidence of PTFT, but since subjects rarely played C following (D,D), their data seems more consistent with some sort of grim strategy. Dal Bó and Frechette (2011) used maximum likelihood to estimate the proportions of subjects using one of six ex-ante relevant strategies. They find that ALLD and TFT account for the majority of their data. Aoyagi and Frechette (2009) study experimental play of a prisoner's dilemma where subjects do not observe their partner's actions but instead observe a noisy symmetric signal of it. The signal is a real number, and has the same distribution under (C,D) and (D,C), so that commonly discussed prisoner's dilemma strategies such as TFT are not implementable. They find that subjects play "trigger strategies" of memory 1, except in the limit no-noise case where signals from two rounds ago also have an impact. Engle-Warnick and Slonim (2006) study the strategies used in a repeated sequential-move trust game by counting how many observations of a subject's play in a given interaction is exactly described by a given strategy. They find that in most of the interactions, the investors choose actions which are consistent with a grim trigger strategy, while the play of the trustees is more diverse. Gabriele Camera, Marco Casari and Maria Bigoni (2010) on the other hand find little evidence of grim trigger strategies on the subject level when subjects are randomly put in groups of four to play an indefinitely repeated PD where in each round they are randomly matched with one subject in the group. Even though behavior at the aggregate looks like grim trigger, individual behavior is far more heterogeneous.

An advantage of studying repeated games with errors is that we can more easily identify different strategies: In the absence of errors a number of repeated game strategies are observationally equivalent, for example if a pair of subjects cooperates with each other in every round, we see no data on how they would have responded to defections. Thus the introduction of errors has a methodological advantage as well as a substantive one, as the errors will lead more histories to occur and thus make it easier to distinguish between histories.

To have any hope of inferring the subjects' strategies from their play, we must focus our attention on a subset of the infinitely many repeated game strategies. We begin with strategies that have received particular attention in the theoretical literature: ALLD, ALLC, Grim[20], TFT, and PTFT. Because one round of punishment is only enough to sustain cooperation in one of our four treatments (when b/c=4) we also include modifications of TFT and PTFT that react to D with two rounds of defection, we call these 2TFT and 2PTFT. We also include the strategy T2 used by Dal Bó and Frechette (2011).[21]

To inform our extension of this strategy set, we asked subjects to describe their strategies in a post-experimental survey. Several regularities emerged from these descriptions. Many subjects reported 'giving the benefit of the doubt' to an opponent on the first defection, assuming that it was a result of noise rather than purposeful malfeasance; only after two or three defections by their partner would they switch to defection themselves.[22] We refer to this slow-to-anger behavior as 'leniency'. None of the strategies mentioned above are lenient; note that leniency requires looking further into the past than permitted by memory-1 strategies. Subjects also varied in the extent to which they reported being willing to return to cooperation following a partner's defection. We refer to this strategic feature as 'forgiveness', which is an often-discussed aspect of TFT (as opposed to Grim, for example); 2TFT also shows forgiveness, as do PTFT and 2PTFT, although only following mutual defection.

In response to the subjects' strategy descriptions, we added several lenient strategies to our analysis. Because our games were on average only 8 rounds in length, we have limited power to explore intermediate levels of forgiveness between TFT and Grim, so we restrict strategies to either forgive after 1 to 3 rounds or to never forgive (as with Grim and its lenient variants).[23]

---

[20] As in Dal Bó & Frechette (2011), our specification of Grim begins by playing C and then switches permanently to D as soon as either player defects.

[21] 2TFT initially plays C, then afterwards plays C if opponent has never played D or if the opponent played C in both of the previous two rounds. A defection by the partner triggers two rounds of punishment defection in both 2TFT and T2. However, T2 automatically returns to C following the two Ds, regardless of the partner's play during this time, while 2TFT only returns to C if the partner played C in both of the "punishment rounds." Additionally, T2 begins its punishment if either player defects, whereas 2TFT responds only to the partner's defection.

[22] Subjects' free-response descriptions of their strategies are reproduced in the online appendix.

[23] More generally the average length of 8 interactions imposes restrictions on our ability to estimate the extent to which subjects condition on long histories. This constraint reflects a tradeoff between our interest in the way subjects use past history and our desire to limit sessions to 90 minutes (to avoid fatiguing the subjects) while allowing them to play enough interactions to have some chance to learn. Fortunately the average length of 8 was enough to provide convincing evidence that subjects can use forgiving strategies with memory greater than 1.

As strategies that are both lenient and forgiving, we include TFT variants that switch to defection only after the other player chooses D multiple times in a row, considering TF2T (plays D if the partner's last two moves were both D) and TF3T (plays D if the partner's last three moves were D). For strategies that are lenient but not forgiving, we include Grim variants that wait for multiple rounds of D (by either player) before switching permanently to defection, considering Grim2 (waits for two consecutive rounds in which either player played D) and Grim3 (waits for three consecutive D rounds). We include three strategies that punish twice (intermediate to TFT's one round of punishment and Grim's unending punishment) but can be implemented by conditioning only on the last 3 rounds. These are T2 and 2TFT, which were discussed above, and 2TF2T ("2 Tits for 2 Tats"), which waits for the partner to play D twice in a row, and then punishes by playing D twice in a row. Because we do not include strategies that punish for a finite number of rounds greater than two, our estimated share of "Grim" strategies may include some subjects who use such strategies with more than two rounds of punishment.

Other subjects indicated that they used strategies which tried to take advantage of the leniency of others by defecting initially and then switching to cooperation. Thus we consider 'exploitive' versions of our main cooperative strategies that defect on the first move and then return to the strategy as normally specified: D-TFT[24], D-TF2T, D-TF3T, D-Grim2 and D-Grim3.[25] Because TF2T appears prevalent in many treatments, we also looked at whether subjects used the strategy that alternates between D and C (DC-Alt), as this strategy exploits the leniency and forgiveness of TF2T. Lastly, some subjects reported playing strategies which give the first impression of being cooperative and then switch to defection, hoping the partner will assume the subsequent Ds are due to error. Therefore we include a strategy which plays C in the first round and D thereafter (C-to-ALLD). Each strategy is described verbally in Table 2; complete descriptions are given in the online appendix.

---

[24] Boyd and Jeffrey P. Lorberbaum (1987) call this strategy "Suspicious Tit for Tat."

[25] ALLD and D-Grim are identical except when the other player plays C in the first round, and you mistakenly also play C in the first round: here D-Grim cooperates while ALLD defects. Thus we do not include D-Grim in our analysis as we do not have sufficient number of observations per subject to differentiate between the two.

Table 2. *Descriptions of the 20 strategies considered.*

Strategy	Abbreviation	Description
Always Cooperate	ALLC	Always play C
Tit-for-Tat	TFT	Play C unless partner played D last round
Tit-for-2-Tats	TF2T	Play C unless partner played D in both of the last 2 rounds
Tit-for-3-Tats	TF3T	Play C unless partner played D in all of the last 3 rounds
2-Tits-for-1-Tat	2TFT	Play C unless partner played D in either of the last 2 rounds (2 rounds of punishment if partner plays D)
2-Tits-for-2-Tats	2TF2T	Play C unless partner played 2 consecutive Ds in the last 3 rounds (2 rounds of punishment if partner plays D twice in a row)
T2	T2	Play C until either player plays D, then play D twice and return to C (regardless of all actions during the punishment rounds)
Grim	Grim	Play C until either player plays D, then play D forever
Lenient Grim 2	Grim2	Play C until 2 consecutive rounds occur in which either player played D, then play D forever
Lenient Grim 3	Grim3	Play C until 3 consecutive rounds occur in which either player played D, then play D forever
Perfect Tit-for-Tat / Win-Stay-Lose-Shift	PTFT	Play C if both players chose the same move last round, otherwise play D
Perfect Tit-for-Tat with 2 rounds of punishment	2PTFT	Play C if both players played C in the last 2 rounds, both players played D in the last 2 rounds, or both players played D 2 rounds ago and C last round. Otherwise play D
Always Defect	ALLD	Always play D
False cooperator	C-to-ALLD	Play C in the first round, then D forever
Exploitive Tit-for-Tat	D-TFT	Play D in the first round, then play TFT
Exploitive Tit-for-2-Tats	D-TF2T	Play D in the first round, then play TF2T
Exploitive Tit-for-3-Tats	D-TF3T	Play D in the first round, then play TF3T
Exploitive Grim2	D-Grim2	Play D in the first round, then play Grim2
Exploitive Grim3	D-Grim3	Play D in the first round, then play Grim3
Alternator	DC-Alt	Start with D, then alternate between C and D

To assess the prevalence of each strategy in our data, we follow Dal Bó and Frechette (2011) and suppose that each subject chooses a fixed strategy at the beginning of the session (or alternatively for the last four interactions, when we restrict attention to those observations)[26], and moreover that in addition to the extrinsically imposed execution error, subjects make mistakes when choosing their intended action, so every sequence of choices (e.g. of intended actions) has positive probability.[27] More specifically, we suppose that if subject *i* uses strategy *s*, her chosen action in round *r* of interaction *k* is C if $s_{ikr}(s) + \gamma\varepsilon_{ikr} \geq 0$, where $s_{ikr}(s) = 1$ if strategy *s* says to play C in round *r* of interaction *k* given the history to that point, and $s_{ikr}(s) = -1$ if *s* says to play D. Here ε_{ikr} is an error term that is independent across subjects, rounds, interactions, and histories, γ parameterizes the probability of mistakes, and the density of the error term is such that the overall likelihood that subject *i* uses strategy *s* is

$$(1) \qquad p_i(s) = \Pi_k \Pi_r \left(\frac{1}{1+\exp(-s_{ikr}(s)/\gamma)}\right)^{y_{ikr}} \left(\frac{1}{1+\exp(s_{ikr}(s)/\gamma)}\right)^{1-y_{ikr}},$$

where y_{ikr} is 1 if the subject chose C and 0 if the subject chose D. [28]

To better understand the mechanics of the specification, suppose that an interaction lasts *w* rounds, that in the first round the subject chose C, the first round outcome was that the subject played C and her partner played D, and in the second round the subject chose D. Then for strategy *s* = TFT, which plays C in the first round, and plays D in the second round following (C,D), the likelihood of the subject's play is the probability of two "no-error" draws. This is the same probability that we would assign to the overall sequence of the subject's play given the play of the opponent - it makes no difference whether we compute the likelihood round by round or for the whole interaction.

For any given set of strategies S and proportions p¸ we then derive the likelihood for the entire sample, namely $\sum_I \ln \sum_{s\in S} p(s)p_i(s)$. Note that the specification assumes that all

---

[26] We found that conducting the MLE supposing that subjects pick a fixed strategy at the beginning of each interaction, as opposed to using the same strategy throughout the session, gave qualitatively similar results to those presented below.

[27] Recall that we, unlike our subjects, observe the intended actions as well as the implemented ones. We use this more informative data in our estimates.

[28] Thus the probability of an error in implementing one's strategy is 1/(1+exp(1/γ)). Note that this represents error in intention, rather than the experimentally imposed error in execution. This formulation assumes that all strategies have an equal rate of implementation error. In the online appendix we show that the MLE estimates of strategy shares are robust to allowing each strategy have a different value of γ.

subjects are ex-ante identical with the same probability distribution over strategies and the same distribution over errors; one could relax this at the cost of adding more parameters. Because *p* describes a distribution over strategies, this likelihood function implies that in a very large sample we expect fraction *p(s)* of subjects to use strategy s, though for finite samples there will be a non-zero variance in the population shares. We use maximum likelihood estimation (MLE) to estimate the prevalence of the various strategies, and bootstrapping to associate standard errors with each of our frequency estimates. We construct 100 bootstrap samples for each treatment by randomly sampling the appropriate number of subjects with replacement. We then determine the standard deviation of the MLE estimates for each strategy frequency across the 100 bootstrap samples.

To investigate the validity of this estimation procedure, we tested it on simulated data. For a given strategy frequency distribution, we assigned strategies to 3 groups of 20 computer agents. We then generated a simulated history of play across 4 interactions by randomly pairing members of each group to play games with representative lengths from the game length sequence used in the experiment (t_1=5, t_2=11, t_3=8, t_4=9). As in the main experimental treatments, we included a 1/8 probability of error, and recorded both the intended and actual action of each agent. We generated simulated data in this way using strategy distributions similar to those estimated from the experimental data (see Table 3), and then used the MLE method described above to estimate the strategy frequencies. The MLE results were consistent with the actual strategy frequencies, giving us confidence in the estimation procedure.[29]

IV. Results

We begin by examining behavior in the 4 treatments with E=1/8; we will then compare this to the behavior in the controls with E=0 and E=1/16 using b/c=4, which is the ratio most favorable to cooperation.

Examining play in the first round of each interaction, as displayed in Figure 2, suggests that there was some learning, except perhaps when b/c=1.5; this is confirmed by the statistical analysis reported in Appendix B. To reduce the potential effects of learning while striking a

[29] See Appendix A for MLE results using simulated data.

balance with the need for data, our analysis will focus on how subjects played in the last four interactions of the session, which is roughly the last third of each session.[30]

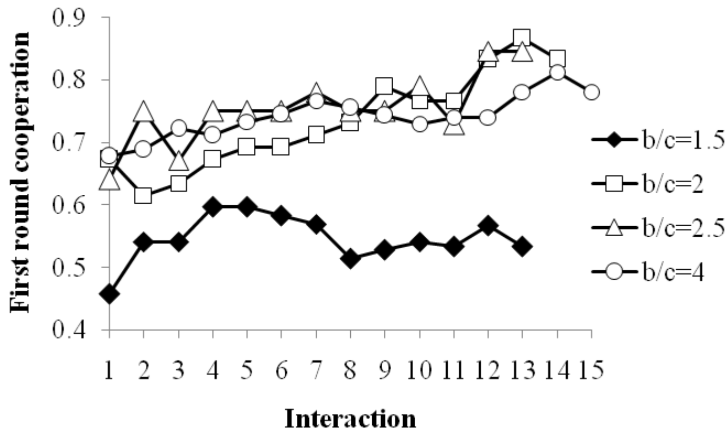


Figure 2. *First round cooperation over the course of the session, by payoff specification.*

QUESTION 1: Is cooperation more frequent in treatments where there are cooperative equilibria? Is risk dominance of TFT over ALLD a good predictor of cooperation?

Figure 3 reports both cooperation in the first round of the last 4 interactions and the average cooperation over the last 4 interactions as a whole, which can depend on the relationship between the two subjects' strategies and also on possible random errors.[31] We see that there is markedly less cooperation when b/c=1.5, both in the first round (1.5 vs. 2, p=0.016; 1.5 vs. 2.5, p=0.003; 1.5 vs. 4, p=0.004) and overall (1.5 vs. 2, p=0.001; 1.5 vs. 2.5, p<0.001; 1.5 vs. 4, p<0.001). Conversely, we see little difference in first round cooperation between the three treatments with cooperative equilibria (2 vs. 2.5, p=0.71; 2 vs. 4, p=0.83; 2.5 vs. 4, p=0.87); and while there is an increase in overall cooperation going from b/c=2 to b/c=2.5, this increase is smaller than that between b/c=1.5 and b/c=2 and is only marginally significant (p=0.058). Moreover there is no significant difference in overall cooperation between b/c=2.5 and b/c=4 (p=0.73). The reason there is about the same amount of initial cooperation in b/c=2 and 2.5 yet

[30] Our results are not sensitive to this particular cutoff. Using either the last 6 or last 2 interactions instead yields very similar results. See the online appendix for details.

[31] For each pairwise b/c comparison, we report the results of a logistic regression over first-round/all individual decisions, with a b/c value dummy as the independent variable, clustered on both subject and interaction pair.

somewhat more overall cooperation in the latter case seems related to the fact that more subjects are more forgiving in the latter treatment, as seen in the discussion of Questions 3-4. Because the largest difference in cooperation occurs between b/c=1/5 and b/c=2, as opposed to between b/c=2 and b/c=2.5, the data do not show the strong support for risk dominance of TFT as the key determinant of the level of cooperation in games with noise that was seen in studies of games without noise.

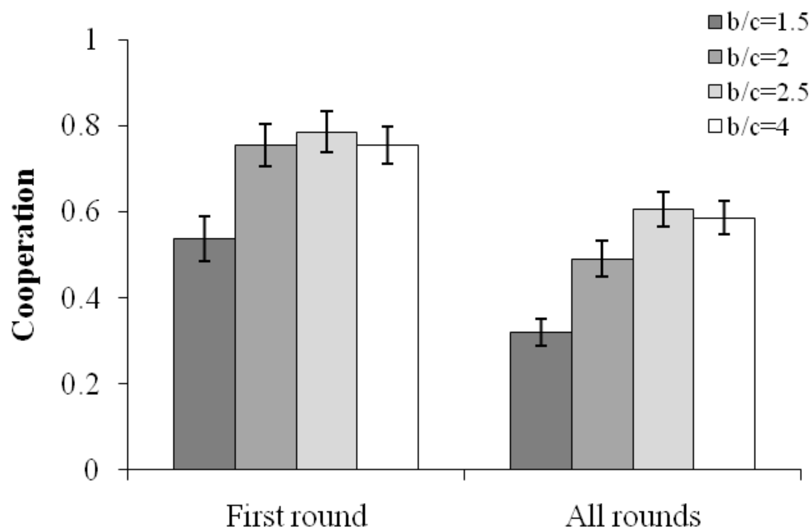


Figure 3. *First round and overall cooperation by payoff specification, averaged over the last 4 interactions of each session. See Table 4 for a list of the cooperation frequencies displayed here. Error bars indicate standard error of the mean, clustered on subject and interaction pair.*

QUESTION 2: Is there substantially more cooperation when cooperation is the outcome of an equilibrium in strategies which base their play on outcomes in previous round?

Indeed, we see a substantial amount of cooperation when b/c=2 and 2.5, even though cooperative equilibria in these treatments require memory 2 or more.[32] To the extent that play resembles an equilibrium of the repeated game, these results are a first sign that the predictions of the memory-1 restriction are not consistent with the data.

[32] This theoretical result is robust to considering memory-1 strategies which forget the state with some non-zero probability. Regardless of the probability of forgetting, TFT and PTFT are not equilibria at E=1/8 and b/c=2 or b/c=2.5. See the online appendix for further discussion.

