Turning the Other Cheek:

Leniency and Forgiveness in an Uncertain World

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

First version: March 24, 2010

This version: May 21, 2010

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. Furthermore, cooperative strategies yielded higher payoffs than uncooperative strategies in the treatments with cooperative equilibria, and successful strategies were "lenient" in not retaliating for the first defection and/or "forgiving" in trying to return to cooperation after inflicting a punishment. In all settings there was considerable strategic diversity, indicating that subjects had not fully learned the distribution of play. There was no difference across treatments in giving in a post-experiment dictator game, and dictator giving was not correlated with cooperation in the treatments with high returns on cooperation. We conclude that cooperation in settings with cooperative equilibria is primarily driven by strategic considerations.

† Lead authors

* Department of Economics, Harvard University

** Program for Evolutionary Dynamics, Berkman Center for Internet & Society, Harvard University

*** Institute for Financial Research (SIFR), Stockholm and Women and Public Policy Program, Harvard Kennedy School.

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, Stephanie Hurder, Magnus Johannesson, Martin Nowak, Parag Pathak, and John Scholz for helpful conversations and comments. Research support was provided by National Science Foundation grant 0646816 and the Dean of the Faculty of Arts and Sciences.

1. Introduction

Repeated games with observed actions have a great many equilibrium outcomes when players are patient, as shown by the 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. Axelrod [1984], Dal Bó [2005], Dreber et al [2008], Dal Bó and Frechette [2009], Duffy and Ochs [2009]) suggest that in these games players do indeed tend to cooperate when there is a cooperative equilibrium, at least if the gains to cooperation are sufficiently large,[2] and there are evolutionary arguments for why this should be the case.[3]

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. The folk theorem also applies to repeated games with this sort of imperfectly observed actions (Fudenberg, Levine and Maskin [1994]), and the evolutionary arguments for cooperative equilibria are even stronger here, 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. To test whether cooperation in the repeated game is driven by strategic or altruistic considerations, we also investigate how play correlates with giving in a dictator

---

[1] Aumann and Shapley [1976], Friedman [1971], Fudenberg and Maskin [1986], Fudenberg, Maskin and Levine [1994].

[2] Dal Bó and Frechette [2009] find that there need not be cooperation when the gain from cooperation is small. Earlier papers had found less cooperative play, 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] See e.g. Axelrod and Hamilton [1991] and Nowak et al [2004].

[4] Binmore and Samuelson [1992] and Fudenberg and Maskin [1990], [1994]. Nowak and Sigmund [1992], and Imhof, Fudenberg and Nowak [2007] also study evolutionary stability in repeated games with noise but they consider a restricted strategy space that rules out grim strategies.

game and with responses to survey questions related to social preferences in general, and to motivations of specific cooperative decisions in our experimental setting.

We present evidence from four different payoff specifications for a repeated prisoner's dilemma, with stage game actions "Cooperate" ("C") and "Defect" ("D") (more neutral language was used in the experiment itself). The four specifications all had the same continuation probability (7/8) and error rate (1/8); the change was in the benefit that playing "C" gives to the other player. Consistent with past work, we find that there is much more cooperation in specifications where cooperation is an equilibrium. 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 past experiments on the prisoner's dilemma without noise, which we review in Section 3, subjects were markedly more lenient (slower to resort to punishment). For example, in the payoff treatments that support cooperation, we find that players played C in 63-67% of the rounds immediately following their partner's first defection, compared to only 13%-40% in the cooperative treatments of Dal Bó and Frechette [2009]. Subjects also showed a considerable level of forgiveness (willingness to give cooperation a second chance), though less than in experiments without noise.

In addition to such descriptive statistics, we more explicitly explore which strategies our subjects used, as in past work by Wedekind and Milinski [1996], Aoyagi and Frechette [2009], and Dal Bó and Frechette [2009] on the repeated prisoner's dilemma as well as Engle-Warnick and Slonim [2006] who studied play 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.

We find considerable strategic diversity in all settings. Relatively few subjects used the strategy "Tit-for-Tat" (TFT) which was prevalent in the Dal Bo and Frechette [2009] no-noise setting, 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 period's outcome, as is commonly assumed in that literature. [5] 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) and modified, lenient versions of the Grim strategy, which wait for two or three defections before abandoning cooperation. These results show that subjects can and do use strategies with more than one period memory, at least in games with noise.

It is not an equilibrium for all subjects to play Tit-for-2-Tats, as in such a population any one subject would do better to play D on the first round. However, it turns out that the strategy "Grim 2"- switch to D forever following two consecutive periods of at least one D by either player, is an equilibrium in the treatment with the highest benefit to cooperation. Moreover, in this treatment, TF2T was the most common cooperative strategy, and had the highest payoff among the strategies that seem to be used, while in the other cooperative treatments the most common strategy, and one of the highest scorers, was the lenient but unforgiving "Grim 3."

The diversity of strategies used is consistent with learning being incomplete: there is "too much" cooperation in the treatment where Always Defect earns the highest payoff, and too much "always defect" in treatments where cooperative strategies do best. This explanation is consistent with the fact that there was only mild evidence of learning over the course of the experiment.[6]

An alternative explanation for this diversity 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

[5] The explicit analysis of evolutionary dynamics becomes quite difficult when longer memories are possible.