QUESTION 3: What strategies do subjects use in the noisy prisoner's dilemma? Do subjects use PTFT when it is an equilibrium strategy?

We now report the results of the MLE analysis of strategy choice, examining the last 4 interactions of each session.[33] We consider 20 strategies in total (Table 2): the fully cooperative strategies ALLC, TFT, TF2T, TF3T, 2TFT, 2TF2T, Grim, Grim2, Grim3, PTFT, 2PTFT and T2, which always play C against themselves in the absence of errors; the fully non-cooperative strategies ALLD and D-TFT, which always play D against themselves in the absence of errors; and the partially cooperative strategies C-to-ALLD, D-TF2T, D-TF3T, D-Grim2, D-Grim3 and DC-Alt, which play a combination of C and D against themselves in the absence of error.[34] Of these, only 11 are present at frequencies significantly greater than 0 in at least one payoff specification: the cooperative strategies ALLC, TFT, TF2T, TF3T, 2TFT, 2TF2T, Grim, Grim2 and Grim3, and the non-cooperative strategies ALLD and D-TFT.[35]

Thus we restrict our attention to these 11 strategies (Table 3). We do not find any evidence of subjects using PTFT in any payoff specification - PTFT never received a positive weight in any of the bootstrapped samples. In the treatments with cooperative equilibria, the most common cooperative strategies TF2T, TF3T, Grim2 and Grim3 are all lenient.

[33] Analysis of simulated data suggests that MLE on 4 interactions lasting on average 8 rounds detects strategies whose frequencies are 5% or higher, but that lower-frequency strategies may not be detected.

[34] Following Dal Bo and Frechette, we consider only pure strategies in the MLE estimation. The evolutionary game theory literature suggests the consideration of "generous tit-for-tat" or GTFT (Nowak and Sigmund 1990), which cooperates with probability strictly between 0 and 1 if the partner defected last period, and otherwise cooperates. Simulations show that this strategy will be identified by our estimation as playing TF3T, but that expanding our MLE procedure to include stochastic strategies can differentiate GTFT from TF3T. Doing so suggests that the majority of leniency observed in our data is not in fact the result of stochastic memory-1 strategies, but is rather due to lenient strategies with longer memories. The online appendix reports these results, which should be viewed as a first step towards understanding how to test for mixed strategies; we hope to explore the issues posed by mixed strategies in greater detail in future work.

[35] See Appendix C for the estimates and standard errors for the full set of 20 strategies. At the request of the referees, we also explored versions of TFT and PTFT which forget the state with some non-zero probability, versions of TFT and Grim which ignore a defection in the first round, and a family of strategies which cooperate until the fraction of D by their partner passes some threshold. None of these strategies were present at frequencies significantly greater than 0. See the online appendix for the details of these robustness checks.

Table 3. *Maximum likelihood estimates using the last 4 interactions of each session. All payoff specifications use error rate E=1/8. Bootstrapped standard errors (shown in parentheses) used to calculate p-values.*
*† Significant at p<0.1, * Significant at p<0.05, ** Significant at p<0.01.*

	b/c=1.5	b/c=2	b/c=2.5	b/c=4
ALLC	0 (0)	0.03 (0.03)	0 (0.02)	0.06† (0.03)
TFT	0.19** (0.05)	0.06 (0.04)	0.09* (0.04)	0.07* (0.03)
TF2T	0.05 (0.03)	0 (0)	0.17* (0.06)	0.20** (0.07)
TF3T	0.01 (0.01)	0.03 (0.03)	0.05 (0.05)	0.09* (0.04)
2TFT	0.06 (0.04)	0.07† (0.04)	0.02 (0.02)	0.03 (0.02)
2TF2T	0 (0.02)	0.11* (0.05)	0.11† (0.06)	0.12* (0.05)
Grim	0.14** (0.04)	0.07 (0.05)	0.11* (0.04)	0.04† (0.02)
Grim2	0.06† (0.03)	0.18** (0.06)	0.02 (0.03)	0.05† (0.03)
Grim3	0.06 (0.03)	0.28** (0.08)	0.24** (0.07)	0.11** (0.04)
ALLD	0.29** (0.06)	0.17** (0.06)	0.14** (0.04)	0.23** (0.04)
D-TFT	0.14** (0.05)	0 (0)	0.05† (0.03)	0 (0)
Gamma	*0.46*** *(0.02)*	*0.5*** *(0.03)*	*0.49*** *(0.03)*	*0.43*** *(0.02)*

Note that in all treatments the MLE assigns a substantial share to strategies that depend on outcomes from more than one period ago. To provide additional evidence for this conclusion, we use a logistic regression to test whether subjects condition play on their partner's decision two rounds ago. Bias introduced by heterogeneity presents a potential challenge for this approach: the other's play two rounds ago interacts with own play two rounds ago to determine the history in the previous round, so other's play two rounds ago could have a spuriously significant coefficient in a heterogeneous population of subjects all of whom use memory 1 strategies. To control for bias introduced by heterogeneity, we include controls for the type of the player making the decision, as in Aoyagi & Frechette (2009).[36] We conduct a logistic regression with correlated random effects, regressing own decision in round t against own play in round t-1, other's play in t-1, own play in t-2 and other's play in t-2, and including controls for b/c ratio and own average frequency of first round cooperation and overall cooperation, both over the last 4 interactions.[37] Consistent with the use of longer memories, we find a significant effect of other's play two rounds ago (coeff=1.01, p<0.001). This result supports the conclusion that many subjects are conditioning on more than only the last round.

QUESTION 4: How do the strategies used vary with the gains to cooperation?

The strategies employed by subjects clearly vary according to the gains from cooperation. This can be seen from descriptive statistics analyzing aggregate behavior as well as from the MLE analysis, both of which are summarized in Table 4. Three trends are apparent.

First, cooperation is significantly lower at b/c=1.5 than at the higher b/c ratios, as shown in Table 4 and visualized in Figure 3. Consistent with this observation, Table 4 also shows that the share of the non-cooperative strategies ALLD and D-TFT is 43% when b/c=1.5, which is substantially and significantly higher than in other b/c conditions (b/c=2, 17%, p<0.001; b/c=2.5, 19%, p<0.001; b/c=4, 23%, p=0.001).[38]

Second, leniency also increases when moving from b/c=1.5 to the higher b/c ratios. To get a measure of leniency distinct from the MLE estimates, we examine all histories in which both subjects played C in all but the previous round, while in the previous round one subject played D.[39] We then ask how frequently the subject who had hitherto cooperated showed leniency by

---

[36] For a general treatment of the topic, see Gary Chamberlain (1980) and James J. Heckman (1981).

[37] When we simulate data for various combinations of memory-1 strategies, we find that this regression returns a significant coefficient on partner's action in t-2 no more often than predicted by chance. We also find that when simulating only memory-1 strategies, the size of the estimated coefficient of play two rounds ago is at least an order of magnitude smaller than the coefficients for play last round, in contrast to the estimates on the experimental data; see the online appendix for a detailed discussion of these issues. Note that as the regression conditions on play two rounds ago, it necessarily omits decisions made in the first two rounds of each interaction.

[38] For each pairwise comparison of aggregated MLE coefficients, we report the results of a two-sample t-test using bootstrapped standard errors of the aggregated coefficients.

[39] We also include second round decisions in which the first round's outcome was CD.

continuing to cooperate despite the partner's defection.[40] At b/c=1.5, 17% of histories show leniency, compared to the significantly higher values of 63% at b/c=2 (b/c=1.5 vs. b/c=2, p=0.001), 67% at b/c=2.5 (b/c=1.5 vs. b/c=2.5, p<0.001) and 66% at b/c=4 (b/c=1.5 vs. b/c=4, p<0.001). No significant difference in leniency exists among the higher b/c ratios (p>0.20 for all comparisons). Thus leniency increases across the transition from b/c=1.5 to b/c=2. Analyzing strategy frequencies paints a similar picture. The combined frequency of the lenient strategies ALLC, TF2T, TF3T, 2TF2T, Grim2 and Grim3 is 18% at b/c=1.5 which is significantly less than at b/c=2 (62%, p<0.001), b/c=2.5 (60%, p<0.001) or b/c=4 (63%, p<0.001).

Table 4. *Descriptive statistics of aggregate behavior, as well as aggregated MLE frequencies from Table 3. All specifications use E=1/8. The descriptive statistics for leniency and forgiveness are defined in the text. For MLE aggregation, all strategies other than ALLD and D-TFT are cooperative; lenient strategies are TF2T, TF3T, 2TF2T, Grim2 and Grim3; and forgiving strategies are TFT, TF2T, TF3T, 2TFT and 2TF2T.*

	b/c=1.5	b/c=2	b/c=2.5	b/c=4
Descriptive statistics				
%C First Round	54%	75%	79%	76%
%C All Rounds	32%	49%	61%	59%
Leniency	29%	63%	67%	66%
Forgiveness	15%	18%	33%	32%
MLE aggregation				
Cooperative strategies	57%	83%	81%	77%
Lenient strategies	18%	62%	60%	63%
Forgiving strategies	31%	29%	44%	57%

The gains to cooperation also influence the frequency of forgiveness. Forgiveness is more complicated to define and measure, as it describes a more complex pattern of behavior: to us it means that the players were initially cooperating, that one of them then defects leading the other player to "punish" the initial defector, and finally that the punishing player relents and returns to cooperation. To develop an operational measure of histories where forgiveness occurs, we first identify all histories in which (i) at least one subject chose C in the first round, (ii) in at least one

[40] For each pairwise b/c comparison of aggregate descriptive statistics, we report the results of a logistic regression over all decisions in lenient/forgiving histories, with a b/c dummy as the independent variable, clustered on both subject and interaction pair.

previous round, the initially cooperative subject chose C while the other subject chose D and (iii) in the immediately previous round the formerly cooperative subject played D. We then ask how frequently this formerly cooperative subject showed forgiveness by returning to C. For example, if the outcome in the first round is (C,D), and the first player plays D in the second round, and C in the third, we would say that the first player had "forgiven" the second player. We find significantly less forgiveness at b/c=1.5 (15%) and b/c=2 (18%) compared to b/c=2.5 (33%) and b/c=4 (32%) (1.5 vs. 2.5, $p<0.001$; 1.5 vs. 4, $p<0.001$; 2 vs. 2.5, $p=0.008$; 2 vs. 4, $p=0.007$). Thus forgiveness increases significantly when b/c increases from 2 to 2.5. This is again confirmed by examining strategy frequencies in the payoff specifications that support cooperation. The forgiving strategies ALLC, TFT, TF2T, TF3T, 2TFT and 2TF2T are less common at b/c=1.5 (31%) and b/c=2 (29%) than at b/c=2.5 (44%) and b/c=4 (57%) (1.5 vs. 2.5, $p=0.061$; 1.5 vs. 4, $p<0.001$; 2 vs. 2.5, $p=0.054$; 2 vs. 4, $p<0.001$; 2.5 vs. 4, $p=0.054$).

QUESTION 5: How do the strategies used vary with the level of noise?

To explore how play varies with the error rate, we now examine our two additional control treatments using b/c=4 with E=1/16 and E=0, and compare them to our results using b/c=4 at E=1/8. We begin by asking whether subjects condition on their partner's play two rounds ago in the last 4 interactions of each session, using a logistic regression with correlated random effects and regressing own decision in round t against other's play in round t-2, own play in t-2, other's play in t-1, own play in t-1, own first round cooperation frequency in the last 4 interactions and own frequency of cooperation in all rounds of the last 4 interactions. As with the E=1/8 treatments, we find a highly significant and sizable dependence on other's play two rounds ago for E=1/16 (coeff=1.221, $p<0.001$); while in the no-error E=0 control, however, we find no significant dependence on other's play two rounds ago (coeff=0.387, $p=0.247$). This provides our first direct evidence that the presence of noise plays an important role in strategy selection, promoting more complicated strategies.

Next we present the MLE results for each strategy in Table 5, as well as aggregated MLE results and descriptive statistics in Table 6. As shown in Table 6, overall cooperation is lower at E=1/8 (59%) than in the lower error conditions, and these differences are statistically significant

(E=1/16, 82%, p<0.001; E=0, 78%, p=0.002)[41]. Considering cooperation in the first round, there is somewhat less cooperation at E=1/8, but the differences are smaller and either not significant or just marginally so (E=1/8, 76%, E=1/16, 87%, E=0, 83%; 1/8 vs. 1/16, p=0.050, 1/8 vs. 0, p=0.346). We find no significant difference at E=0 compared to E=1/16 in either overall cooperation (p=0.486) or first round cooperation (p=0.338). Complementing this aggregate analysis, the share of the non-cooperative strategies ALLD and D-TFT is significantly larger at E=1/8 (23%) compared to E=1/16 (11%, p=0.001), and larger at E=1/8 than E=0 (16%, p=0.191) although the difference between E=0 and E=1/8 is not significant.[42] There is also no significant difference between E=0 and E=1/16 (p=0.297).

Turning to leniency, we examine cooperation frequency in the subset of histories in which leniency is possible, as described above in response to Question 5. At E=0, 42% of the eligible histories show leniency, compared to the significantly higher value of 77% at E=1/16 (p=0.001) and the marginally significantly higher value of 66% at E=1/8 (p=0.052). We also find that marginally significantly more eligible histories showed leniency at E=1/16 than at E=1/8 (p=0.071). We see similar results when analyzing strategy frequencies. The combined frequency of lenient strategies ALLC, TF2T, TF3T, 2TF2T, Grim2 and Grim3 is significantly lower at E=0 (40%) than at E=1/16 (82%, p<0.001) or E=1/8 (63%, p<0.001). As with the analysis of histories, we also see significantly more leniency at E=1/16 than at E=1/8 (p<0.001).

Considering cooperation frequency in histories with the potential for forgiveness, as described above in the response to Question 5 we also see significantly less forgiveness at E=0 (19%) compared to E=1/16 (47%, p<0.001) or E=1/8 (32%, p<0.001). We find no significant difference in forgiveness between E=1/16 than at E=1/8 (p=0.298). Examining the aggregated MLE frequencies, the forgiving strategies ALLC, TFT, TF2T, TF3T, 2TFT and 2TF2T are less common at E=0 (52%) compared to E=1/16 (78%, p=0.007) and at E=1/8 (57%) compared to E=1/16 (p=0.001). There is no significant difference between forgiving strategies at E=0 and E=1/8 (p=0.670).

[41] For each pairwise comparison of aggregate descriptive statistics, we report the results of a logistic regression over all decisions in lenient/forgiving histories, with an error rate dummy as the independent variable, clustered on both subject and interaction pair.
[42] For each pairwise comparison of aggregated MLE coefficients, we report the results of a two-sample t-test using bootstrapped standard errors of the aggregated coefficients.

Table 5. *Maximum likelihood estimates for our E=0, E=1/16 and E=1/8 conditions using the last 4 interactions of each session. All specifications use b/c=4. Bootstrapped standard errors (shown in parentheses) used to calculate p-values.*
*† Significant at p<0.1, * Significant at p<0.05, ** Significant at p<0.01.*

	E=0	E=1/16	E=1/8
ALLC	0.24*	0	0.06†
	(0.10)	(0.04)	(0.03)
TFT	0.14†	0.04	0.07*
	(0.08)	(0.04)	(0.03)
TF2T	0	0.24*	0.20**
	(0.02)	(0.10)	(0.07)
TF3T	0	0.42**	0.09*
	(0.04)	(0.09)	(0.04)
2TFT	0.15*	0	0.03
	(0.07)	(0)	(0.02)
2TF2T	0	0.08	0.12*
	(0)	(0.06)	(0.05)
Grim	0.15†	0.03	0.04†
	(0.08)	(0.02)	(0.02)
Grim2	0.16†	0.09	0.05†
	(0.09)	(0.05)	(0.03)
Grim3	0	0	0.11**
	(0.06)	(0)	(0.04)
ALLD	0.07†	0.05	0.23**
	(0.04)	(0.03)	(0.04)
D-TFT	0.09*	0.05	0
	(0.04)	(0.03)	(0)
Gamma	*0.35***	*0.44***	*0.43***
	(0.03)	*(0.03)*	*(0.02)*

In summary, we find that substantial levels of leniency and forgiveness are not unique to the high error rate of E=1/8, but are also present at the lower error rate of E=1/16. When the error rate is zero, leniency is much less frequent. The somewhat greater leniency and forgiveness at E=1/16 compared to E=1/8 is surprising; investigating this issue further is an interesting topic for future study.

To further explore the difference in play between error and no-error games, we reanalyze data from Dal Bó & Frechette (2011) and Dreber et al. (2008) using our strategy set from Table

2. Doing so finds TFT to be the most common cooperative strategy in all but one payoff specification.[43] Additionally, the aggregate frequency of strategies with memory at most 1 (namely ALLC, TFT, D-TFT, and ALLD) is 76% in the games without noise (including our E=0 control), compared to only 33% in our games with noise; this difference is largely driven by lenient strategies, most of which by definition look back more than one round, and have an aggregate frequency of 13% without noise compared to 57% with noise.[44]

The importance of noise for promoting leniency is also reflected in the post-experimental questionnaire. Many subjects reported cooperating following their partner's first defection because they assumed it was due to error.

Table 6. *Descriptive statistics of aggregate behavior, as well as aggregated MLE frequencies. All specifications use b/c=4. The descriptive statistics for leniency and forgiveness are defined in the text. For MLE aggregation, all strategies other than are ALLD and D-TFT are cooperative; lenient strategies are TF2T, TF3T, 2TF2T, Grim2 and Grim3; forgiving strategies are TFT, TF2T, TF3T, 2TFT and 2TF2T.*

	E=0	E=1/16	E=1/8
Descriptive statistics			
%C First Round	83%	87%	76%
%C All Rounds	78%	82%	59%
Leniency	42%	77%	66%
Forgiveness	19%	47%	32%
MLE aggregation			
Cooperative strategies	84%	89%	77%
Lenient strategies	40%	82%	63%
Forgiving strategies	53%	78%	57%

[43] See Appendix D for MLE results.

[44] Note that here we do find some evidence of longer-memory strategies, while our test based on adding only partner's play two periods ago found an insignificant effect. This may be in part due to the fact that the MLE includes strategies like "Grim" that have a longer memory, and in part to the fact that the no-noise case provides less information about play at many histories. The results are qualitatively equivalent when restricting our attention to the no-noise payoff specifications where TFT risk-dominates ALLD, 62% of strategies use memory at most 1, and lenient strategies have weight 20%.

V. Discussion

To relate play in the experiment to theoretical predictions, we would like to understand the extent to which the observed distribution of play approximates an equilibrium, and to the extent that play is not an equilibrium, what sorts of alternative strategies would perform better. To that end, we used simulations to compute the expected payoff matrix for the strategies that had non-negligible shares in the MLE estimation, along with a few "exploitive" strategies that struck us as good responses to the commonly used lenient strategies.[45] The resulting payoff matrices are displayed in Appendix E. We will use this table to compute the expected payoff to each strategy given the estimated frequencies, but first we use it to make some observations about the equilibria of the game. In particular, any strategy that is not a Nash equilibrium in this payoff matrix cannot be a Nash equilibrium in the full game. The converse is of course false, but we will then check which of the strategies that are equilibria of the payoff matrix also are equilibria of the full game.

Using that calculated payoff matrix we see that the lenient-and-forgiving strategy TF2T, which was common when b/c=2.5 or 4, is not an equilibrium in any treatment: it can be invaded by DC-Alt (the strategy that alternates between D and C) in all payoff specifications, as well as by ALLD when b/c=1.5 and by various "exploitive" strategies that start with D in the other treatments.[46] Note also that TFT is never an equilibrium, although it is fairly common when b/c=1.5: it is invaded by ALLD when b/c=1.5 and by ALLC (!) at other b/c values. This is a reflection of the fact that errors can move TFT into an inefficient 2-cycle.

Of course ALLD is an equilibrium in every treatment, and as discussed in Section II, PTFT is an equilibrium at b/c=4. Perhaps surprisingly, it turns out that Grim2 is also an equilibrium when b/c=4, even though it is not an equilibrium in the game without errors.[47]

---

[45] Analytic computations of the payoff for two different strategies playing each other is complicated due to the combination of discounting and noise, especially if the strategies look back more than one round and/or have many implicit "states."

[46] This is because the one round of punishment provided by TF2T, multiplied by the increased probability of punishment associated with the first D, is too small to outweigh the short-run gain to deviation. When b/c becomes sufficiently large, it does pay to conform to TF2T at histories where the strategy says to cooperate, but then it is also optimal to play C at histories where TF2T says to play D. Mathematica computations show that TF2T is not an equilibrium in any of our treatments.

[47] Without errors it would be better to play D in the first round and subsequently play Grim2, as the continuation payoff after one D is the same as after no D's at all. However in the presence of errors, the expected continuation payoff to Grim2 is lower in the round following a D, and numerical calculations show that when b/c=4 and $\delta=7/8$, Grim2 is an equilibrium provided that the error probability is between 0.0332 and 0.2778. Moreover, the exploitive

Moreover, the range of error probabilities for which Grim2 is an equilibrium increases with b/c. Intuitively, there is more reason to be lenient when the rewards to cooperation are greater, which is consistent with the way overall leniency in the data increases with b/c.

Of course these equilibrium calculations do not tell us what strategies can persist in a mixed-strategy equilibrium, and they do not tell us which strategies have good payoffs given the actual distribution of play. Table 7 shows the expected payoff of each strategy given the prevailing strategy frequencies. In the b/c=1.5 treatment, where ALLD is most prevalent, ALLD is also the best response to the prevailing strategy frequencies. Furthermore, the average earnings per round of subjects for whom ALLD is the strategy with the greatest likelihood are significantly higher than other subjects' earnings (coeff=0.108, p=0.028). ALLD does about as well as the average subject in the b/c=2 treatment (coeff=-0.039, p=0.672), and is significantly worse at b/c=2.5 (coeff=-0.400, p<0.001) and b/c=4 (coeff=-0.914, p<0.001).[48] We also see that subjects showed good judgment in avoiding PTFT, which performs very poorly in all treatments.

In the treatments with cooperative equilibria, lenient strategies perform very well. Within each of these treatments, the highest payoff strategy that is played is lenient (b/c=2, Grim2; b/c=2.5, 2TF2T; b/c=4, TF2T). Furthermore, all common lenient strategies (frequency of 10% or higher) earn within 1% of the highest payoff earned by any strategy played in that treatment, except for 2TF2T at b/c=2, which earns 1.6% less than the highest payoff. Various start-with-D strategies would have been the highest earners (b/c=2, D-Grim3; b/c=2.5 and b/c=4, D-TF2T), but these strategies were not played. Perhaps this is because these exploitive strategies do not fit well with subjects' intuitions and heuristics about cooperative play, and only out-performed the cooperative strategies by a very small margin (best cooperative strategy earns 5.3% less than best exploitative strategy at b/c=2, 0.9% less at b/c=2.5 and 0.2% less at b/c=4). Given the incomplete learning we observe in all treatments, it may therefore not be such a surprise that subjects did not discover the benefit of these exploitive strategies. Furthermore, given the roughly equal payoffs, subjects might reasonably prefer lenient cooperative strategies to those that exploit. Exploring the lack of exploitative strategies is an important direct for future work.

---

but lenient strategy D-Grim2 is also an equilibrium when b/c=4 and δ=7/8, although it is not used. See the online appendix for the Grim2 and D-Grim2 equilibrium calculations.

[48] We report the results of a linear regression over profit in all rounds of all interactions, with an ALLD dummy as the independent variable, clustered on both subject and interaction pair.

Table 7. *Observed frequencies and resulting expected payoffs for each strategy. Highest payoff strategy among those that were used is shown in bold; highest payoff strategy among all strategies considered is underlined.*

	b/c=1.5		b/c=2		b/c=2.5		b/c=4	
	Frequency	Expected payoff	Frequency	Expected payoff	Frequency	Expected payoff	Frequency	Expected payoff
ALLC		-1.25	0.03	6.92		13.27	0.06	28.13
TFT	0.19	2.40	0.06	8.71	0.09	14.64	0.07	29.01
TF2T	0.05	1.53		8.69	0.17	14.65	0.20	**29.67**
TF3T	0.01	0.90	0.03	8.44	0.05	14.53	0.09	29.56
2TFT	0.06	2.87	0.07	8.59	0.02	13.58	0.03	27.08
2TF2T		1.86	0.11	8.89	0.11	**14.72**	0.12	29.62
GRIM	0.14	3.02	0.07	8.40	0.11	12.33	0.04	23.99
GRIM2	0.06	2.37	0.18	**9.03**	0.02	13.98	0.05	27.90
GRIM3	0.06	1.79	0.28	9.02	0.24	14.67	0.11	29.23
ALLD	0.29	<u>**3.73**</u>	0.17	8.53	0.14	11.33	0.23	21.04
D-TFT	0.15	2.89		9.19	0.05	14.66		28.76
PTFT		0.72		6.34		12.05		25.36
D-TF2T		1.93		9.14		<u>14.87</u>		<u>29.73</u>
D-Grim3		2.34		<u>9.54</u>		14.83		28.92

Based on the expected payoffs in Table 7, perhaps the largest surprise is not the success of leniency and forgiveness, but rather the high proportion of subjects playing ALLD, particularly at b/c=4. The reason that low performing strategies such as ALLD can persist despite receiving low expected payoffs is probably that the complexity of the environment makes it difficult to learn the optimal response. Even though ALLD is not a best response to what people are really doing, ALLD *is* a best response to a belief that everyone else plays ALLD or any other history-independent strategy, and because of the noisy observation of intended play, subjects who have such false beliefs may not learn that more cooperative strategies yield a higher payoff. Consistent with this, 12% of subjects defected in more than 85% of all rounds in all interactions at b/c=2, 9% of subjects at b/c=2.5 and 16% of subjects at b/c=4. This accounts for a substantial fraction of the players classified as ALLD by the MLE, and suggests that these subjects almost never experimented with cooperation, preventing them from learning about its benefits. Furthermore, examining the play of these stubborn defectors, we find no positive correlation between first round cooperation and the previous partner's cooperation in the first round of the previous interaction; in fact the relationship is negative, although not significant (coeff=-0.081, p=0.882).[49] Thus meeting a first-round cooperator does not increase these subjects' probability of cooperating in future interactions. This is reminiscent of heterogeneous self-confirming equilibrium (Fudenberg and Levine 1993), and the diversity of strategies is consistent with heterogeneous self-confirming equilibrium in the absence of noise; in the presence of noise similar situations can persist for a while.[50] A similar logic applies to Grim, which is a best response to the belief that a substantial fraction of the population plays Grim while the rest plays ALLD - a subject who always uses Grim may not learn about the benefits of being more lenient.

We find no difference in first round cooperation between b/c=2 and b/c=2.5, and that the increase in overall cooperation as b/c increases from 1.5 to 2 is larger than the increase in moving from b/c=2 to b/c=2.5, even though ALLD risk-dominates TFT at b/c=2 but not b/c=2.5. Thus the risk dominance criterion has at best limited predictive power regarding cooperation in games with noise.

To explore possible non-strategic motivations for leniency and forgiveness, we examined the subjects' social preferences using a post-experiment dictator game and a set of survey questions from social psychology. In Dreber, Fudenberg and Rand (2010), we show that dictator giving is not correlated with cooperation in histories where there is the possibility of leniency, and not consistently correlated with cooperation in histories with the possibility for forgiveness. Dictator giving is also uncorrelated with both first-round cooperation and overall cooperation in the specifications with cooperative equilibria (where leniency and forgiveness are common). Furthermore, while lenient and forgiving strategies earn high expected monetary payoffs, the Ernst Fehr and Klaus M. Schmidt (1999) model of inequity aversion gives little utility to these strategies. We use this, the survey data and additional analysis to argue that social preferences do not seem to be a key factor in explaining the leniency and forgiveness observed in our experiments.

---

[49] When considering only subjects whose realized actions (as opposed to intended actions) resulted in over 85% D, we also find no correlation between previous partner's first round cooperation and own cooperation in the first round of the present interaction (coeff=-0.350, p=0.828).

[50] Evolutionary models such as the replicator dynamic, when applied to repeated games by restricting the strategy set can converge to steady states with multiple strategies present, as in Feldman and Thomas (1987). These polymorphic steady states, however, require that all of the active strategies obtain the same payoff, which is not a good approximation of the situation here.

VI. Conclusion

We conclude that subjects do tend to cooperate in noisy repeated games when there is a cooperative equilibrium, that they cooperate even when there are no cooperative equilibria in memory-1 strategies, and that they cooperate even when TFT is risk-dominated by ALLD. This shows that conclusions based on evolutionary game theory models that incorporate the memory-1 restriction need not apply to play in laboratory experiments, and that subjects can and do use strategies with more complexity. We also see that strategies such as TF2T that involve leniency and forgiveness are both common and rather successful in the sense of obtaining high payoffs given the actual distribution of play, even though it is not an equilibrium for all agents to play TF2T: In an uncertain world, it can be payoff-maximizing to be slow to anger and fast to forgive.

References

Abreu, Dilip, David Pearce, and Ennio Stachetti. 1990. "Towards a Theory of Discounted Repeated Games with Imperfect Monitoring." *Econometrica*, 58(5): 1041–1064.

Aoyagi, Masaki, and Guillaume Frechette. 2009. "Collusion as public monitoring becomes noisy: Experimental evidence." *Journal of Economic Theory*, 144(3): 1135–1165.

Aumann, Robert J., and Lloyd S. Shapley. 1994. "Long-term Competition - a Game Theoretic Analysis." *Essays in Game Theory in Honor of Michael Maschler*, edited by N. Megiddo, Springer, New York, 1994, pp. 1-15.

Axelrod, Robert. 1984. *The Evolution of Cooperation.* New York: Basic Books.

Axelrod, Robert, and William D. Hamilton. 1981. "The Evolution of Cooperation," *Science*, 211(4489): 1390–1396.

Binmore, Ken, and Larry Samuelson. 1992. "Evolutionary Stability in Repeated Games Played by Finite Automata." *Journal of Economic Theory*, 57(2): 278–305.

Blonski, Matthias, and Giancarlo Spagnolo. 2004. "Prisoners' other Dilemma." http://www.gianca.org/Index.htm.

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

Boyd, Robert. 1989. "Mistakes Allow Evolutionary Stability in the Repeated Prisoner's Dilemma." *Journal of Theoretical Biology*, 136(1): 47–56.

Boyd, Robert, and Jeffrey P. Lorberbaum. 1987. "No Pure Strategy is Stable in the Repeated prisoner's dilemma Game." *Nature*, 327: 58–59.

Camera, Gabriele, Marco Casari, and Maria Bigoni. 2010. "Cooperative Strategies in Groups of Strangers: An Experiment." Krannert School of Management Working Paper 1237.

Chamberlain, Gary. 1980. "Analysis of Covariance with Qualitative Data." *Review Economic Studies*, 47(1): 225–238.

Dal Bó, Pedro. 2005. "Cooperation Under the Shadow of the Future: Experimental Evidence from Infinitely Repeated Games." *American Economic Review*, 95(5): 1591–1604.

Dal Bó, Pedro, and Guillaume Frechette. 2011. "The Evolution of Cooperation in Infinitely Repeated Games: Experimental Evidence." *American Economic Review*, 101(1): 411–429.

Dreber Anna, Drew Fudenberg, and David G. Rand. 2010. "Who Cooperates in Repeated Games?" http://www.economics.harvard.edu/faculty/fudenberg.

Dreber Anna, David G. Rand, Drew Fudenberg, and Martin A. Nowak. 2008. "Winners Don't Punish." *Nature*, 452: 348–351.

Duffy, John, and Jack Ochs. 2009. "Cooperative Behavior and the Frequency of Social Interaction." *Games and Economic Behavior*, 66(2): 785–812.

Engle-Warnick, Jim, and Robert L. Slonim. 2006. "Inferring Repeated-Game Strategies from Actions: Evidence from Trust Game Experiments." *Economic Theory*, 28(3): 603–632.

Feldman, Marcus W., and Ewart A.C. Thomas. 1987. "Behavior-Dependent Contexts for Repeated Plays of the Prisoner's Dilemma II: Dynamical Aspects of the Evolution of Cooperation." *Journal of Theoretical Biology*, 128(3): 297–315.

Fehr, Ernst, and Klaus M. Schmidt. 1999. "A Theory of Fairness, Competition, and Cooperation." *Quarterly Journal of Economics*, 114(3): 817-868.

Feinberg, Robert M., and Thomas A. Husted. 1993. "An Experimental Test of Discount-Rate Effects on Collusive Behavior in Duopoly Markets." *Journal of Industrial Economics*, 41(2): 153–60.

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

Friedman, James W. 1971. "A Noncooperative Equilibrium for Supergames." *Review of Economic Studies*, 38(113): 1–12.

Fudenberg, Drew, and David K. Levine. 1993. "Self-Confirming Equilibrium." *Econometrica*, 61(3): 523–546.

Fudenberg, Drew, David K. Levine, and Eric Maskin. 1994. "The Folk Theorem with Imperfect Public Information." *Econometrica*, 62(5): 997–1040.

Fudenberg, Drew, and Eric Maskin. 1986. "The Folk Theorem in Repeated Games with Discounting or with Incomplete Information." *Econometrica*, 54(3): 533–556.

Fudenberg, Drew, and Eric Maskin. 1990. "Evolution and Cooperation in Noisy Repeated Games." *American Economic Review*, 80(2): 274–279.

Fudenberg, Drew, and Eric Maskin. 1994. "Evolution and Noisy Repeated Games." http://www.economics.harvard.edu/faculty/fudenberg.

Green, Edward J., and Robert H. Porter. 1984. "Noncooperative Collusion under Imperfect Price Information." *Econometrica* 52(1): 87–100.

Heckman, James J. 1981. "Structural Analysis of Discrete Data with Econometric Applications." In *The incidental parameters problem and the problem of initial conditions in estimating a discrete time-discrete data stochastic process,* MIT Press, Cambridge, MA, pp. 179–195.

Imhof, Lorens. A., Drew Fudenberg, and Martin A. Nowak. 2007. "Tit-for-Tat or Win-Stay, Lose-Shift?" *Journal of Theoretical Biology,* 247(3): 574–580.

Murnighan, J. Keith, and Alvin E. Roth. 1983. "Expecting Continued Play in Prisoner's Dilemma Games." *Journal of Conflict Resolution*, 27(2): 279–300.

Nowak, Martin A., and Karl Sigmund. 1990. "The Evolution of Stochastic Strategies in the Prisoner's Dilemma." *Acta Applicandae Mathematicae*, 20(3): 247–265.

Nowak, Martin A., and Karl Sigmund. 1993. "A Strategy of Win-Stay, Lose-Shift that Outperforms Tit-for-Tat in Prisoner's Dilemma." *Nature*, 364: 56–58.

Nowak, Martin A., Akira Sasaki, Christine Taylor, and Drew Fudenberg. 2004. "Emergence of Cooperation and Evolutionary Stability in Finite Populations." *Nature*, 428: 646–650.

Posch, Martin. 1999. "Win–Stay, Lose–Shift Strategies for Repeated Games—Memory Length, Aspiration Levels and Noise." *Journal of Theoretical Biology*, 198(2): 183–195.

Roth, Alvin E., and J. Keith Murnighan. 1978. "Equilibrium Behavior and Repeated Play of the prisoner's dilemma." *Journal of Mathematical Psychology*, 17(2): 189 –98.

Wedekind, Claus, and Manfred Milinski. 1996. "Human Cooperation in the Simultaneous and the Alternating Prisoner's Dilemma: Pavlov versus Generous Tit-for-Tat." *Proceedings of the National Academy of Sciences (USA)*, 93: 2686–2689.