[6] We used a continuation probability of 7/8, instead of the ½ and ¾ in Dal Bó and Frechette, to investigate the extent to which players condition on observations before the previous period. When the continuation probability is ½, many interactions will last three or fewer periods, which makes it hard to study how far back players look in choosing their actions. One consequence of the higher continuation probability is that each subject played fewer repeated games (between 8 and 15) and so had less of an opportunity to learn.

later experimental session.[7] Because giving in the dictator game did not vary with the treatment, while cooperation did, it is clear that variations in social preferences did not cause the different levels of cooperation we observed. To investigate the relationship in more detail, we analyzed the correlation between a subject's giving in the dictator game and two measures of cooperative play. We find that giving is not correlated with playing C in the first period of the repeated game, except perhaps in the treatment where cooperation is not an equilibrium. We do find some evidence that giving is correlated with the overall frequency of cooperation, at least in the two lowest payoff treatments. This suggests that when cooperation is not payoff maximizing, the altruists (who give in the dictator game) are more likely to open with cooperation, and to be lenient and/or forgiving; but when the selfish payoffs strongly support cooperation, altruism does not seem to be a key explanatory variable.

We also asked subjects to rate self-interested versus altruistic motivations for being lenient (i.e. cooperating following a period in which they had cooperated while their partner had defected). Across all payoff specifications, a large majority of cooperative subjects reported maximizing their long-term payoff as a more important motivator of leniency than either a desire to increase their partner's payoff, to do the morally right thing or to avoid upsetting their partner. The desire to maximize payoff was a stronger motivation for cooperation in the payoff specifications which had higher cooperation rates, while other motivations were either unrelated to cooperation or constant across specifications. We also find that women are less likely to cooperate in the first round, and cooperate less overall.

2. 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 a known constant error probability $E=1/8$ that an

[7] Overall giving in the dictator game is roughly constant across treatments and comparable to that found in past studies (average donation: b/c=1.5, 18.7%; b/c=2, 16.7%; b/c=2.5, 16%, b/c=4, 19.5% - ranksum p>0.10 for all pairwise comparisons), which suggests that the preceding repeated game had little effect.

intended move is changed to the opposite move. Subjects are 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 are only notified of the other player's actual move, not the other's intended move. Subjects are 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 Table 1 where the payoffs are denoted in points. Cooperation and defection take the "benefit-cost" form, where cooperation means paying a cost c to give a benefit b to the other player; b/c took the values 1.5, 2, 2.5, and 4 in our four different treatments. Defection is passive, and means 0 units for both parties. Subjects were presented with both the benefit/cost representation and the resulting pre-error payoff matrix, in neutral language (the choices were labeled option A and option B as opposed to the "C vs. D" choice that is standard in the prisoner's dilemma). We use 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 a $6 dictator game (see below). To allow for negative stage-game payoffs, subjects began the session with an "endowment" of 50 units (in addition to the show–up fee).[8] On average subjects made $22 per session, with a range from $14 to $36. Sessions lasted approximately 90 minutes. [9]

[Tables 1 and 2]

Comment [DF1]: Table 2 is the expected payoff matrix. We want it in the paper but since it is large I'm not sure if we want it in the text- we might, I;m not sure.

A total of 278 subjects voluntarily participated at the Harvard Decision Science Laboratory in Cambridge, MA. In each session, 16-32 subjects interacted anonymously via computer using the software Z-Tree (Fischbacher [2007]) in a sequence of infinitely repeated Prisoner's Dilemmas (see Table 2 above for summary statistics on the different treatments). We conducted a total of 12 sessions between September 2009 and February 2010. 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 use the turnpike protocol as in Dal Bó [2005]. Subjects are divided into two groups, A and B. A-subjects only interact with B-subjects and vice versa. This implies a

[8] No subject ever had fewer than 19 units, and only 4 out of 278 subjects ever dropped below 40 units.
[9] 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.

matching where no subject will ever play twice with another subject, or with a subject who has played with a subject they have played with. Thus, this setup eliminates possible contamination effects: subjects cannot influence the play of future subjects they interact with. Subjects were informed about this setup. The maximum number of interactions in a session with N subjects is thus N/2. 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. [10]

Following the end of the series of repeated Prisoner's Dilemmas, subjects played a dictator game where they were asked to divide $6 between themselves and an anonymous recipient that was not a participant in the PD but would be recruited at a later date. Subjects were informed that the recipient would receive no payment other than what the subject chose to give. Lastly, subjects answered survey questions (included in the online appendix) related to motivations for specific cooperative decisions in our experimental setting, the types of strategies they have used, and demographics.

3. Theoretical and Experimental Background

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 "Always Defect," (ALLD) while the other treatments all allow cooperative equilibria.[11] 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 "Play C if yesterday's outcome was (C,C) or (D,D) and otherwise play D" is an equilibrium. This strategy is called "Win-Stay, Lose-Shift" or "Perfect Tit-for-Tat" (PTFT) as for high enough discount factors it is subgame-perfect in the

[10] Due to technical difficulties with the computer software, however, the actual sequence of game lengths deviated somewhat in certain sessions.

[11] 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.

prisoner's dilemma without noise while TFT is typically not.[12] When b/c=2 or 2.5, PTFT is not a Nash equilibrium in the game with errors. Intuitively, the cost of contributing exceeds the expected discounted benefit of the partner contributing tomorrow, so one period of punishment is not sufficient to deter defection.[13] However the game still has cooperative equilibria that prescribe two periods of punishment for each observed defection.