Appendix A – MLE strategy frequency estimates using simulated data

Table A1. *Maximum likelihood estimates for simulated histories. For each b/c ratio, the first column shows the actual frequency in the simulated data, and the second column shows the MLE estimate. Bootstrapped standard errors shown in parentheses.*
*† Significant at p<0.1, * Significant at p<0.05, ** Significant at p<0.01*

	b/c=1.5		b/c=2		b/c=2.5		b/c=4	
	Actual	Estimated	Actual	Estimated	Actual	Estimated	Actual	Estimated
ALLC	0	0	0.03	0.02	0	0	0.05	0.05*
		(0)		(0.02)		(0)		(0.03)
TFT	0.18	0.18**	0.08	0.08**	0.08	0.08**	0.10	0.10**
		(0.05)		(0.04)		(0.04)		(0.04)
TF2T	0.03	0.03†	0	0	0.17	0.19**	0.18	0.18**
		(0.02)		(0)		(0.05)		(0.05)
TF3T	0	0	0.03	0.02	0.05	0.08*	0.08	0.04†
		(0)		(0.03)		(0.05)		(0.03)
2TF2T	0	0	0.10	0.10**	0.13	0.09*	0.12	0.13**
		(0)		(0.04)		(0.05)		(0.05)
Grim	0.22	0.22**	0.10	0.10**	0.12	0.12**	0.03	0.03†
		(0.05)		(0.04)		(0.04)		(0.02)
Grim2	0.03	0.03†	0.20	0.19**	0.02	0.02	0.07	0.07*
		(0.02)		(0.05)		(0.02)		(0.03)
Grim3	0.08	0.08**	0.28	0.31**	0.25	0.24**	0.12	0.14**
		(0.03)		(0.06)		(0.06)		(0.04)
PTFT	0	0	0	0	0	0	0	0
		(0)		(0)		(0)		(0)
2PTFT	0	0	0	0	0	0	0	0
		(0)		(0)		(0)		(0)
2TFT	0	0	0	0	0	0	0	0
		(0)		(0)		(0)		(0)
T2	0	0	0	0	0	0	0	0
		(0)		(0)		(0)		(0)
ALLD	0.28	0.28**	0.17	0.17**	0.13	0.13**	0.25	0.25**
		(0.06)		(0.06)		(0.05)		(0.06)
C-to-ALLD	0	0	0	0	0	0	0	0
		(0)		(0)		(0)		(0)
D-TFT	0.17	0.17**	0	0	0.05	0.05†	0	0
		(0.05)		(0)		(0.03)		(0)
D-TF2T	0	0	0	0	0	0	0	0
		(0)		(0)		(0)		(0)
D-TF3T	0	0	0	0	0	0	0	0
		(0)		(0)		(0)		(0)
D-Grim2	0	0	0	0	0	0	0	0
		(0)		(0)		(0)		(0)
D-Grim3	0	0	0	0	0	0	0	0
		(0)		(0)		(0)		(0)
DC-Alt	0	0	0	0	0	0	0	0
		(0)		(0)		(0)		(0)
Gamma		0.02		0.02		0		0.05**
		(0.01)		(0.01)		(0.01)		(0.01)

Appendix B – Evidence of learning

First, we investigate the extent of learning over the session for each payoff specification using E=1/8.[51] We do this by examining cooperation in the first round of each interaction (Figure 1), because this reflects each subject's strategy independent of the play of their current partner. There is no significant relationship between interaction number and first round cooperation when b/c=1.5 (coeff=0.006, p=0.788), a significant positive relationship when b/c=2 (coeff=0.089, p=0.001), and a non-significant relationship with a nonetheless rather sizable positive coefficient when b/c=2.5 (coeff=0.056 p=0.100) and b/c=4 (coeff=0.034, p=0.166). Examining learning at the individual level, we see a significant positive correlation between first round cooperation and the previous partner's cooperation in the first round of the previous interaction (coeff=0.335, p<0.001).[52] Thus cooperative partners tends to make one more cooperative, although the effect size is moderate (first round cooperation after meeting a defector, 65%; after meeting a cooperator, 72%). In the E=1/16 control, we find a similar relationship between first round cooperation and the previous partner's first round decision (coeff=0.513, p=0.047; after defection, 81%; after cooperation 88%), although there is no change in first round cooperation across interactions (coeff=0.005, p=0.898). For E=0, the opposite is true: there is no significant relationship between first round cooperation and the previous partner's opening move (coeff=-0.073, p=0.818), but there is a significant increase in cooperation across interactions (coeff=0.129, p=0.034). Thus there is also evidence of learning in both controls.

[51] We report the results from a logistic regression over all individual first round decisions, with the interaction number as the independent variable. To account for the non-independence of observations from a given subject, and from subjects within a given pairing, we clustered on both subject and interaction pair.

[52] The positive correlation between first round cooperation and the previous partner's cooperation in the first round of the previous interaction remains significant (coeff=0.291, p=0.001) when controlling for interaction number and b/c ratio, and we find no significant interaction either interaction number (coeff=-0.016, p=0.484) or b/c ratio (coeff=0.0003, p=0.998). Furthermore, we continue to observe this positive relationship when restricting our analysis to the last 4 rounds of each interaction (coeff=0.342, p=0.001). Thus the effect of meeting cooperative partners does not appear to vary across interaction or payoff specification.

Appendix C – MLE strategy frequencies using full 20 strategy set

Table A2. *Maximum likelihood estimates for the last 4 interactions of each session, all 20 strategies. Bootstrapped standard errors in parentheses.*
*† Significant at p<0.1, * Significant at p<0.05, ** Significant at p<0.01*

	b/c=1.5 E = 1/8	b/c=2 E = 1/8	b/c=2.5 E = 1/8	b/c=4 E = 1/8	b/c=4 E = 1/16	b/c=4 E = 0
ALLC	0 (0)	0.03 (0.03)	0 (0.02)	0.05† (0.03)	0 (0.05)	0.24* (0.10)
TFT	0.19** (0.05)	0.07 (0.04)	0.09* (0.04)	0.07* (0.03)	0.04 (0.03)	0.15* (0.07)
TF2T	0.05† (0.03)	0 (0)	0.16* (0.07)	0.19** (0.06)	0.22* (0.09)	0 (0.03)
TF3T	0.01 (0.01)	0.03 (0.03)	0.05 (0.04)	0.09* (0.04)	0.40** (0.11)	0 (0.03)
2TFT	0.06 (0.04)	0.07† (0.04)	0.02 (0.02)	0.03 (0.02)	0 (0)	0.16* (0.07)
2TF2T	0 (0.02)	0.11* (0.05)	0.11† (0.06)	0.12* (0.06)	0.08 (0.07)	0 (0)
Grim	0.14** (0.05)	0.05 (0.05)	0.11* (0.05)	0.02 (0.02)	0.02 (0.02)	0.12 (0.08)
Grim2	0.05† (0.03)	0.16* (0.07)	0.02 (0.03)	0.05† (0.03)	0.09† (0.05)	0.16* (0.08)
Grim3	0.06† (0.03)	0.27** (0.08)	0.24** (0.08)	0.11* (0.04)	0 (0)	0 (0.05)
PTFT	0 (0)	0 (0)	0 (0)	0 (0)	0.02 (0.02)	0 (0)
2PTFT	0 (0)	0.03 (0.03)	0 (0)	0 (0)	0.01 (0.02)	0 (0)
T2	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0.01)
ALLD	0.27** (0.05)	0.17** (0.06)	0.14** (0.05)	0.21** (0.04)	0.05 (0.03)	0.06† (0.03)
C-to-ALLD	0 (0.01)	0.01 (0.02)	0 (0)	0.01 (0.01)	0.02 (0.02)	0.02 (0.02)
D-TFT	0.10** (0.04)	0 (0)	0.03 (0.03)	0 (0)	0.02 (0.02)	0 (0.02)
D-TF2T	0 (0)	0 (0)	0.01 (0.02)	0 (0)	0.02 (0.01)	0 (0.03)
D-TF3T	0.01 (0.01)	0 (0)	0 (0)	0 (0)	0 (0)	0.04 (0.03)
D-Grim2	0.05† (0.03)	0.01 (0.01)	0 (0.01)	0.01 (0.01)	0 (0.01)	0 (0)
D-Grim3	0 (0)	0 (0)	0.01 (0.01)	0 (0)	0.03 (0.03)	0.04 (0.03)
DC-Alt	0 (0)	0 (0)	0 (0)	0.01 (0.01)	0 (0)	0 (0)
Gamma	*0.46*** *(0.02)*	*0.49*** *(0.03)*	*0.49*** *(0.03)*	*0.43*** *(0.02)*	*0.43*** *(0.02)*	*0.34*** *(0.02)*

Appendix D – MLE strategy frequencies for previous experiments without noise

Table A3. *Maximum likelihood estimates for Dal Bó and Frechette (2011), the two-option control games from Dreber et al. (2008) and a b/c=4 no-error control session. For δ=1/2, the last 16 interactions were analyzed, and for δ=3/4, the last 8 interactions. Bootstrapped standard errors in parentheses. Note that MLE results for our no-noise treatment is shown in the main text, Table 5.*
*† Significant at p<0.1, * Significant at p<0.05, ** Significant at p<0.01.*

	Dal Bo and Frechette 2011						Dreber et al. 2008	
Payoff	R=32	R=40	R=48	R=32	R=40	R=48	b/c=1.5	b/c=2
δ	0.5	0.5	0.5	0.75	0.75	0.75	0.75	0.75
ALLC	0	0	0.01	0	0	0.02	0	0
	(0)	(0.01)	(0.02)	(0)	(0.02)	(0.04)	(0)	(0)
TFT	0.07†	0.06	0.24**	0.23**	0.21†	0.55**	0.15*	0.40**
	(0.04)	(0.04)	(0.06)	(0.07)	(0.12)	(0.15)	(0.06)	(0.15)
TF2T	0	0.02	0.16*	0.11	0.22**	0	0	0
	(0)	(0.02)	(0.07)	(0.07)	(0.06)	(0)	(0)	(0)
TF3T	0	0	0.01	0	0	0.06	0	0
	(0)	(0.01)	(0.02)	(0)	(0.02)	(0.04)	(0)	(0)
2TFT	0	0.06	0	0	0.35**	0.09	0	0
	(0)	(0.04)	(0)	(0.02)	(0.13)	(0.09)	(0)	(0)
2TF2T	0	0	0	0	0	0	0	0
	(0)	(0)	(0.04)	(0.04)	(0.06)	(0)	(0)	(0)
Grim	0	0	0	0	0.04	0.2	0.07	0.21
	(0)	(0.02)	(0)	(0.02)	(0.06)	(0.12)	(0.04)	(0.16)
Grim2	0	0.01	0.02	0	0	0.02	0	0
	(0)	(0.01)	(0.02)	(0)	(0.01)	(0.04)	(0)	(0)
Grim3	0	0.01	0.02	0	0	0.06†	0	0
	(0)	(0.01)	(0.02)	(0)	(0.01)	(0.04)	(0)	(0)
ALLD	0.91**	0.76**	0.49**	0.66**	0.11*	0	0.64**	0.3**
	(0.04)	(0.06)	(0.07)	(0.07)	(0.05)	(0)	(0.1)	(0.11)
D-TFT	0.02	0.08†	0.04	0	0.08†	0	0.14	0.09
	(0.02)	(0.04)	(0.03)	(0)	(0.04)	(0)	(0.09)	(0.07)
Gamma	*0.34***	*0.49***	*0.4***	*0.45***	*0.32***	*0.28***	*0.36***	*0.42***
	(0.04)	*(0.04)*	*(0.04)*	*(0.04)*	*(0.03)*	*(0.03)*	*(0.03)*	*(0.03)*

Appendix E – Calculated payoff matrices

Figure A1. *Row player's payoff is shown, averaged over 10^5 randomly simulated games using c=1, e=1/8 and δ=7/8. Best response in each column is shown in bold.*

b/c=1.5	ALLC	TFT	TF2T	TF3T	2TFT	2TF2T	GRIM	GRIM2	GRIM3	ALLD	D-TFT	PTFT	D-TF2T	D-Grim3
ALLC	6.99	5.03	6.79	6.98	3.52	6.62	-4.17	3.05	6.05	-10.99	2.78	0.20	4.53	3.46
TFT	8.31	5.09	7.83	8.22	2.89	7.31	2.01	4.16	6.56	-1.81	2.86	4.70	6.33	4.39
TF2T	7.13	5.03	6.90	7.11	3.08	6.74	0.57	4.83	6.16	-3.97	2.80	2.76	4.72	3.44
TF3T	7.01	5.03	6.80	6.99	3.42	6.64	-0.55	4.41	6.38	-5.61	2.78	1.77	4.55	3.93
2TFT	9.32	5.13	7.90	8.88	3.62	6.89	2.70	5.05	6.03	-0.81	2.93	6.97	6.43	3.23
2TF2T	7.24	5.05	6.92	7.19	3.23	6.65	1.08	5.03	6.31	-3.19	2.81	3.97	4.76	3.72
GRIM	14.44	5.15	7.32	8.99	**4.13**	6.55	3.27	4.96	6.49	-0.66	2.94	8.45	5.72	4.07
GRIM2	9.63	5.10	6.96	7.59	3.73	6.68	2.15	5.69	6.32	-2.23	2.86	6.80	4.79	3.82
GRIM3	7.63	5.06	6.89	7.13	3.65	6.66	1.13	5.26	6.68	-3.65	2.82	5.31	4.65	4.33
ALLD	**18.99**	**5.22**	8.45	10.91	3.71	7.29	**3.48**	5.83	7.97	**1.00**	**2.97**	**10.75**	6.19	5.57
D-TFT	9.81	5.14	9.12	9.69	1.94	8.34	1.31	5.25	7.70	-0.32	2.90	4.89	7.05	**5.82**
PTFT	11.53	5.10	7.88	9.29	1.86	6.75	-0.18	2.10	4.23	-5.51	2.86	5.85	6.37	1.08
D-TF2T	8.64	5.06	8.26	8.60	1.70	7.98	-0.54	5.52	7.52	-2.46	2.84	2.57	6.07	4.94
D-Grim3	9.35	5.10	8.31	8.65	2.63	7.90	0.57	6.16	**8.06**	-2.05	2.84	6.29	6.05	5.81

b/c=2	ALLC	TFT	TF2T	TF3T	2TFT	2TF2T	GRIM	GRIM2	GRIM3	ALLD	D-TFT	PTFT	D-TF2T	D-Grim3
ALLC	14.00	**11.37**	13.71	13.97	**9.35**	13.50	-0.93	8.73	12.71	-9.99	**8.37**	4.95	10.71	9.31
TFT	15.31	10.18	14.53	15.17	6.68	13.70	5.28	8.67	12.51	-0.81	6.64	9.57	12.15	9.01
TF2T	14.13	11.19	13.81	14.10	8.09	13.54	3.83	10.52	12.62	-2.97	8.10	7.57	10.87	8.84
TF3T	14.01	11.34	13.71	13.98	9.01	13.50	2.73	10.14	13.07	-4.61	8.32	6.55	10.72	9.75
2TFT	16.31	9.37	13.86	15.58	7.23	12.30	5.96	9.59	11.12	0.19	5.46	11.91	10.98	6.22
2TF2T	14.23	11.00	13.78	14.17	7.92	13.30	4.36	10.69	12.78	-2.20	7.83	9.06	10.82	9.11
GRIM	21.47	9.03	11.95	14.18	7.68	10.90	**6.53**	8.78	10.83	0.33	5.23	13.46	8.95	6.73
GRIM2	16.62	9.83	13.08	13.91	7.99	12.68	5.43	11.38	12.20	-1.22	6.69	11.71	9.88	8.61
GRIM3	14.63	10.72	13.51	13.96	8.26	13.18	4.39	10.95	13.36	-2.65	7.59	10.18	10.45	10.12
ALLD	**25.99**	7.62	11.93	15.23	5.61	10.38	5.33	8.43	11.29	**2.00**	4.61	**15.00**	8.93	8.10
D-TFT	16.80	9.36	15.70	16.61	4.29	14.48	3.30	9.52	13.44	0.69	5.82	8.98	12.42	10.46
PTFT	18.52	10.03	13.71	15.57	5.75	12.39	3.10	6.11	8.92	-4.50	6.54	11.70	11.23	4.21
D-TF2T	15.64	10.64	15.10	15.57	5.29	14.67	1.43	10.73	13.87	-1.47	7.35	6.67	12.15	10.31
D-Grim3	16.33	9.91	14.66	15.42	5.51	14.12	2.53	11.38	**14.62**	-1.05	6.90	10.49	11.66	**11.62**

b/c=2.5	ALLC	TFT	TF2T	TF3T	2TFT	2TF2T	GRIM	GRIM2	GRIM3	ALLD	D-TFT	PTFT	D-TF2T	D-Grim3
ALLC	20.99	**17.71**	20.63	20.96	**15.19**	20.36	2.36	14.41	19.41	-8.99	**13.98**	9.68	16.88	15.10
TFT	22.31	15.27	21.25	22.11	10.48	20.09	8.56	13.22	18.48	0.19	10.42	14.43	17.95	13.70
TF2T	21.14	17.35	20.71	21.08	13.09	20.36	7.10	16.20	19.05	-1.96	13.42	12.41	17.00	14.23
TF3T	21.00	17.64	20.66	20.96	14.61	20.36	6.01	15.79	19.78	-3.61	13.86	11.31	16.89	15.57
2TFT	23.32	13.60	19.84	22.27	10.87	17.73	9.25	14.13	16.20	1.19	7.99	16.84	15.51	9.21
2TF2T	21.23	16.93	20.63	21.15	12.61	19.95	7.60	16.41	19.21	-1.20	12.85	14.15	16.87	14.52
GRIM	28.45	12.94	16.57	19.36	11.25	15.25	**9.80**	12.60	15.17	1.33	7.50	18.44	12.17	9.41
GRIM2	23.61	14.55	19.21	20.23	12.26	18.69	8.69	17.05	18.10	-0.23	10.50	16.63	14.95	13.35
GRIM3	21.62	16.38	20.09	20.79	12.86	19.72	7.63	16.63	20.04	-1.65	12.35	15.06	16.23	15.98
ALLD	**32.98**	10.04	15.40	19.53	7.52	13.49	7.15	11.05	14.60	**3.00**	6.28	**19.25**	11.66	10.62
D-TFT	23.80	13.59	22.28	23.53	6.61	20.61	5.26	13.82	19.17	1.69	8.70	13.06	17.78	15.11
PTFT	25.52	14.98	19.53	21.86	9.66	18.03	6.35	10.11	13.60	-3.50	10.23	17.55	16.07	7.42
D-TF2T	22.64	16.21	21.94	22.56	8.88	21.38	3.39	15.97	20.22	-0.47	11.88	10.77	18.22	15.71
D-Grim3	23.35	14.72	21.03	22.19	8.37	20.36	4.50	16.61	**21.23**	-0.05	10.94	14.69	17.28	**17.42**

b/c=4	ALLC	TFT	TF2T	TF3T	2TFT	2TF2T	GRIM	GRIM2	GRIM3	ALLD	D-TFT	PTFT	D-TF2T	D-Grim3
ALLC	41.99	**36.74**	41.40	41.91	**32.73**	40.97	12.21	31.44	39.43	-6.02	**30.75**	23.91	35.43	32.53
TFT	43.28	30.54	41.39	42.94	21.86	39.32	18.36	26.85	36.29	3.19	21.74	29.00	35.47	27.67
TF2T	42.11	35.82	41.40	42.01	28.08	40.78	16.95	33.25	38.47	1.04	29.42	26.79	35.44	30.37
TF3T	42.01	36.58	41.42	41.93	31.43	40.96	15.83	32.78	39.79	-0.62	30.52	25.66	35.42	32.98
2TFT	44.30	26.31	37.72	42.35	21.69	33.96	19.02	27.69	31.42	4.20	15.60	31.61	29.17	18.21
2TF2T	42.24	34.77	41.16	42.05	26.64	39.89	17.43	33.47	38.60	1.80	27.87	29.39	35.02	30.67
GRIM	49.44	24.61	30.43	34.91	21.91	28.36	**19.63**	24.14	28.21	4.32	14.39	33.42	21.80	17.40
GRIM2	44.63	28.74	37.55	39.15	25.06	36.79	18.53	**34.13**	35.78	2.76	21.94	31.42	30.20	27.70
GRIM3	42.63	33.34	39.90	41.25	26.68	39.27	17.49	33.68	40.11	1.35	26.64	29.70	33.63	33.44
ALLD	**53.99**	17.23	25.86	32.45	13.22	22.76	12.64	18.89	24.55	**6.00**	11.25	31.99	19.86	18.19
D-TFT	44.80	26.28	42.03	44.28	13.66	38.96	11.14	26.70	36.40	4.69	17.45	25.37	33.88	28.98
PTFT	46.52	29.77	37.02	40.72	21.30	34.95	16.14	22.07	27.55	-0.51	21.34	**35.09**	30.65	16.83
D-TF2T	43.61	32.94	42.46	43.49	19.68	41.47	9.31	31.65	39.34	2.54	25.40	23.10	36.46	31.81
D-Grim3	44.34	29.11	40.11	42.47	16.99	39.04	10.42	32.34	**40.93**	2.96	23.03	27.29	34.07	**34.84**