Standard equilibrium analysis thus 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 [1992] and Imhof, Fudenberg and Nowak [2007] predict cooperation when b/c=4, and moreover predict that subjects will play PTFT. Since these analyses restrict attention to memory-1 strategies, however, they predict defection when b/c= 2 or 2.5. The Fudenberg and Maskin [1990], [1993] 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.

In addition to theoretical considerations, experimental work on repeated games without errors suggests that cooperation is more likely when it is more beneficial, and in particular that the existence of a cooperative equilibrium is necessary but not sufficient for there to be a substantial amount of cooperation (e.g., Dal Bó [2005], Dal Bó and Frechette [2009], Dreber et al [2008], Roth and Murnighan [1978], Murnighan and Roth [1983], Feinberg and Husted [1993], Duffy and Ochs [2004]). Dal Bó [2005] finds that there is cooperation 27% of the time when $\delta = 1/2$, and 37% of the time when $\delta = 3/4$. Dal Bó and Frechette [2009] find little cooperation in cases where the cooperative strategy TFT is risk-dominated by ALLD, which is the case when the present value of the gain from cooperation is small, but find much more cooperation when TFT risk-dominates ALLD. Additionally, Dal Bó and Frechette [2009] find that cooperation

---

[12] In the game without errors, PTFT is a subgame-perfect equilibrium if $c < \delta(b-c)$ or $\delta > 1/(b/c-1)$ so in particular it is subgame perfect when b/c=2.5

[13] 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. We also find that PTFT is an equilibrium when b/c=4.

increases (up to 76% in the most favorable treatment) as the discount factor and static payoff to (C,C) increase.

When the payoff matrix of the prisoner's dilemma takes the benefit-cost form used in this paper, so that the static gain to D is independent of the opponent's strategy, the condition for TFT risk dominating ALLD in the game without errors takes a simple form, as can be seen from the row player's payoffs for the associated symmetric 2x2 game, which are

	TFT	$ALLD$
TFT	$(b-c)/(1-\delta)$	$-c$
$AllD$	b	0

It follows that TFT risk dominates ALLD if $(b-c)/(1-\delta)-c>b$, or $\delta>\delta^{*}=2c/(b+c)$. The control two-action treatments of Dreber et al [2008] have this form, with b/c = 1.5 or 2, and $\delta=3/4$. Thus TFT risk-dominates ALLD when b/c=2 but not when b/c=1.5, and consistent with the predictions of Dal Bò & Frechette [2009], cooperation was much more prevalent in the high-b/c condition.[14]

Since we use $\delta=7/8$ in this paper, TFT would risk dominate ALLD in all payoff conditions in the absence of noise. However, noise lowers the payoff of TFT against itself sufficiently that TFT is not an equilibrium when b/c=1.5. Inspection of the payoff tables for the games with noise (in the online appendix) shows that TFT is an equilibrium of the 2x2 game when b/c=2 but that it is only risk dominant if b/c=2.5 or 4. Thus to the extent that this heuristic extends to games with noise it predicts substantially more cooperation when b/c=2.5 than when b/c=2.[15]

These concerns lead to our first experimental set of experimental questions:

---

[14] In Dreber et al [2008], there was twice as much cooperation in the high-gains condition than in the low-gains condition (43% vs. 21%). In the low gains condition, first round cooperation decreases over time (First 4 interactions, 63% C; last 4 interactions, 19% C), whereas cooperation is stable in the high gains condition (First 4 interactions, 53% C, last 4 interactions, 58% C).

[15] Note that using Grim as the cooperative strategy instead of TFT gives exactly the same risk-dominance conditions in games without noise. However, using PTFT results in a higher discount factor threshold as when this strategy meets ALLD it ends up playing C every other period; the critical discount factor is then $2c/(b-c)$, so that cooperation is only predicted when b/c=4.

QUESTION 1: Is cooperation more frequent in treatments where it is an equilibrium?

QUESTION 2: Do players cooperate when cooperation cannot be supported with a memory-1 strategy?

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, thus forcing subjects to choose a strategy that depends only on the previous period's outcome is much too restrictive, and even allowing for all memory-two strategies gives too large a set to explicitly present. Of course the data cannot discriminate between all possible repeated game strategies, and we used a combination of 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 the strategies used in repeated games. In the repeated prisoner's dilemma, Wedekind and Milinski [1996] note that players rarely play C in the period 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 [2009] 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 players 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). They find that subjects play "trigger strategies" of memory 1, except in the limit no-noise case where signals from two periods ago also have an impact; note that the usual Prisoner's Dilemma strategies such as TFT are not implementable in their setting as no particular signals are associated with the actions of any given player. 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 72% of the interactions the actions of the player A's is consistent with a grim trigger strategy, while the play of the player B's is more diverse.

Of course 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 period 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 histories to occur and thus make it easier to distinguish between histories.

QUESTION 3: What strategies do players use in the noisy prisoner's dilemma?

QUESTION 4: Do players use PTFT when it is an equilibrium strategy?

QUESTION 5: How do the strategies used vary with the gains to cooperation?

QUESTION 6: How do the strategies used differ from those used in games with perfectly observed actions?

There is now considerable evidence that some subjects are not motivated only by their own monetary payoffs in the game, but also by the payoff received by others. The fraction of subjects with such altruistic or, more generally, "social," preferences varies from experiment to experiment, and in some cases (such as a one-shot prisoner's dilemma played against a series of different opponents) also decreases markedly over the course of a given experimental session (Andreoni [1995a], [1995b], Bereby-Meyer and Roth [2006], Cooper et al [1996], Fehr and Gächter [2000], [2002]). This raises the question of how much of the observed cooperation is due to alternative preferences as opposed to strategic play by self-interested players. To address this question, we had subjects play a dictator game at the end of the session. If altruism is a main force leading to cooperative play, we would expect that subjects who were more generous in the dictator game played more cooperatively in the repeated game.

QUESTION 7: How does giving in the dictator game correlate with cooperation in the first round of each interaction and with the overall frequency of cooperation?

To further explore motivations for cooperation, we had subjects complete a series of self-report measures. Subjects indicated the extent to which their motivation for cooperating following each of CC, CD, DC and DD was to (i) maximize their long-term payoff, (ii) help the other player earn money, (iii) do the morally right thing or (iv) avoid upsetting the other player. We also collected data on various demographic variables which may be relevant to cooperation, such age, gender and major (Croson and Gneezy [2009], Frank et al [1993]).

QUESTION 8: How do demographic/survey data correlate with cooperation?

4. Methodology

To have any hope of inferring the subjects' strategies from their play, we must focus our attention on a subset of the infinitely many possible strategies. We begin with those strategies which have received particular attention in the theoretical literature: Always Defect (ALLD), Always Cooperate (ALLC), Grim[16], Tit for tat (TFT), and Perfect Tit for Tat (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 [2009].[17]

---

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

[17] 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 periods. 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

To inform our extension of this strategy set, we asked subjects to describe their strategies in a post-experimental questionnaire. 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. 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. Thus we restrict our attention to strategies that depend only on play in the last three periods, with the exception of Grim and its variants, which use at most the last 3 periods when deciding whether to switch to permanent defection.

For 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 Grim 3 (waits for three consecutive D rounds). We include three strategies that punish twice (intermediate to TFT's one period of punishment and Grim's unending punishment) but can be implemented by conditioning only on the last 3 periods. 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

---

"punishment rounds." Additionally, T2 begins its punishment if either player defects, whereas 2TFT responds only to the partner's defection.

row, and then punishes by playing D twice in a row. Because we do not include strategies that punish for a finite number of periods greater than two, our estimated share of "Grim" strategies may include some players 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[18], D-TF2T, D-TF3T, D-Grim2 and D-Grim3. [19] Because TF2T appears prevalent in many treatments, we also looked at whether subjects used the strategy 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 3; complete descriptions are given in the online appendix.

[Table 3]

To assess the prevalence of each strategy in our data, we follow Dal Bó and Frechette [2009] 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)[20], and moreover that in addition to the extrinsically imposed execution error, subjects make mistakes when choosing their intended action, so every sequence of choices has positive probability. 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

[18] Boyd and Lorberbaum [1987] call this strategy "Suspicious Tit for Tat."
[19] 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.
[20] 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.

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, matches, 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.

To better understand the mechanics of the specification, suppose that an interaction lasts w rounds, that in the first period the subject chose C, the first period outcome was that the subject played C and her partner played D, and in the second period the subject chose D. Then for strategy *s* TFT, which plays C in the first period, and plays D in the second period following (C,D), the likelihood of this sample 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 $\sum_I \ln\left(\sum_{s\in S} p(s)p_i(s)\right)$,and we then use maximum likelihood estimation (MLE) to estimate the prevalence of the various strategies. Note that the specification assumes that all 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. Finally, we use 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 experiment, 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 each of the 4 strategy distributions estimated from the experimental data (see Table 4 below), 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. [21]

5. Results

QUESTION 1: Is cooperation more frequent in treatments where it is an equilibrium?

We begin by examining cooperation in the first round of each interaction (Figure 1), and use it to investigate the extent of learning over the session for each payoff specification. [22] There is no significant relationship between interaction number and first round cooperation when b/c=1.5 (coeff =0.006, p=0.85), a significant positive relationship when b/c=2 (coeff=0.089, p=0.009), a non-significant relationship when b/c=2.5 with a nonetheless rather sizable positive coefficient (coeff=0.056 p=0.32), and a significant positive relationship when b/c=4 (coeff=0.034, p=0.029). We conclude that some learning occurred when b/c=2 and b/c=4, and since it may have occurred in the other sessions as well, our analysis will focus on how subjects played in the last four interactions of the session.

[Figure 1]

Figure 2 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. [23] We see that there is markedly less cooperation when b/c=1.5, both in the first round (1.5

[21] See Appendix A for MLE results using simulated data.
[22] 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 session, we clustered on both subject and session.
[23] 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 session.

vs 2, p=0.017; 1.5 vs 2.5, p<0.001; 1.5 vs 4, p=0.003) and overall (1.5 vs 2, p=0.020; 1.5 vs 2.5, p<0.001; 1.5 vs 4, p<0.001). Conversely, we see little difference in cooperation between the three treatments where cooperation is an equilibrium, either in the first round (2 vs 2.5, p=0.66; 2 vs 4, p=0.99; 2.5 vs 4, p=0.58) or overall (2 vs 2.5, p=0.069; 2 vs 4, p=0.183; 2.5 vs 4, p=0.714). The reason there is about the same amount of initial cooperation in b/c= 2 and 2.5 yet more cooperation in the latter case seems related to the fact that more players are forgiving in the latter treatment, as seen in the discussion of Questions 3-4.