Online Appendix

Slow to Anger and Fast to Forgive:

Cooperation in an Uncertain World

Appendix 0-A: Sample instructions

Instructions:

Thank you for participating in this experiment.

Please read the following instructions carefully. If you have any questions, do not hesitate to ask us. Aside from this, no communication is allowed during the experiment.

This experiment is about decision making. You will be randomly matched with other people in the room. None of you will ever know the identity of the others. Everyone will receive a fixed show-up amount of $10 for participating in the experiment. In addition, you will be able to earn more money based on the decisions you and others make in the experiment. Everything will be paid to you in cash immediately after the experiment.

You will interact numerous times with different people. Based on the choices made by you and the other participants over the course of these interactions, you will receive between $0 and $30, in addition to the $10 show-up amount.

You begin the session with 50 units in your account. Units are then added and/or subtracted to that amount over the course of the session as described below. At the end of the session, the total number of units in your account will be converted into cash at an exchange rate of 30 units = $1.

The Session:

The session is divided into a series of interactions between you and other participants in the room.

In each interaction, you play a random number of rounds with another person. In each round you and the person you are interacting with can choose one of two options. Once the interaction ends, you get randomly re-matched with another person in the room to play another interaction.

The setup will now be explained in more detail.

The round

In each round of the experiment, the same two possible options are available to both you and the other person you interact with: A or B.

The payoffs of the options (in units)

Option	You will get	The other person will get
A:	−2	+8
B:	0	0

If your move is A then you will get −2 units, and the other person will get +8 units.

If you move is B then you will get 0 units, and the other person will get 0 units.

Calculation of your income in each round:

Your income in each round is the sum of two components:
· the number of units you get from the move you played
· the number of units you get from the move played by the other person.

Your round-total income for each possible action by you and the other player is thus

		Other person A	Other person B
You	A	+6	-2
	B	+8	0

For example:
If you play A and the other person plays A, you would both get +6 units.
If you play A and the other person plays B, you would get -2 units, and they would get +8 units.
If you play B and the other person plays A, you would get +8 units, and they would get -2 units.
If you play B and the other person plays B, you would both get 0 units.

Your income for each round will be calculated and presented to you on your computer screen.

The total number of units you have at the end of the session will determine how much money you earn, at an exchange rate of 30 units = $1.

Each round you must enter your choice within 30 seconds, or a random choice will be made.

A chance that the your choice is changed

There is a 7/8 probability that the move you choose actually occurs. But with probability 1/8, your move is changed to the opposite of what you picked. That is:

> When you choose A, there is a 7/8 chance that you will actually play A, and 1/8 chance that instead you play B. The same is true for the other player.
>
> When you choose B, there is a 7/8 chance that you will actually play B, and 1/8 chance that instead you play A. The same is true for the other player.

Both players are informed of the moves which actually occur. Neither player is informed of the move chosen by the other. Thus with 1/8 probability, an error in execution occurs, and you never know whether the other person's action was what they chose, or an error.

For example, if you choose A and the other player chooses B then:

- With probability (7/8)*(7/8)=0.766, no changes occur. You will both be told that your move is A and the other person's move is B. You will get -2 units, and the other player will get +8 units.

- With probability (7/8)*(1/8)=0.109, the other person's move is changed. You will both be told that your move is A and the other person's move is A. You both will get +6 units.

- With probability (1/8)*(7/8)=0.109, your move is changed. You will both be told that your move is B and the other person's move is B. You will both get +0 units.

- With probability (1/8)*(1/8)=0.016, both your move and the other person's moves are changed. You will both be told that your move is B and the other person's move is A. You will get +8 units and the other person will get -2 units.

Random number of rounds in each interaction

After each round, there is a 7/8 probability of another round, and 1/8 probability that the interaction will end. Successive rounds will occur with probability 7/8 each time, until the interaction ends (with probability 1/8 after each round). Once the interaction ends, you will be randomly re-matched with a different person in the room for another interaction. Each interaction has the same setup. You will play a number of such interactions with different people.

You will not be paired twice with the same person during the session, or with a person that was previously paired with someone that was paired with you, or with someone that was paired with someone that was paired with someone that was paired with you, and so on. Thus, the pairing is done in such a way that the decisions you make in one interaction cannot affect the decisions of the people you will be paired with later in the session.

Summary

To summarize, every interaction you have with another person in the experiment includes a random number of rounds. After every round, there is a 7/8 probability of another round. There will be a number of such interactions, and your behavior has no effect on the number of rounds or the number of interactions.

There is a 1/8 probability that the option you choose will not happen and the opposite option occurs instead, and the same is true for the person you interact with. You will be told which moves actually occur, but you will not know what move the other person actually chose.

At the beginning of the session, you have 50 units in your account. At the end of the session, you will receive $1 for every 30 units in your account.

You will now take a very short quiz to make sure you understand the setup.

The session will then begin with one practice round. This round will not count towards your final payoff.

Appendix O-B – Demographic statistics by session

b/c	error	Female	Economics	Age
1.5	1/8	52%	21%	20.9
1.5	1/8	48%	18%	21.0
1.5	1/8	50%	10%	19.6
Average		*50%*	*17%*	*20.5*
2	1/8	45%	14%	20.5
2	1/8	43%	10%	20.4
Average		*44%*	*12%*	*20.5*
2.5	1/8	38%	13%	21.3
2.5	1/8	64%	9%	19.5
2.5	1/8	52%	17%	20.7
Average		*52%*	*13%*	*20.4*
4	1/8	62%	17%	22.8
4	1/8	69%	6%	22.1
4	1/8	61%	13%	21.8
4	1/8	42%	30%	20.3
Average		*59%*	*16%*	*21.7*
4	1/16	35%	17%	22.7
4	1/16	44%	19%	21.0
4	1/16	69%	21%	21.2
Average		*47%*	*19%*	*21.6*
4	0	41%	6%	24.9
4	0	44%	6%	22.8
4	0	36%	8%	21.0
Average		*41%*	*7%*	*23.2*

Appendix O-C – Sequence of game lengths

b/c	e	t0†	t1	t2	t3	t4	t5	t6	t7	t8	t9	t10	t11	t12	t13	t14	t15
1.5	**1/8**	3	8	7	10	7	8	9	5	11	9	8	7	8	16		
1.5	**1/8**	3	8	7	10	7	8	9	5	11	9	8					
1.5	**1/8**	3	8	7	10	7	8	9	5	11	9	8					
2	**1/8**	3	8	7	10	7	8	9	5	11	9						
2	**1/8**	3	8	7	10	7	8	9	5	11	9	8	7	8	16	4	
2.5	**1/8**	3	8	7	10	7	9	10	5	11							
2.5	**1/8**	3	8	7	10	7	9	10	5	11	9	8	7				
2.5	**1/8**	3	8	7	10	7	8	9	5	11	9	8	7	8	11		
4	**1/8**	3	8	7	10	1	8	9	5	11	9	8	7	8			
4	**1/8**	3	8	7	10	8	9	6	11	7							
4	**1/8**	3	8	7	10	7	8	9	5	11	9	8	7	8	16	4	9
4	**1/8**	3	8	7	10	7	8	9	5	11	9	8					
4	**1/16**	3	8	7	10	7	8	9	5	11	9						
4	**1/16**	3	8	7	10	7	8	9	5	11	9	8	7	8			
4	**1/16**	3	8	7	10	7	8	9	5								
4	**0**	3	8	7	10	7	8	9	5	11	9						
4	**0**	3	8	7	10	7	8	9									
4	**0**	3	8	7	10	7	8	9	5	11	9						

† t0 was a practice round that did not count toward the players' earnings.

The sequence of games in each session is shown in the preceding table. The starting place in the sequence of random game lengths that was used in the experiment was picked by the programmer, and the sequence following the chosen starting place had an unusually low number of short games. Although the average probability over all rounds of the game continuing was not significantly different from 7/8 (t-test: $p>0.10$ for all sessions), the overall distribution of game lengths differed significantly from what would be expected using a geometric distribution (Chi^2 goodness of fit test, $p<0.05$ in all sessions).

This raises the concern that over the course of the session, subjects might have come to believe that the game was more likely to end in later rounds than in earlier ones, and adjusted their play accordingly. Particularly, we might expect that the tendency for cooperation to decrease over the course of an interaction would be greater in later interactions; or put differently, that the relationship between round number and cooperation will become more negative as interaction number increases.

It is not clear why such an effect would alter our findings, but nonetheless we check for evidence of this occurring. To do so, we ran a logistic regression with robust standard errors clustered on subject and session, including controls for b/c ratio and error rate. In addition to round number and interaction number as independent variables, we also add a [round X interaction] term. A significant negative coefficient on the [round X interaction] term would indicate that in later interactions, subjects are less likely to cooperate in later rounds, suggesting that after subjects have had time to learn that games are more likely to end in later rounds, they become more likely to defect in those later rounds. However, we find no evidence of a such relationship between round number and interaction number, (the coefficient for the [round X interaction] term in the regression is not significantly different from 0), either when considering all histories (coeff=-0.003, p=0.330), only considering histories where in the previous round both players cooperated (coeff=-0.004, p=0.418), only considering histories with the possibility of leniency (coeff=-0.004, p=0.698) or only considering histories with the possibility of forgiveness (coeff=-0.001, p=0.832). This suggests that increasing experience with the game length distribution did not affect subjects' probability to cooperate in later rounds, and in particular did not affect their levels of leniency or forgiveness.

A second complication with the game lengths is that due to technical difficulties with the computer software, the actual sequence of game lengths deviated somewhat from the pre-generated sequence in 4 out of the 18 sessions. We did not inform subjects of this error (which we were unaware of at the time) and they were most likely not aware of it either; no subject commented on the issue, and our experience is that subjects in our experiments do communicate with us when they are aware of software errors.

Appendix O-D – Equilibrium calculations

O-D.1 Equilibrium calculation for TFT with error

If both players use TFT then all histories fall into one of 4 phases defined by play in the previous round, while what happened 2 rounds ago doesn't matter either for current play or for continuation:

- P1: CC yesterday. Here the strategy says to play C.
- P2: CD yesterday. Here the strategy says to play D.
- P3: DC yesterday. Here the strategy says to play C.
- P4: DD yesterday. Here the strategy says to play D

Regardless of the current phase, next round's phase is P1 if both play C; P2 if player 1 plays C while player 2 plays D; P3 if player 1 plays D while player 2 plays C; and P4 if both play D.

Let the payoffs from following TFT in these phases be v_1, v_2, v_3, v_4 respectively, and let the error rate and discount factor be *e* and δ respectively. Then

$$v_1 = (1-e)^2(b - c + \delta v_1) + e(1-e)(b - c + \delta(v_2 + v_3)) + e^2\delta v_4$$

$$v_2 = e(1-e)(b - c + \delta(v_1 + v_4)) + (1-e)^2(b + \delta v_3) + e^2(-c + \delta v_2)$$

$$v_3 = e(1-e)(b - c + \delta(v_1 + v_4)) + (1-e)^2(-c + \delta v_2) + e^2(b + \delta v_3)$$

$$v_4 = e^2(b - c + \delta v_1) + e(1-e)(b - c + \delta(v_2 + v_3)) + (1-e)^2\delta v_4 .$$

For example, consider v_1. Here both players intend to play C. Thus with probability $(1-e)^2$, no errors occur, both players play C and remain in phase 1 for the next round; and therefore player 1 earns (b-c) now plus the value of v_1 discounted by δ. With probability e(1-e), player 1 (only) makes an error and accidentally plays D, shifting to phase 3 for the next round; therefore player 1 earns b today plus the value of v_3 discounted by δ. Also with probability e(1-e), player 2 (only) makes an error, exploiting player 1 and shifting to phase 2 for the next round; therefore player 1 earns -c today plus the value of v_2. Together this results in the 2nd term $e(1-e)(b-c+\delta(v_2+v_3))$. Finally, with probability e^2 both players error and play D, shifting to phase 4 in the next round; here both earn 0 today, plus the value of v_4 discounted by δ.

It is enough to consider deviations in P1 and P2 to show that TFT is never an equilibrium in the presence of noise; for TFT to be an equilibrium, it is necessary (but not sufficient) to have

$$v_1 \geq e(1-e)(b-c+\delta v_1)+(1-e)^2(b+\delta v_3)+e^2(-c+\delta v_2)+e(1-e)\delta v_4$$

and

$$v_2 \geq (1-e)^2(b-c+\delta v_1)+e(1-e)(b+\delta v_3)+e(1-e)(-c+\delta v_2)+e^2\delta v_4 .$$

However, these two conditions are mutually exclusive. Thus TFT is never an equilibrium. [53]

O-D.2 Equilibrium calculation for PTFT with error

If both players use PTFT then all histories fall into one of 2 phases:

- P1: CC or DD yesterday. Here the strategy says to play C, and what happened 2 days ago doesn't matter either for current play or for continuation. Next round's phase is P1 if both play C or both play D, else P2.
- P2: CD or DC yesterday. Here the strategy says to play D, and what happened 2 days ago again doesn't matter either for current play or for continuation. Next round's phase is P1 if both play C or both play D, else P2.

Let the payoffs from following PTFT in these phases be v_1, v_2

respectively, and let the error rate and discount factor be ***e*** and δ respectively. Then

$$v_1 = (1-e)^2(b-c)+e(1-e)(b-c)+\delta((1-e)^2+e^2)v_1+2e(1-e)v_2)$$

[53] Moreover, a forgetful version of TFT that forgets the current state and picks a new state randomly with some non-zero probability is also never an equilibrium. For TFT, forgetting is equivalent to increasing the error rate *e*, as follows. From the point of view of an individual TFT player, there are two states: s1 (the other played C last round) and s2 (the other played D last round). In state s1, TFT plays C with probability 1-*e* and play D with probability *e*; in state s2, TFT plays C with probability *e* and play D with probability 1-e. Now imagine that players have a probability 2*p* of forgetting the state and randomly picking a new state, such that with probability *p* a player switches state. When the state is forgotten in s1 (with probability *p*), the player gets 'confused' and switches to state s2, and therefore intends to defect. Thus a TFT player in state s1 plays C with probability (1-*p*)(1-*e*)+*ep* and plays D with probability *p*(1-*e*)+(1-*p*)*e*. This is equivalent to non-forgetful TFT with error rate *e*'=*e*+*p*-2*ep*. The same is true when considering a player that forgets in state s2, who plays C with probability *p*(1-*e*)+(1-*p*)*e* and plays D with probability (1-*p*)(1-*e*)+*ep*. As shown above, TFT is never an equilibrium regardless of the value of *e*. Therefore forgetful TFT is never an equilibrium.

$$v_2 = (1-e)^2\delta v_1 + e^2(b-c+\delta v_1) + e(1-e)((b-c)+2\delta v_2).$$

Following PTFT is clearly optimal in P2, so it is enough to check for a one-stage deviation in P1. Thus PTFT is an equilibrium if

$$v_1 \geq (1-e)^2(b+\delta v_2) + e^2(-c+\delta v_2) + e(1-e)(b-c+2\delta v_1)$$

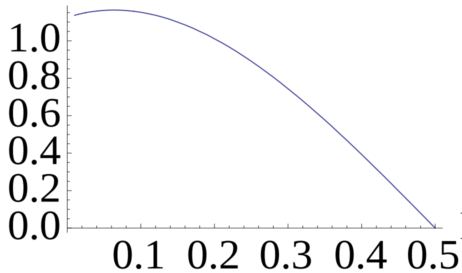
For *c*=2 and *δ*=7/8, we evaluate this expression for relevant values of *b/c* and *e*.

e	*b/c*	v_1	v_2	Phase 1 deviation payoff	Is PTFT an equilibrium?
1/8	1.5	5.85	5.1	6.98	No
1/8	2	11.7	10.2	12.46	No
1/8	2.5	17.55	15.3	17.94	No
1/8	4	35.11	30.61	34.39	Yes
1/16	4	40.69	35.44	38.93	Yes
0	4	48	42	44.75	Yes

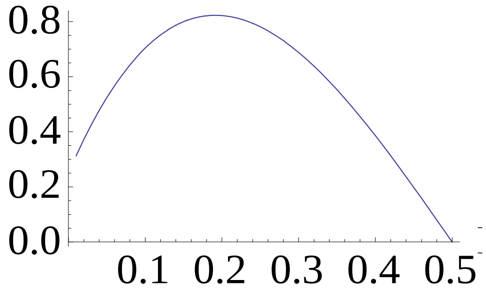
To more fully explore that range of *e* values over which PTFT is an equilibrium, we plot the payoff advantage of deviating in P1 as a function of *e*, for δ = 7/8, *c* = 2 and various *b*. We see that the for payoff specifications where PTFT is not an equilibrium at *e* = 1/8, increasing *e* does not lead to PTFT becoming an equilibrium. [54]

[54] A forgetful version of PTFT that forgets the current state and picks a new state randomly with some non-zero probability is also not an equilibrium at b/c=1.5, 2 or 2.5, as forgetting for PTFT is equivalent to increasing the error rate *e*, for similar reasons as for TFT. As shown above, increasing the error rate e cannot make PTFT an equilibrium for b/c=1.5, b/c=2 or b/c=2.5 with δ=7/8, and therefore forgetful PTFT is not an equilibrium for those value.