[Figure 2]

QUESTION 2: Do players cooperate when cooperation cannot be supported with a memory-1 strategy?

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. These results are a first sign that the predictions of the memory-1 restriction are not consistent with the data.

QUESTION 3: What strategies do players use in the noisy prisoner's dilemma?

QUESTION 4: Do players 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. We consider 20 strategies in total (Table 3): 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. Of these, only 9 are present at frequencies significantly greater than 0 in at least one payoff specification: the cooperative strategies

TFT, TF2T, TF3T, 2TF2T, Grim, Grim2 and Grim 3,and the non-cooperative strategies ALLD and D-TFT.[24]

Thus we restrict our attention to these 9 strategies (Table 4). 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, 2TF2T, Grim2 and Grim3 are all lenient.

[Table 4]

QUESTION 5: 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 5. Three trends are apparent.

[Table 5]

First, cooperation is significantly lower at b/c=1.5 than at the higher b/c ratios, as shown in Table 5 and visualized in Figure 2. Consistent with this observation, Table 5 also shows that the share of the non-cooperative strategies ALLD and D-TFT is 44% 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). [25]

Second, leniency also increases with the gains of cooperation. To get a measure of leniency distinct from the MLE estimates, we examine all histories in which both players played C in all but the previous period, while in the previous period one player played D.[26] We then ask how frequently the player who had hitherto cooperated showed leniency by continuing to cooperate despite the partner's defection. [27] At b/c=1.5, 17% of histories

---

[24] See Appendix B for the estimates and standard errors for the full set of 20 strategies.

[25] 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.

[26] We also include 2nd round decisions in which the first round's outcome was CD.

[27] 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 session.

show leniency, compared to the significantly higher values of 63% at b/c=2 ($p<0.001$), 67% at b/c=2.5 ($p<0.001$) and 66% at b/c=4 ($p<0.001$). No significant difference in leniency exists among the higher b/c ratios ($p>0.5$ 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 non-lenient strategies Grim and TFT is 38% at b/c=1.5, which is significantly greater than at higher b/c vales (b/c=2, 19%, p=0.002; b/c=2.5, 21%, p=0.004; b/c=4, 13%, $p<0.001$). The opposite is true of the lenient strategies Grim2, Grim3, TF2T, TF3T and 2TF2T which have a combined frequency of 18% at b/c=1.5, compared to 64% at b/c=2 ($p<0.001$), 60% at b/c=2.5 ($p<0.001$) and 64% at b/c=4 ($p<0.001$).

The gains to cooperation also influence the frequency of forgiveness. To measure forgiveness, we examine all histories in which (i) at least 1 player chose C in the first round, (ii) in at least one previous period, the initially cooperative player chose C while the other player chose D and (iii) in the immediately previous period the formerly cooperative player played D. We then ask how frequently this formerly cooperative player showed forgiveness by returning to C. 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=0.005; 2 vs 4, $p<0.001$). 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 unforgiving strategies Grim, Grim2 and Grim3 are significantly more common at b/c=2 (57%) than at b/c=2.5 (38%, p=0.032) or b/c=4 (22%, $p<0.001$). Conversely, the forgiving strategies TFT, TF2T and TF3T are less common at b/c=2 (16%) than at b/c=2.5 (31%, p=0.044) and b/c=4 (43%, $p<0.001$).

QUESTION 6: How do the strategies used differ from those used in games with perfectly observed actions?

The prevalence of leniency in our experiment with noise stands in contrast to games without noise. Reanalyzing data from Dal Bò & Frechette [2009] and Dreber et al [2008] using our strategy set from Table 4 finds TFT to be the most common cooperative strategy in all payoff specifications.[28] Additionally, the aggregate frequency of strategies with memory at most one [29] is 90% in the games without noise, compared to only 49% in our games with noise; this difference is largely driven by lenient strategies, which by definition look back more than 1 period, and have aggregate frequency of 10% without noise compared to 51% with noise.[30] Further evidence for strategies with longer than 1 period memory comes from conditioning on the other player's play 2 periods ago. 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.

To confirm that the difference in leniency was due to noise and not other aspect of the experiment, we ran a control treatment where actions were implemented without error, using b/c=4 and δ=7/8 (18 subjects, 41% female, 9 interactions, average interaction length 8.22 rounds). In stark contrast to our results with errors, we found that the most common cooperative strategy was the non-lenient Grim (35% Grim without errors, 4% Grim with errors). [31] As in the other no-error experiments, 88% of strategies in our no-error condition had memory at most 1, compared to 36% in the b/c=4 condition with error; and while 64% of strategies were lenient at b/c=4 with errors, only 42% were lenient without noise. Thus the differences between our results with error and previous experiments without error do not seem to be a result of differences in subject pool, game parameters or details of the experimental protocol.

In addition to the increased leniency in the error treatments, there is also less forgiveness: Without errors, 67% of cooperative strategies have 1-2 periods of punishment, with the rest being a form of "Grim," while with errors only 51% of the cooperative strategies are forgiving. This is consistent with players being aware of incentives: Since lenient strategies are less likely to punish, when they do the eventual punishment must be longer; this is why Grim2 is an equilibrium, but 2TF2T is not.

[28] See Appendix C for MLE results.
[29] In our strategy set these strategies are ALLC, ALLD, TFT, Grim and D-TFT.
[30] The results are qualitatively equivalent when restricting our attention to the no-noise payoff specifications where TFT risk-dominates ALLD: 87% of strategies use memory at most 1, and lenient strategies have weight 13%.
[31] See Appendix C for MLE results.

QUESTION 7: How does giving in the dictator game correlate with cooperation in the first round of each interaction and with the overall frequency of cooperation?

To answer this question, we ran logistic regressions over first-round/all rounds in the last 4 interactions, with b/c ratio and subject's donation in the dictator game as independent variables, clustering on subject and session. First we consider cooperation in the first round of each interaction, which is an indication of whether the subject is playing a fundamentally cooperative or non-cooperative strategy. We find no significant relationship between the amount given in the dictator game and first round cooperation (coeff=0.062, p=0.417). To test whether the relationship between altruism and cooperation varies across b/c ratios, we next included an interaction term between b/c ratio and dictator donation. Doing so produces a marginally significant negative interaction term (coeff=-0.217, p=0.067), and evaluating the net coefficient[32] of dictator donation for each b/c ratio finds a significant effect of donation at b/c=1.5 (net coeff=0.19, p=0.0439) and no significant effect at higher b/c ratios (b/c=2: net coeff=0.09, p=0.23; b/c=2.5: net coeff=-0.02, p=0.80; b/c=4: net coeff=-0.35, p=0.15).

Next we consider overall cooperation in the last 4 interactions. We find marginally significant positive relationship between dictator donation and overall cooperation (coeff=0.092, p=0.058). As with first-round cooperation, including an interaction between b/c ratio and dictator donation shows a marginally significant negative interaction between b/c ratio and dictator donation (coeff=-0.10, p=0.066). Examining the net coefficient of dictator donation for each b/c ratio finds a significant effect of donation at b/c=1.5 (net coeff=0.16, p=0.006) and b/c=2 (net coeff=0.11, p=0.014), and no significant effect at b/c=2.5 (net coeff=p=0.21) or b/c=4 (net coeff=, p=0.28).

The results on first round cooperation suggest that altruism is not motivating the choice between cooperative and non-cooperative strategies, except perhaps in the b/c=1.5 specification where cooperation is not an equilibrium. The results on overall cooperation suggest that altruism may play some role in choosing which cooperative strategy to play,

[32] The net coefficient is the main effect coefficient for dictator donation plus the product of the interaction term coefficient and the b/c value in question.

particularly in the lower b/c treatments. Taken together, this analysis shows that if altruism is important in our study, it is so only in the payoff specifications with the least cooperative play, and thus altruistic preferences cannot explain why cooperation increases with b/c or why we see leniency and forgiveness in our high b/c treatments.

QUESTION 8: How do demographic/survey data correlate with cooperation?

We included survey questions to elicit the motivation behind subjects' choice of playing C vs D. For example, subjects were given questions such as "Imagine that last round you played C while the other played D. When you choose to now play C, to what extent is it motivated by (i) earning the most points in the long run (ii) helping the other person earn points, (iii) feeling it's the moral thing to do or (iv) not wanting to upset the other person." This enables us to look more closely at both leniency and forgiveness. Excluding subjects whose choices are best described by ALLD [33], we focus on to what extent alternative (i) "earning the most points in the long run" is the main motivator for playing C. We find that (i) is higher than all other motivations in all four b/c treatments (see Table A4 in Appendix D). That is, even though subjects may give some importance to all motivations, between 63% to 93% subjects rated (i) higher than (ii), 74% to 90% rated (i) higher than (iii), and 78% to 98% rated (i) higher than (iv). Earning the most points in the long run thus seems to be an important motivation behind something forgiveness-related like when subjects play C after the other person just played D.

We asked the same question about motivation for playing C in the other three possible states (C after CC, C after DC, C after DD). To look at the importance of each specific motivator across the four states, we make a composite measure which is the sum of (i) over all four states (CC, CD, DC, DD), the sum of (ii), the sum of (iii) and the sum of (iv). Regressing each of our cooperation measures separately against all four of these composite cooperation motivations while controlling for each payoff specification (and excluding subjects who are ALLD), we find that "earning the most points in the long run" is a significant predictor of first-round cooperation (coeff=0.161, $p<0.001$) and overall

[33] The likelihood that subject *s* uses strategy *i* can be calculated as in Eqn 1. Using the appropriate value of γ from Table 2, we can thus determine which individual subjects are best described by ALLD for a given payoff specification (i.e. ALLD has a higher likelihood than any other strategy for that particular subject).

cooperation (coeff=0.124, p<0.001), as is “not wanting to upset the other person” (1st round: coeff=.069, p=0.23; overall: coeff=0.056, p<0.001). Note however that coefficients for the self-interested motivation (i) are more than twice as large for either cooperation measure.