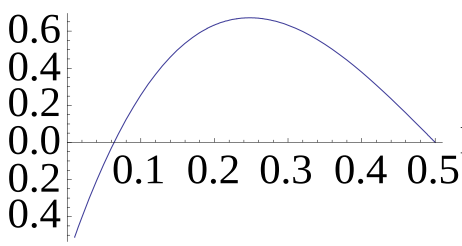
Advantage of Phase 1 Deviation, b.c 1.5



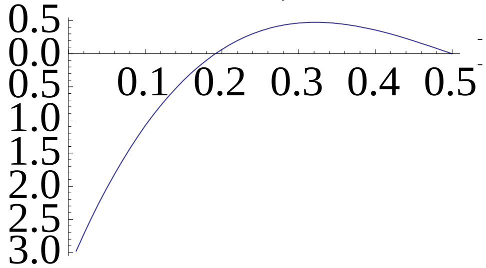
Advantage of Phase 1 Deviation, b.c 2



Advantage of Phase 1 Deviation, b.c 2.5



Advantage of Phase 1 Deviation, b.c 4



O-D.3 Payoffs of 2x2 game between TFT and ALLD and between Grim and ALLD with error probability e=1/8 and δ=7/8

TFT vs. ALLD

b/c = 1.5

	TFT	ALLD
TFT	5.09	-1.81
ALLD	5.23	1

b/c = 2

	TFT	ALLD
TFT	10.18	-0.81
ALLD	7.625	2

b/c = 2.5

	TFT	ALLD
TFT	15.27	0.19
ALLD	10.03	3

b/c = 4

	TFT	ALLD
TFT	30.55	3.19
ALLD	17.25	6

Grim vs. ALLD

b/c = 1.5

	Grim	ALLD
Grim	3.27	-0.66
ALLD	3.49	1

b/c = 2

	Grim	ALLD
Grim	6.54	0.34
ALLD	5.32	2

b/c = 2.5

	Grim	ALLD
Grim	9.82	1.34
ALLD	7.15	3

b/c = 4

	Grim	ALLD
Grim	19.63	4.34
ALLD	12.64	6

O-D.4 Payoffs of 2x2 game between TFT and ALLD and between Grim and ALLD with error probability e=1/16 and δ=7/8

TFT vs. ALLD

b/c = 1.5

	TFT	ALLD
TFT	5.87	-2.02
ALLD	4.27	0.50

b/c = 2

	TFT	ALLD
TFT	11.73	-1.52
ALLD	6.03	1.00

b/c = 2.5

	TFT	ALLD
TFT	17.60	-1.02
ALLD	7.79	1.50

b/c = 4

	TFT	ALLD
TFT	35.20	0.48
ALLD	13.06	3.00

Grim vs. ALLD

b/c = 1.5

	Grim	ALLD
Grim	4.29	-1.34
ALLD	3.27	0.50

b/c = 2

	Grim	ALLD
Grim	8.58	-0.84
ALLD	4.69	1.00

b/c = 2.5

	Grim	ALLD
Grim	12.87	-0.34
ALLD	6.11	1.50

b/c = 4

	Grim	ALLD
Grim	25.73	1.16
ALLD	10.38	3.00

O-D.5 Equilibrium calculation for Grim with error

If both players use Grim then all histories fall into one of 2 phases:

- P1: No D yesterday. Here the strategy says to play C, and what happened 2 days ago doesn't matter either for current play or for continuation. Next round's phase is P1 if both play C else P2
- P2: at least one D in the last round: play D. This phase is absorbing.

Let the payoffs from following Grim in these phases be v_1, v_2 respectively, and let the error rate and discount factor be *e* and δ respectively.

Then $v_1 = (1-e)^2(b-c) + e(1-e)(b-c) + \delta[(1-e)^2 v_1 + (2e - e^2)v_2]$,

$v_2 = e^2(b-c) + e(1-e)(b-c) + \delta v_2 \rightarrow v_2 = (e^2(b-c) + e(1-e)(b-c)) / (1-\delta)$

Following Grim is clearly optimal in P2, so it is enough to check for a one-stage deviation in P1. Thus Grim is an equilibrium if

$v_1 \geq (1-e)^2(b + \delta v_2) + e(1-e)(b - c + \delta v_1) + e(1-e)(\delta v_2) + e^2(-c + \delta v_2)$

For *c*=2 and δ=7/8, we evaluate this expression for relevant values of *b/c* and *e*.

e	*b/c*	v_1	v_2	Phase 1 deviation payoff	Is Grim an equilibrium?
1/8	1.5	3.27	1	3.47	No
1/8	2	6.54	2	5.43	Yes
1/8	2.5	9.82	3	7.40	Yes
1/8	4	19.63	6	13.30	Yes
1/16	4	25.73	3	11.17	Yes
0	4	48	0	8	Yes

O-D.6 Equilibrium calculation for Grim2 with error

If both players use Grim2 then all histories fall into one of 3 phases:

- P1: No D yesterday. Here the strategy says to play C, and what happened 2 days ago doesn't matter either for current play or for continuation. Next round's phase is P1 if both play C else P2
- P2: D yesterday but none the day before: play C. Next round's phase is P1 if both play C else P3.
- P3: at least one D in each of the last two rounds: play D. This phase is absorbing.

Let the payoffs from following Grim2 in these phases be v_1, v_2, v_3 respectively, and let the error rate and discount factor be *e* and *δ* respectively.

Then $v_1 = (1-e)^2(b-c) + e(1-e)(b-c) + \delta[(1-e)^2 v_1 + (2e-e^2)v_2]$,

$v_2 = (1-e)^2\ \ b - c + \delta v_1\ \ + e(1-e)(b-c) + (2e-e^2)\delta v_3]$

$v_3 = e^2(b-c) + e(1-e)(b-c) + \delta v_3 \rightarrow v_3 = (e^2(b-c) + e(1-e)(b-c))/(1-\delta)$

Following Grim2 is clearly optimal in P3, so it is enough to check for one-stage deviations in P1 and P2. Thus Grim2 is an equilibrium if

$v_1 \geq (1-e)^2(b+\delta v_2) + e(1-e)(b-c+\delta v_1) + e(1-e)(\delta v_2) + e^2(-c+\delta v_2)$ and

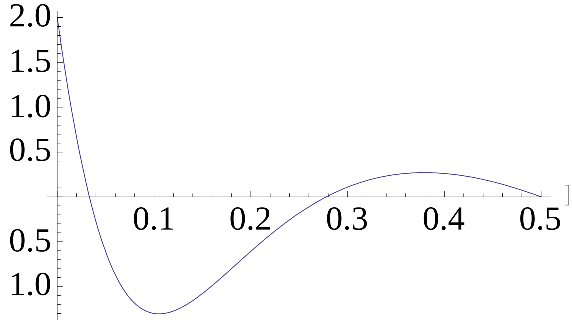
$v_2 \geq (1-e)^2(b+\delta v_3) + e(1-e)((b-c)+\delta v_1) + e(1-e)(\delta v_3) + e^2(-c+\delta v_3)$

For *c*=2 and *δ*=7/8, we evaluate this expression for relevant values of *b/c* and *e*.

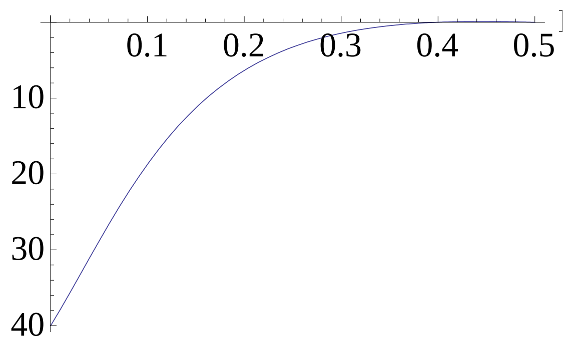
e	*b/c*	v_1	v_2	v_3	Phase 1 deviation payoff	Phase 2 deviation payoff	Is Grim2 an equilibrium?
1/8	1.5	5.69	4.89	1	6.73	3.70	No
1/8	2	11.38	9.78	2	11.96	5.90	No
1/8	2.5	17.07	14.68	3	17.20	8.10	No
1/8	4	34.14	29.35	6	32.89	14.69	Yes
1/16	4	41.85	38.12	3	40.92	11.99	Yes
0	4	48	48	0	50	8	No

To more fully explore that range of e values over which Grim2 is an equilibrium when b=8, c=2 and δ=7/8, we plot the payoff advantage of deviating in each state as a function of e. Numerical calculation shows that Grim2 is an equilibrium when $0.033 < e < 0.278$.

Advantage of Phase 1 Deviation



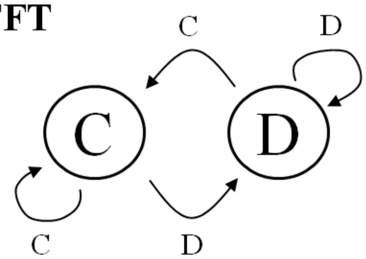
Advantage of Phase 2 Deviation



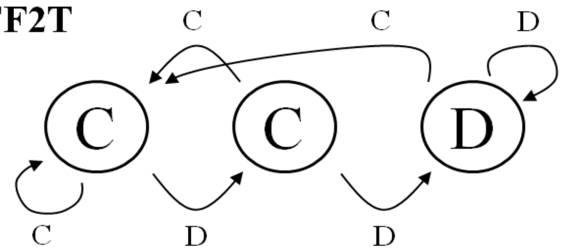
Here we define each strategy included in our analysis. Each phase is represented by a circle, with the strategy's move in that phase shown in the center of the circle, and transitions out of the phase indicated with arrows.

We begin with the strategies where transitions between phases depend only on the partner's move in the previous round: TFT, TF2T, TF3T, 2TFT, 2TF2T, D-TFT, D-TF2T and D-TF3T. For clarity we indicate only the partner's move with each transition arrow. All strategies begin in the leftmost phase.

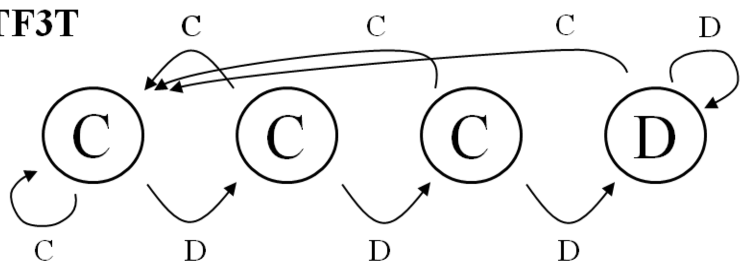
TFT



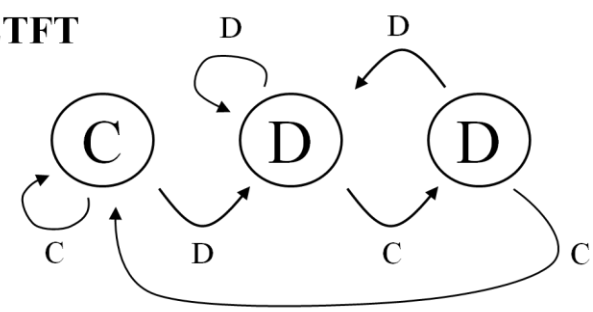
TF2T



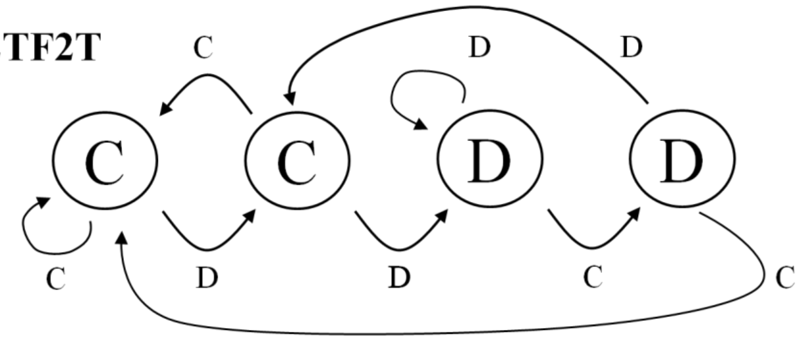
TF3T



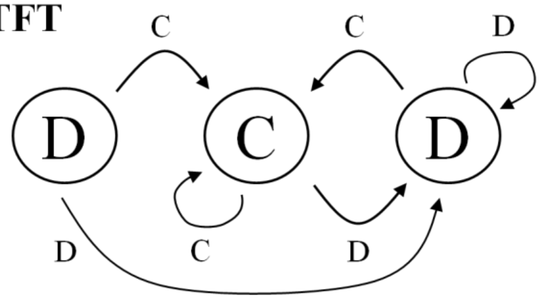
2TFT



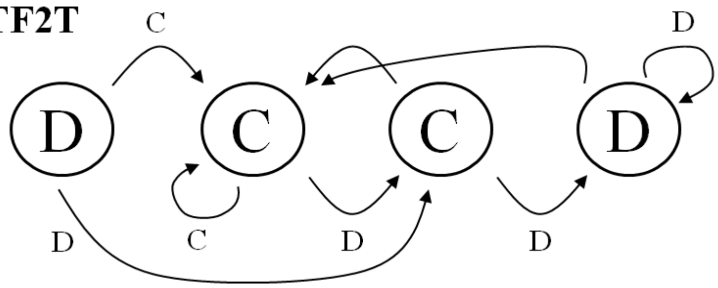
2TF2T



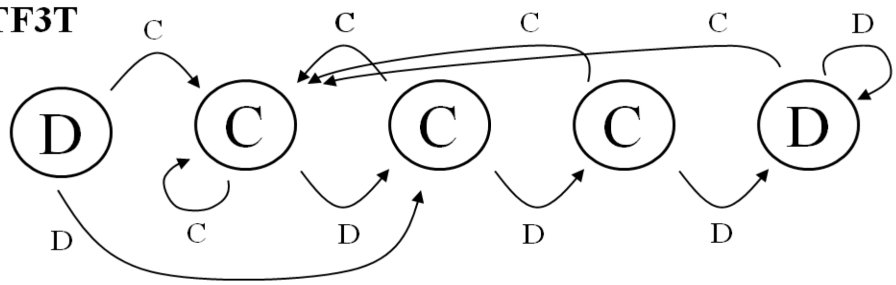
D-TFT



D-TF2T

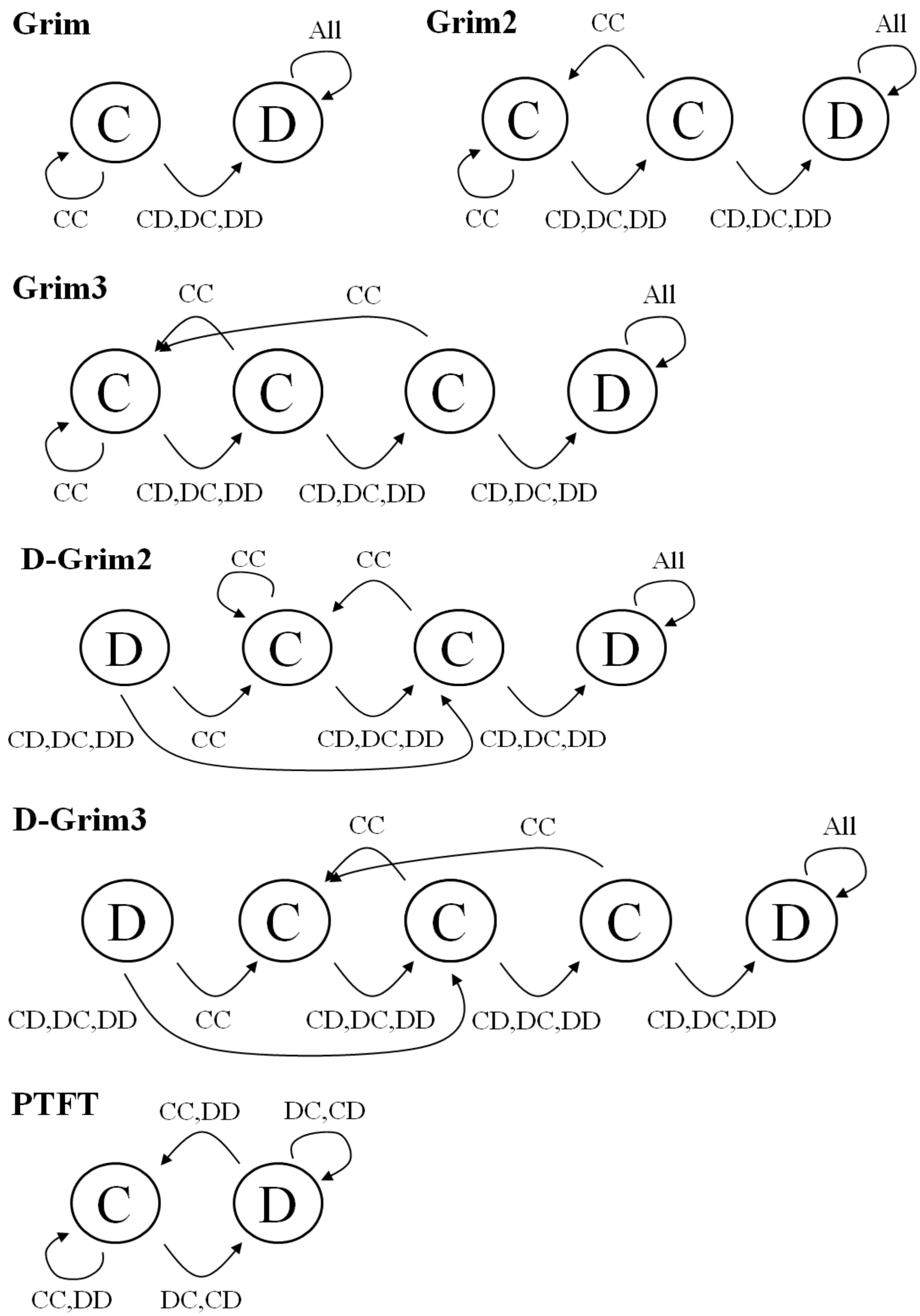


D-TF3T

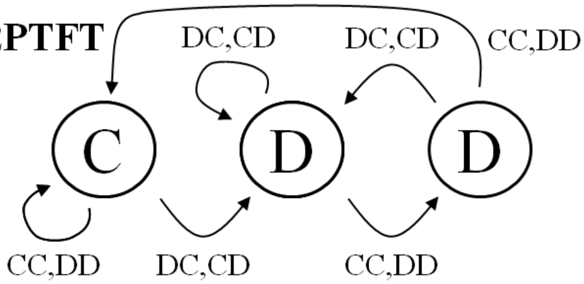


Next we define the strategies where transitions depend on both players' actions in the previous round: Grim, Grim2, Grim3, D-Grim2, D-Grim3, PTFT, 2PTFT and T2. The last round histories associated with each transition are indicated by the pair X_iX_j where X_i is the strategy's move last round and X_j is the partner's move last round (i.e. CD represents histories where the strategy played C last round while the partner played D).