Motivations may play different roles in the different payoff specifications, so we next performed additional regression including interaction terms between the b/c ratio and the two significant motivations (i and iv). We find a significant positive interaction between b/c ratio and the profit-maximization motivation (i) for both first-round cooperation (coeff=.049, p=0.032) and overall cooperation (coeff= .016, p=0.017), and no significant interaction for the other-regarding motivation (iv) (1st round: coeff= -.028, p=0.495; overall: coeff= -.010, p=0.544). Thus payoff maximization is a more important motivator in the higher payoff specifications, while not wanting to upset the other player is equally important across specifications. This suggests that the desire to maximize monetary payoff is the most important motivator of the increased cooperation and leniency in the cooperation-supporting treatments.

Finally, we ask whether cooperation varies based on gender and major, comparing economics and psychology students to those studying all other disciplines [34] We find that women cooperate significantly less than men, both in the 1st round (coeff=-0.598, p=0.001) and overall (coeff=-.324, p=0.001), while major has no significant effect (1st round: coeff=-.069, p=0.80; overall: coeff=-0.173, p=0.365). We also find no significant interaction between gender and b/c ratio (1st round: coeff=.145, p=0.445; overall: coeff= -.064, p=0.461). Thus our data suggest that women consistently cooperate less than men across all payoff specifications.

Previous work has consistently found women to be more risk averse than men (e.g., Croson and Gneezy [2009]), thus differences in risk preferences could be responsible for the gender effect we observe. To evaluate this possibility, we perform two additional analyses controlling for risk preferences using (i) a hypothetical investment question, and (ii) a risk perception or general risk taking question. Both questions are

---

[34] At Harvard it is common to study the prisoner’s dilemma in both fields. Thus we report the results of logistic regression over first-round/all individual decisions in the last 4 interactions, with independent variables for b/c ratio, gender (1=female) and major (1=economics/psychology), clustered on both subject and session.

from Dohmen et al [forthcoming]. Gender remains significant despite controlling for either risk measure (Gender with lottery: 1st round C, coeff=-.623, p=0.002; overall C, coeff=-.338, p=0.001; Gender with risk perception: 1st round C, coeff=-.671, p=0.008; overall C, coeff=-.334, p=0.003). Thus risk preferences do not appear to be driving the gender difference in cooperation seen in our data.

6. Discussion

To relate play in the experiment to theoretical predictions we first use simulations to compute the expected payoff matrix for the strategies that had non-negligible shares in the MLE estimation. (Table 4).[35] We also include several exploitive strategies that did not seem to be used but have some theoretical appeal as responses to commonly used lenient strategies. The resulting payoff matrix is displayed in Table 6.

[Table 6]

Note first that any strategy that is not a Nash equilibrium in this payoff table cannot be a Nash equilibrium in the full game, but that the converse is false, as an equilibrium of this reduced game (a "reduced-game equilibrium") might be vulnerable to a strategy that we have not included. Recall also that ALLD will be an equilibrium for any b/c ratio, and that we have already computed that TFT is a reduced-game equilibrium except when b/c=1.5, and that PTFT is only an equilibrium when b/c=4.

Looking at the payoff table, we see that Grim is a reduced-game equilibrium except if b/c=1.5, so there is no cooperative equilibrium in this condition. Perhaps surprisingly, Grim 2 is a reduced-game equilibrium if b/c=4; moreover it turns out that Grim2 is a full equilibrium in this condition, even though it is not an equilibrium in the game without errors. Without errors it would be better to play D in the first period 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 period following a D, and numerical calculations show that when b/c=4 and δ=7/8, Grim2 is an equilibrium provided that the error probability is between 0.0332 and 0.2778. Moreover, the range of error probabilities for which Grim2 is an equilibrium

[35] 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 period and/or have many implicit "states."

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.

We also see that the exploitive but lenient strategy D-Grim2 is a reduced-game equilibrium in all payoff specifications, and that D-Grim3 is a reduced-game equilibrium in all payoff specifications other than b/c=1.5. Mathematica computations show that while D-Grim2 is a full equilibrium at b/c=4, D-Grim3 is not, and neither exploitive strategy is a full equilibrium in any of the lower b/c payoff specifications.

The payoff table also shows 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. This is because the one-period 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. [36]

Note that the payoff table does not show the payoff for PTFT, even though as noted earlier this strategy is an equilibrium when b/c=4. We exclude it because it does not seem to be used and does not do very well against the observed distribution of play. Note also that TFT is never an equilibrium when all strategies in the MLE are considered, 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 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]

Table 7 shows the expected payoff of each strategy given the prevailing strategy frequencies. In the treatment when b/c=1.5, where ALLD is most prevalent, ALLD is also the best response to the prevailing strategy frequencies, and the average earnings per

[36] 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 there is no b/c value for which TF2T is an equilibrium.

round of subjects for whom ALLD is the strategy with the greatest likelihood are significantly higher than other subjects' earnings (coeff=0.177, p<0.001). ALLD does about as well as the average subject when b/c=2 (coeff=0.032, p=0.328), and is significantly worse at b/c=2.5 (coeff=-0.372, p=0.001) and b=4 (coeff=-1.129, p<0.001). [37]

In the treatment b/c=4, where TF2T was the most common cooperative strategy, a number of the exploitive "start with D" strategies slightly beat TF2T in pairwise play, but TF2T scores highest against the population as whole. [38] This is because the deceptive strategies do poorly against TFT and variants of Grim. In a sense, the lenient and forgiving strategy TF2T, which is not an equilibrium in its own, is "protected" by less lenient and/or forgiving strategies, so it is benefiting from "punishment costs" (in the form of reduced cooperation) borne by the less lenient/forgiving strategies. This is reminiscent of the "free riding on other's punishment" found in other settings by e.g. Falk et al [2005], Fehr and Gächter [2000] and Rand et al [2009] and discussed theoretically by e.g. Boyd et al [2003]. A similar effect occurs with the very lenient Grim3, which is the most common strategy at b/c=2 and b/c=2.5. Despite losing in pairwise comparisons with various start-with-D strategies, Grim3 scores highest against the population at b/c=2, and very close to highest at b/c=2.5 (Grim3 is the 3rd highest scorer at b/c=2.5, and only 0.1 behind the winner D-TFT's earnings of 14.86).