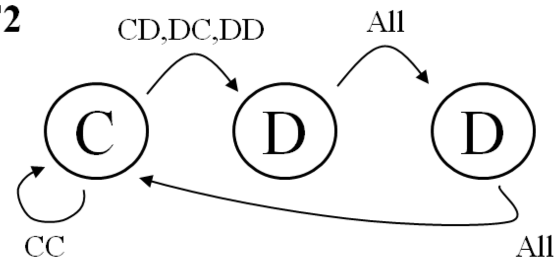
When only one transition out of a phase exists, irrespective of either player's action, the transition is labeled "All".



2PTFT

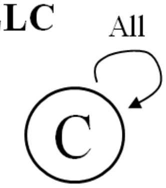


T2

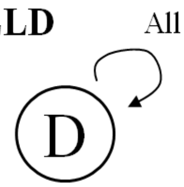


Lastly, we define the strategies whose transitions do not depend on previous histories of play: ALLC, ALLD, C-to-ALLD and DC-Alt.

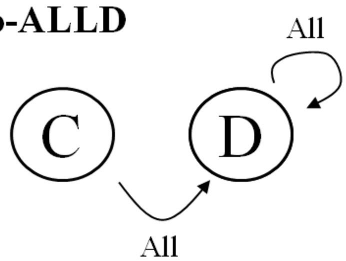
ALLC



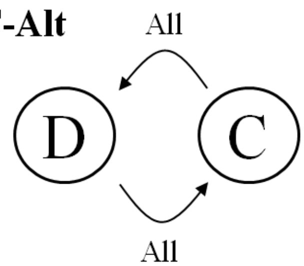
ALLD



C-to-ALLD



DC-Alt



Appendix O-F Robustness of strategy analysis

In the main text, we analyze the last 4 interactions of each session to minimize the effects of learning. Here we replicate our main analyses considering instead the last 2 or last 6 interactions, and find little difference, as shown in the table below. Regardless of the cutoff, we find that cooperation and leniency increase substantially going from b/c=1.5 to b/c=2, while forgiveness is changes little between b/c=1.5 and b/c=2, and then steadily increases from b/c=2 to b/c=4.

Last 2 Interactions	b/c=1.5	b/c=2	b/c=2.5	b/c=4
Descriptive statistics				
%C First Round	53%	79%	78%	77%
%C All Rounds	33%	45%	61%	64%
Leniency	30%	64%	67%	69%
Forgiveness	18%	16%	27%	45%
MLE aggregation				
Cooperative strategies	59%	84%	83%	78%
Lenient strategies	21%	53%	62%	61%
Forgiving strategies	27%	24%	41%	58%

Last 4 Interactions	b/c=1.5	b/c=2	b/c=2.5	b/c=4
Descriptive statistics				
%C First Round	54%	75%	79%	76%
%C All Rounds	32%	49%	61%	59%
Leniency	29%	63%	67%	66%
Forgiveness	15%	18%	33%	32%
MLE aggregation				
Cooperative strategies	57%	83%	81%	77%
Lenient strategies	18%	62%	60%	63%
Forgiving strategies	31%	31%	44%	57%

Last 6 Interactions	b/c=1.5	b/c=2	b/c=2.5	b/c=4
Descriptive statistics				
%C First Round	54%	75%	78%	75%
%C All Rounds	32%	49%	60%	58%
Leniency	30%	59%	68%	64%
Forgiveness	14%	19%	31%	32%
MLE aggregation				
Cooperative strategies	59%	82%	83%	77%
Lenient strategies	23%	65%	63%	64%
Forgiving strategies	36%	37%	46%	57%

Our MLE estimation procedure assumes that the probability of mental error in strategy implementation, parameterized by γ, is equal across strategies. It is possible, however, that some strategies are more difficult to implement than others and therefore γ may vary across strategies. Here we replicate our MLE estimates from Table 3, now allowing each strategy to have a different γ. We find little difference.

	b/c=1.5	b/c=2	b/c=2.5	b/c=4
	E=1/8	E=1/8	E=1/8	E=1/8
ALLC	0	0.0195	0.0486	0.062
TFT	0.1894	0.0668	0.0864	0.0601
TF2T	0.0473	0.019	0.2109	0.2137
TF3T	0.0122	0.0186	0.0343	0.0616
2TFT	0.069	0.0558	0.0234	0.0381
2TF2T	0.0048	0.0903	0.0887	0.1195
Grim	0.1551	0.1324	0.1054	0.0297
Grim2	0.0567	0.1122	0.0362	0.0902
Grim3	0.0429	0.3147	0.1685	0.1188
ALLD	0.2362	0.1408	0.1396	0.1669
D-TFT	0.1865	0.0299	0.0579	0.0396
γ-ALLC	*0.6517*	*0.3617*	*0.0438*	*0.2101*
γ-TFT	*0.4615*	*0.505*	*0.4034*	*0.3392*
γ-TF2T	*0.4329*	*3.8609*	*0.6046*	*0.4275*
γ-TF3T	*0.2886*	*0*	*0.4152*	*0.296*
γ-2TFT	*0.7757*	*0.4042*	*0.6942*	*0.5697*
γ-2TF2T	*0.4689*	*0.4188*	*0.419*	*0.3959*
γ-Grim	*0.4602*	*0.6609*	*0.6798*	*0.4077*
γ-Grim2	*0.4536*	*0.4062*	*0.6554*	*1.2917*
γ-Grim3	*0.4474*	*0.5662*	*0.408*	*0.4535*
γ-ALLD	*0.3165*	*0.2935*	*0.356*	*0.2086*
γ-D-TFT	*0.6247*	*1.4884*	*0.8278*	*10*
Cooperative strategies	58%	83%	80%	79%
Lenient strategies	16%	57%	59%	67%
Forgiving strategies	32%	27%	49%	56%

We now show that our MLE estimates are robust to including two alterative classes of simple strategies. Firstly, we consider forgetful memory-1 strategies (we noted above that these strategies are not equilibria at b/c=2 or b/c=2.5, but subjects might be playing them anyway). To include F-TFT and F-PTFT in the MLE, we take advantage of the fact that forgetting is equivalent to experiencing a higher error rate. Thus we add an additional parameter γ_F to the MLE which represents the additional probability of mental error for forgetful players (relative to non-forgetful players). The MLE terms for F-TFT and F-PTFT are therefore

$$p_i(s) = \prod_k \prod_r \left(\frac{1}{1 + \exp(-s_{ikr}(s) / (\gamma + \gamma_F))} \right)^{y_{ikr}} \left(\frac{1}{1 + \exp(s_{ikr}(s) / \gamma + \gamma_F)} \right)^{1-y_{ikr}}$$

where $p_i(s)$ is the likelihood of strategy s given the history of subject i, y_{ikr} is the actual decision of subject i in round r of interaction k (0=D, 1=C), s_{ikr} is the predicted move of strategy s (-1=D, 1=C) and γ parameterizes the mental error rate of non-forgetful strategies. A referee also suggested that subjects might be playing simple strategies which simply ignore their partner's first move before triggering (i.e. always cooperate in the first 2 interactions). To test for this, we also include the strategy C-TFT with always plays C for the first 2 periods then switches to TFT, and C-Grim with always plays C for the first 2 periods then switches to Grim. As in the main text, we estimate the frequency of each strategy using MLE and determine whether a strategy is present at frequency significantly greater than 0 using a t-test with bootstrapped standard errors.

As can be seen in the following table, we find little of any of these strategies. Bootstrapping standard errors shows that none are present at levels significantly greater than 0 ($p > 0.05$ for all 4 strategies in all 4 payoff specifications).

	b/c=1.5	b/c=2	b/c=2.5	b/c=4
ALLC	0	0.02	0.01	0.06
TFT	0.15	0.04	0.08	0.06
TF2T	0.05	0	0.15	0.17
TF3T	0.01	0.01	0.05	0.08†
2TF2T	0	0.11	0.11	0.12
Grim	0.13	0.02	0.07	0.01
Grim2	0.05	0.15	0.02	0.01
Grim3	0.05	0.28	0.23	0.12
PTFT	0	0	0	0
2PTFT	0	0.03	0	0
2TFT	0.03	0.07	0.02	0.03
T2	0	0	0	0
ALLD	0.27	0.17	0.14	0.18
C-to-ALLD	0	0.02	0	0.01
D-TFT	0.09	0	0.02	0
D-TF2T	0	0	0.02	0
D-TF3T	0.01	0	0	0
D-Grim2	0.05	0	0	0
D-Grim3	0	0	0.01	0
DC-Alt	0	0	0	0
C-TFT	0.03	0.03	0	0.02
C-Grim	0.01	0.04	0.03	0.03
F-TFT	0.07	0.03	0.06	0.04
F-PTFT	0	0	0	0.05
Gamma	*0.43*	*0.48*	*0.46*	*0.37*
Gamma_F	*0.6*	*2.17*	*1.09*	*1.92*

A referee also suggested the strategy that cooperates until the fraction of D moves by the partner passes some threshold, at which point it switches permanently to defection. To test for this possibility, we re-analyze the data including this family of strategies. We include 9 strategies which stop cooperating once the fraction of Ds by the partner is greater than 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% or 90%. We find that none of these strategies are present at a frequency significantly greater than 0 in any payoff specification ($p>0.05$ for all). The MLE results are shown in the following table:

	b/c=1.5	**b/c=2**	**b/c=2.5**	**b/c=4**
ALLC	0	0.012	0	0.04
TFT	0.1648	0.065	0.088	0.064
TF2T	0.0459	0	0.1719	0.194
TF3T	0.0099	0.038	0.0383	0.083
2TFT	0.0033	0.097	0.1038	0.093
2TF2T	0.126	0.066	0.1147	0.028
Grim	0.0497	0.187	0.0093	0.023
Grim2	0.0321	0.245	0.1997	0.065
Grim3	0.054	0.07	0	0.013
ALLD	0.2848	0.173	0.1407	0.227
D-TFT	0.1321	0	0.0478	0
FracD-10%	0	0	0	0.029
FracD-20%	0	0	0	0
FracD-30%	0.0273	0	0	0.016
FracD-40%	0.0224	0	0	0
FracD-50%	0	0.029	0.0859	0
FracD-60%	0.0446	0	0	0.077
FracD-70%	0.0031	0	0	0.048
FracD-80%	0	0	0	0
FracD-90%	0	0.017	0	0
Gamma	*0.453*	*0.497*	*0.49*	*0.423*

Appendix O-G MLE estimates including stochastically forgiving strategies

Our main analysis considers only deterministic strategies. Here we extend the MLE formulation to include mixed strategies. The original formulation is

$$p_i(s) = \prod_k \prod_r \left(\frac{1}{1+\exp(-s_{ikr}(s)/\gamma)} \right)^{y_{ikr}} \left(\frac{1}{1+\exp(s_{ikr}(s)/\gamma)} \right)^{1-y_{ikr}}$$

where $p_i(s)$ is the likelihood of strategy s given the history of subject i, y_{ikr} is the actual decision of subject i in round r of interaction k (0=D, 1=C), s_{ikr} is the predicted move of strategy s (-1=D, 1=C) and γ parameterizes the mental error rate. We replace this with a new formulation

$$p_i(s) = \prod_k \prod_r \left[s_{ikr} \left(\frac{1}{1+\exp(-1/\gamma)} \right) + (1-s_{ikr}) \left(\frac{1}{1+\exp(1/\gamma)} \right) \right]^{y_{ikr}} \cdot \left[(1-s_{ikr}) \left(\frac{1}{1+\exp(-1/\gamma)} \right) + s_{ikr} \left(\frac{1}{1+\exp(1/\gamma)} \right) \right]^{1-y_{ikr}}$$

where s_{ikr} now represents the probability that strategy s cooperates given the history preceding round r of interaction k (0 to 1).

We use this new formulation to consider stochastic conditional strategies. In particular we explore a family of ‘generous’ TFT (GTFT) strategies which have received significant attention in the evolutionary game theory literature (e.g. Nowak and Sigmund 1990). These strategies are stochastically forgiving. Like TFT, GTFT always responds to C with C. In response to an opponent’s D, however, GTFT stochastically cooperates with probability *q*.

First we analyze simulated data to see whether the MLE can differentiate between a GTFT that forgives defection 20% of the time and TF3T. We simulate a session with 10 ALLD players, 10 TF3T players, and 10 GTFT-2 players. We simulate 4 interactions, each lasting 8 rounds, and find that the MLE correctly assigns 1/3 to ALLD, TF3T and GTFT-2. Thus the MLE is able to distinguish between memory-1 stochastic forgiveness and longer deterministic forgiveness.

We now reanalyze our data using the 11 strategies from the main text Table 3 plus 9 GTFT variants- those which forgive defection with 10% probability (GTFT-1), 20% probability (GTFT-2), … 90% probability (GTFT-9). We find a somewhat sizable fraction of people playing stochastically forgiving memory 1 strategies (between 10 and 19% in the treatments with cooperative equilibria). However, the inclusion of these strategies doesn't undermine the importance of lenient strategies with more than 1 period of memory – the longer memory lenient strategies are much more common:

	b/c=1.5	**b/c=2**	**b/c=2.5**	**b/c=4**
GTFT-X	11%	12%	10%	19%
TF2T+TF3T+2TF2T+Grim2+Grim3	16%	57%	51%	43%

Thus we conclude that exploring stochastic strategies is an interesting avenue for future research, but that our main findings related to leniency are robust to including stochastic forgiveness.

Appendix O-H Logistic regression analyzing dependence on play 2 rounds ago

In the main text, we use a logistic regression to provide evidence that subjects' decisions are influenced by the partner's move two rounds ago as well as decisions in the previous round. A potential concern with this methodology lies in the possibility for such correlations to occur in heterogeneous populations of subjects with different strategies each of which conditions on at most play in the previous period. This is because the partner's decision two periods ago can give information about a subject's type even if it does not directly influence her decision. For example, consider a population of ½ ALLD and ½ TFT, and imagine that in the previous round both subject splayed C. If Player A played C two rounds ago then Player A is much more likely to be a TFT player than an ALLD player. This is because last round Player A played C in response to her partner's C two rounds ago, and this is consistent with TFT but not ALLD. Therefore Player A is also likely to play C now. If Player A's partner played D two rounds ago, however, then Player A is equally likely to be TFT or ALLD, because both strategies intend to play D in response to D. Therefore Player A is equally likely to play C or D in current round. And so if Player A's partner played C two rounds ago, she is more likely to play C now then if her partner played D last round, even though her strategy does not look back two periods.

Including controls for player type can help address this issue. For example, in a population of ALLD and TFT players, first round cooperation can do a good job of cleanly differentiating types. First round cooperation does not differentiate between TFT and GTFT, however, but GTFT players will have higher overall cooperation. Thus we include controls for both first round cooperation and overall cooperation in our regression.

To explore how pervasive of a problem bias stemming from heterogeneity might be for our analysis, and how effective our controls are for mitigating it, we conduct various simulations. We consider 5 different populations:

i. 1/3 ALLD, 1/3 TFT, 1/3 GTFT-5
ii. 1/3 ALLD, 1/3 TFT, 1/3 F-TFT (forgets the state with 10% probability)
iii. 1/2 ALLD, 1/2 TFT

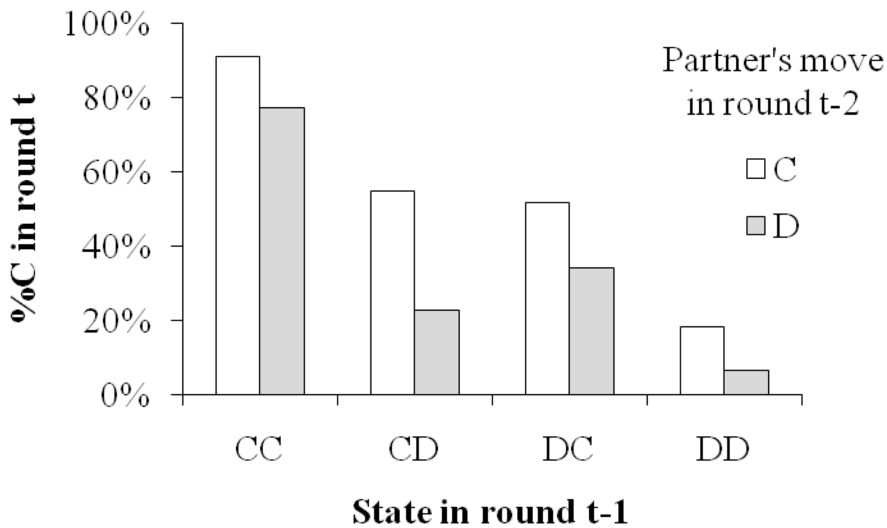
iv. 1/2 75%-TFT 25%-CoinFlip, 1/2 25%-TFT 75%-CoinFlip (on each decision, these strategies randomly choose to either play according to TFT or select a move at random (i.e. CoinFlip) – one half the population plays TFT 75% of the time, the other half picks randomly 75% the time)

v. 1/2 ALLD, 1/2 TF2T

Agents in populations (i), (ii), (iii) and (iv) do not condition on play two periods ago, while 1/2 of the agents in population (v) do. For each population, we simulate 3 sessions of 30 players each, playing 4 interactions each last 8 periods (for maximum comparability to our data). For each population we conduct 25 simulation replicates, and for each replicate we perform the logistic regression with correlated random effects described in the main text (own decision as a function of partner's move 2 rounds ago, own move 2 rounds ago, other's move last round, own move last round, own frequency of first round cooperation, and own frequency of overall cooperation). Across the 100 simulations using only memory 1 strategies, we find 4 instances in which other's play two rounds ago is significant at the $p<0.05$ level, consistent with what is expected due to chance. In contrast, we consistently find a highly significant effect of play two periods ago for every replicate of (v), where agents are in fact conditioning on play two rounds ago.

Examining the size of the coefficient for partner's move 2 rounds ago, not just the p-value, gives further insight into the relative importance of earlier play. We find that in all 4 scenarios where agents are not actually conditioning on play 2 rounds ago, the coefficient for partner's move 2 rounds ago is at least an order of magnitude smaller than the coefficients for play 1 round ago (usually several orders of magnitude smaller). In the scenario where many agents are actually conditioning on play 2 rounds ago, conversely, all coefficients are of the same order of magnitude. In our analysis of the data from our experiments, we find the latter result rather than the former: a large coefficient on partner's move 2 rounds ago, on the same order of magnitude as the other coefficients.
These results suggest that we are in fact picking up subjects explicitly conditioning on the outcome 2 periods ago, rather than only finding spurious correlations due to heterogeneity and stochastic strategies. Graphically displaying these relationships shows

a substantial and consistent effect of play two rounds ago, across all states in the previous round:



Appendix O-I Self-reported strategy descriptions

The post-experimental questionnaire included a free-response question in which subjects had the option of describing the strategy they used in the game. The responses of those subject who answered are reproduced here.

b/c=1.5, E=1/8

- I almost always started by choosing B and continued to play B for the subsequent rounds
- I chose B almost all the time because you are guaranteed not to lose any points with B
- I chose A to start with. If the other person also chose A, I switched to B the next round. If they too chose B, I stuck with it. If they chose A, I would make a 50/50 decision between A and B
- Once B was played by either player (including by me if my choice was switched) I played B for the rest of the interaction
- I chose B. I continued to chose B unless I felt the daring to press A. I chose B 98
- I'd start with B and continue if he played B. if he did A, I'd switch
- I started out choosing B, but if someone chose A twice I'd play A (I think). If someone was playing straight Bs, I'd occasionally play an A in an effort to guilt them into playing an A (at which point I'd have switched back to straight Bs)
- sometimes I started with B but rarely
- i started by choosing B. if the other person played A in the next round, i started playing A until the other person picked B in a later round. If, after that round, they chose A again, i continued to play A
- i chose B every time unless for two rounds in a row the other person played A
- i chose B most (if not all my rounds) because i wanted to maximize my points and minimize my losses
- random
- i chose A once then chose B the rest of the time
- i played A until they played B twice. Then i would switch to B. if they dont switch back to B, i would keep playing B or possibly do a one-off and switch to A (only if i have positive (or more than 2) points)
- i played B every chance i had because i know that the conservative strategy can win. However, i think i played A twice over the entire experiment because i was bored of pressing B
- always play A. switch to B if other player plays B more than 2 times in a row
- i basically always chose B
- always choose B. you cannot lose
- if it became clear we were going tit for tat with each other I’d try and break out of the cycle by playing A even if his or her last move were B

- i start by choosing B. if the other person plays A i will then choose A the next round. If the other person plays A twice in a row or every other time i then play A

b/c=2, E=1/8

- I chose B every time because it had the least risk
- start with A. always chose A unless the person goes B twice in a row
- you want to convey trustworthiness so when/if your response gets changed/changes you will remain with both As subsequently. I chose A until the other person played B 2x, unless it came in the middle of a string of As. I would return to A if my partner did
- I started with A. if I got B back, I would usually keep playing B unless they switched to A. if I got A back, I would usually play A, but slip in a few B's. hoping they would appear accidental.
- I played A until I thought the person was a B idiot, which was based on frequency not total amount of B's or A's chosen
- basically, I start with A and assume A from my opponent until otherwise noted. If opponent played B. I would switch to B until /unless my opponent went back to A indicating the B was a random switch by the comp from A.
- random plays: until the other player plays B. Then, play B afterwards.
- usually, I started by choosing B, then switched to A if the person played A twice in a row
- if the other person gave me 4 points and lost 2 points in the previous round I would try to do the same. if the other person consistently chose B I didn't want to chose A because I would only lose points and gain none. I would chose B after a round where I played A and the other played B, because in the previous round I gave up points to the other person so I expect them to do the same for me in this round
- AAABAABABB. always 3 As in the beginning unless the person played all Bs, in which case I did all Bs
- never played A
- along with A, build trust/rapport that would lead to best outcome for both
- wanted to avoid a B vs. B stalemate
- if I felt the round was ending and there was no incentive to further build the relationship, I might pick B. the other person might have chosen A, but his choice switched by chance. So to reciprocate his intensions, I might choose A
- I tried to use the tit for tat strategy to earn the most points in the long run. the first time the other player played B, I gave them the benefit of the doubt in that the computer had chosen for them.
- A makes the most sense. Both have incentive to gain whereas B only one person has incentive or both get nothing.
- lull them into thinking you were playing fairly (how barbaric)
- when I saw a string of "A" I felt I could trust them. String of "B"s, I didn't
- I would most likely play A for a few rounds as long as the other person was doing the same. Every now and then I would throw in a B to maximize my points. I would also play B if the other person had previously played B. I would never

choose B twice in a row because that would most likely lead to both persons choosing B for the remainder of the experiment. I would start each trial by choosing A. I would choose A until I decided to switch to gain extra points, or until the other person chose B. if that happened I would choose B the next round. If I chose B and the other person chose A, then I would always choose A the next round. rare exceptions: i would play B again only sometimes if the trial had been going for a while and i thought it would be the last round. if i ever chose B twice in a row, then i would choose B for the rest of the experiment unless it was really early and the other person played A twice

b/c=2.5, E=1/8

- my strategy was as follows: -start with A -after one B by other players, assume it was an error and continue with As -after two consecutive Bs by other player, assume he is a selfish jerk and stay with Bs for rest of interaction. Only when errors mixed up my responses and confused things did I have to become creative.
- I want to build trust with the other person. I wanted to give the other person the benefit of the doubt that his or her answer was changed by the computer.
- i didn’t want to get taken advantage of.
- i wouldn’t choose B after AA. After AB, I figured it could have been the glitch that made them possibly choose B 7/8 of the time - i wouldn’t choose B, I'd give the other person one more chance. I always chose A first, and if my action was B it was because the 7/8 glitch made me or the other person continually played B.
- i never chose A, i always chose B.
- benefit of the doubt: maybe he meant to do A so he doesn’t yet deserve retaliation
- i would play B after AA if i had a feeling the computer would change it to A.
- there was a possibility that the other person was picked A but B was what the result. everybody wanted to make as much money as they could and if they all picked B every round, nobody would get anything so it would be important to start the first few rounds with A that way I could let the other person know that if we cooperate, we can go a long way
- B could have been randomly thrown out and not their choice. to see if they randomly selected B and would throw more A choices and work cooperatively, I might choose A.
- choose A initially to earn trust so that I can eventually use B to earn more later.
- didn't really think about other person, but felt taking risk to gain money was better than getting $0
- basically I took into consideration karma. I would click A again just to see if it wasn't the 1/8 probability that made the other person choose B, and knowing this determined the rest of my responses.
- collusion in the long run in game theory can lead to a greater maximum of welfare instead of Nash equilibrium
- I figured that if the other person had been choosing A, they would choose it again and I'd get 3 units or my 1/8th probability of my choice coming out the opposite had not yet occurred so I took my chances that it would w/ A.

- if there was no reason to believe the other person chose B on purpose/would continue to choose B, I would choose A.
- I would cooperate as long as I thought the other person was. If I had done several (2-3) rounds of A I would then do B because it would look more like it was unintentional.
- It depended a lot on the person I played. one person chose B every time so I had to do the same in order to not get taken advantage of. Other times I would mostly press A and throw in a few Bs if the person was cooperative. X .
- I pretty much always played A and assumed if the person played B it was an error. If a person played B twice in a row I would sometimes play B in the next round.
- Gain trust by starting with A for 4 rounds then change to B.
- If they play B more than once, attribute it to intention not error and respond accordingly.
- I would usually choose B during the first round, then A for a few rounds, and then B once or twice later on.
- I would always start with A, and play A the second round. If the person played B both times, I switched to B. If the person played A at least 1 of those items, I would continue with A until the 6th or 7th round. then I'd play B.
- I punished for multiple B choices, not when it was just one and could have been accidental.
- I just wanted to make the most money per round possible.
- sometimes I chose A to establish it as a pattern so deviations were more likely to be interpreted as error.
- If they had more than 2 rounds with B, I went with B too, as little chance this was due to error. Or i went with B if their choice of B was greater than 1/8.
- played A consistently until other person played B three times, then switched to B until end of round.
- Played for a number of periods and realized that the average number of periods was between 7-10, so played B to maximize profit closer to the end.
- i acted fully in self-interest. However this is a PD so my best interest goes alongside what of my "partner". the interpretation of a B depends on how many rounds of A-A have occurred. If only 1, i will think it is more probable that it was intentional. If we have gone through A-A 5 times or more, i feel it is very unlikely it was intentional.
- i gave very little thought to the feelings of the other player at all. I played B every time to maximize my points and assumed they were doing the same so that any playing of A was unintentional.
- I rarely chose A b/c it earned you least points and seemed most high-risk. risk averse: playing B nets me more points and rarely costs me.
- I gave everyone the benefit of a doubt 1 round before during and after they played B. defensively if after 3 of their B rounds in a row I would [illeg] B as well.
- by choosing A, both players were in a mutually beneficial state, but only as long as they both kept choosing.
- I tried tit for 2 tat.

- some people played B all the time so I didn't want to fall into that perpetual cycle of losing. if the person had played B before, I probably didn't think it was chance.
- mostly tit for tat strategy with the understanding that a defect might have been by error and also hoping the other would assume my defection was an error when the other was a cooperator.

b/c=4, E=1/8

- I appreciated their choosing A (I would always assume they chose it). hoped a B was a 1 time thing on their part and so didn't want to sour future relations.
- I selected B at times because I doubted the trustworthiness of the other player. [if?] I decided to trust the other person would consistently select A
- depends on how many times the player has picked B.
- I thought their B may have been the 1/8 that the answer was switched. If so, they might come back to A in this round and we'd be even again. if they kept playing B, I'd play B so I wouldn't lose more points and so they wouldn't get more points
- If I were continuously playing A then they played B out of nowhere I'd play B the next round out of revenge. No need to be greedy. if they were to continue to play B then I would to, that way no one wins or gains.
- I would choose A unless the other person picked B. If the other person picked B, the next round I would also pick B. if the person then picked A I would think that it was an error and then continue to click A. if the person pressed B, i would continue to choose B. if I had accidentally picked B, i would pick A and make it up to the other person. then if they had chosen B i would assume they were checking me and continue to choose A.
- I never chose B if the previous round was both As. B could have been the error so I was willing to give the other person benefit of the doubt.
- if I chose A, and the other person chose B, the next time I chose B as a defense mechanism so I didn't lose anymore points
- in most of the interactions, if the other person did not play A at first, they eventually played A consistently after seeing that I had played A for the start two or more times in a row. Therefore, a mutual strategy seemed to be derived where both players would play A every time in order to maximize the number of units gained in the whole interaction (mutual altruistic response). Additionally, whenever this strategy was employed and the screen showed one time that the other person chose B, i automatically attributed it to error and continued to play A for the following rounds.
- trying to get equilibrium of all-As
- it felt terrible being betrayed after a long "A" streak. assume at first that their B could be a mistake. .
- if both played A, 12 total points earned. That's why I played it.
- keep the other person honest. Compensate for errors.
- when someone, myself or the other got B I stuck with B straight through.
- after I got a B from someone I always gave them the benefit of the doubt that it was due to error. After that if I got another B I would usually give B as well.

- there was a chance it was the computer if they chose A & I chose A, then we could both get six points, which outweighs the risk of losing 2 point, which is something like 7 cents. If they were consistently choosing B (for 3 rounds) then I would switch to B as well. However. I did so knowing we could both be better off with As. I always started with A for at least the 1st 2 rounds to see how much the other person wanted to cooperate. In rounds that we both began choosing B I never made as much money.
- I always played A unless they had played B.
- as the number of rounds increased. I was more likely to play B in the later rounds (expected number of rounds per interaction=8).
- B was the dominant strategy but gave better payoffs. I always chose B.
- since this game was completely anonymous and there was a chance that my choice would be changed, I felt no incentive to be nice and choose A .
- I started with A to demonstrate good faith and hoped that the other person would too so we could establish picking A. I had originally been skeptical and started with B, but then we both ended up choosing B for each round.
- if you are perceived as trustworthy, the interaction benefits both parties.

b/c=4, E=1/16

- Random chance might have produced their Bs so I want to try to possibly salvage the relationship. I would simply ignore occasional B's in a long string of A's
- The only way to earn money was to work together most of the time. I wasn't sure if B was the players intended decision.
- It could have been an error that B was played.
- Depends on if B was just one occurrence, or if they had played lots of Bs before
- I started with B in several rounds and stuck with it. The last few rounds I started with A.
- I would continue for a string of As, but if the other player played a B 2\3 times recently I would switch to Bs.
- didn't choose B at all during all interactions
- play A. If they go two Bs in a row play B (once or twice) then get back to A's
- if we both chose A, we typically stuck with that for the remainder of tee interaction.
- play A until other person chooses B X2, then play B and see what the outcome is. Play A next round etc.
- I never chose B on purpose, and to avoid getting into a locked system of B's I never played B.
- I chose A to begin with and then about every fourth time chose B to get more earnings. I then went back to A so as to blame the previous round on the1\16th chance of A change. If the other person played B twice in response, I chose B the rest of the time.
- Always stat with B. If other player is B, stay B until other player chooses A. If other player is A, stay A. If switch to B, give benefit of doubt due to 1\16 switch. If repeats, go to B.
- I always put A regardless of any choice given to me.

- I chose A for the whole interaction except for one B or if the other played B many times.
- I chose A until a few rounds passed. I then would throw in a B to make it seem as though the computer randomly chose.
- I played A as much as possible and then switched to B if they picked B twice in a row. Occasio.lly I'd try picking an A twice if we played for a long time at 0 to see if they wanted to actually switch to A.
- I started by choosing A, and continued to do so. If the other player chose B, I would choose B the next round. Then I would go back to A. I never chose B unprovoked, and if the computer generated that response for me, I would choose A the next round.
- I started by choosing A. Then I chose A always unless the other player chose B twice. If so, I chose B next round. If they chose B in that round, I chose B next. If they chose A in that round, I chose A next.
- I started choosing A until the other person choose B two times. Then I switched to B until the other person chose A. I immediately returned to choosing A as soon as the other person chose A. (A merciful strategy)
- I started play A 3 times to attempt to signal that I wanted to collaborate. If the other did not switch from B, I would change to B to attempt to change his/her play to A.
- If we got on a spree of both A's I'd pick B to get the +8 and then switch back. It eventually didn't really affect my earnings considerably.
- Started choosing B and then switched to choosing A or B once I ended up with 0 continuously.

b/c=4, E=0

- If they were helping me, I was willing to help them, 1\16 of a dollar isn't a lot to lose. If they chose B first, I felt like they were being greedy and I wanted to show them that strategy wouldn't pay off in the long run
- Always B.
- I hate when things don't make sense. To choose B in after we both chose A would hurt me because the other player would stop choosing A and I'd earn no points. Whether my motives are self-serving or altruistic, A is the most logical. If they've previously chosen B many times, I don't like them, and I don't want them to earn any points
- first I chose A. Then I started with B after seeing the first interaction, I would choose either A or B.
- going for B was not helping with long term sustained benefit.
- If the other person did A I would follow with A
- Both players win if both pick A, no one likes negative points
- If it were one of the middle rounds, I would definitely choose A. If one of the later rounds, most likely B.
- Played A every time. Only had 1 person one time choose B. Played A next round to give them a chance. I think if they had chosen B again I might have considered switching to B but my plan was to always play A.

- I chose A to take more money, collectively, from the university I would prefer another student have it that the university.
- Choosing anything other than A 100% of the time is just mean and vindictive. Its just a study why not help people get rich?
- I played A 3 to 4 times after which I only played B
- I started by choosing A and switched to B if the other player played B in the previous round, but would remain on B as long as the other player did too.
- I started by choosing B. then if the other person chose A in the next round I would choose A. If the other person chose B too, then I would continue playing B.