Based on the expected payoffs in Table 7, perhaps the largest surprise is not the success of leniency and forgiveness, but rather the high level of ALLD players, particularly at b/c=4. The reason that low performing strategies such as ALLD and Grim 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 and Grim are not best responses to what people are really doing, ALLD is a best response to a belief that everyone else plays D, and Grim is a best response to the belief that a substantial fraction of the population plays Grim while the rest plays ALLD. Moreover, subjects who play optimally given such false beliefs may not receive observations that

[37] We report the results of a linear regression over profit in all round of the last 4 interactions, with an ALLD dummy as the independent variable, clustered on both subject and session.
[38] In addition to the exploitive strategies shown in Table 7, at b/c=4 simulated payoffs show that D-TF2T earns 29.684 against the observed distribution, and D-TF3T earns 29.466.

show them that more cooperative strategies yield a higher payoff- for example a player who always uses Grim may not learn about the benefits of being more lenient. And because these non-lenient strategies persist, exploiter strategies don't do very well.[39]

We find a robust gender difference in play in the Prisoner's Dilemma. Women cooperate less than men when we look at cooperation overall, the very first round, or the first round of the last four interactions. Previous literature on gender differences in the Prisoner's Dilemma shows mixed results, with some experiments finding that women cooperate more than men while others find the opposite. Croson and Gneezy [2009] review these results and suggest that these inconsistencies depend on women being more sensitive to subtle cues in the experimental context than men, which perhaps applies to our results too.

Because cooperation "succeeded" in all three treatments where it is an equilibrium, and there is about the same level of cooperation at b/c=2 (where ALLD risk dominated Grim) as at b/c=2.5 (where Grim risk-dominates ALLD) our results suggest that the pairwise risk dominance may not be a good predictor of cooperation levels in games with noise, unlike games without noise (e.g. Dreber et al. [2008], Dal Bò & Frechette [2009]). Based only on our data, the simplest explanation is that players will cooperate to a significant extent whenever cooperation is an equilibrium. If this generalization is correct, then it would show that noisy observations will if anything make cooperation easier, due to subjects' tendency to give each other the benefit of the doubt.

7. Conclusion:

We conclude that subjects do tend to cooperate in noisy repeated games when cooperation is an equilibrium strategy, and that they do so even when there are no cooperative equilibria in memory-1 strategies, and 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

[39] This is reminiscent of heterogeneous self-confirming equilibrium (Fudenberg and Levine [1993], and the diversity of strategies would be consistent with self-confirming equilibrium in the absence of noise. In the presence of noise similar situations can persist for a while.

pretty 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. It can be an equilibrium for all agents to use the lenient but unforgiving strategy Grim2, but this strategy was not used in the treatment where it is an equilibrium.

Some subjects are altruistic, but strategic considerations and not altruism provide the best explanation of how play changes with the payoff matrix of the stage game. Although leniency and forgiveness may at first seem to be the result of altruistic tendencies, our results show that in an uncertain world, it can be payoff-maximizing to forgive and forget.

References

Andreoni, J.A. [1995a] "Warm-Glow versus Cold-Prickle: The Effects of Positive and Negative Framing on Cooperation in Experiments," *Quarterly Journal of Economics*, 110, 1–21.

Andreoni, J.A. [1995b] "Cooperation in Public Goods Experiments: Kindness or Confusion?'" *American Economic Review*, 85, 891-904.

Aoyagi, M. and G. Frechette [2009] "Collusion as public monitoring becomes noisy: Experimental evidence" *Journal of Economic Theory*, 144, *1135-1165*

Axelrod, R. and W.D. Hamilton [1981] "The Evolution of Cooperation," *Science*, 211, 1390–1396.

Bereby Meyer, Y. and A.E. Roth [2006] "The Speed of Learning in Noisy Games: Partial Reinforcement and the Sustainability of Cooperation," *American Economic Review*, 96, 1029-1042.

Binmore, K. and L. Samuelson [1992] "Evolutionary Stability in Repeated Games Played by Finite Automata," *Journal of Economic Theory*, 57, 278–305.

Boyd, R., H. Gintis, S. Bowles, and P.J. Richerson [2003] "The Evolution of Altruistic Punishment," *Proceedings of the National Academy of Sciences (USA)*, 100, 3531–3535.

Boyd, R. and J.P. Lorberbaum [1987] "No Pure Strategy is Stable in the Repeated Prisoner's Dilemma Game," *Nature*, 327, 58–59.

Cooper, R., D.V. DeJong, R. Forsythe and T.W. Ross [1996] "Cooperation without Reputation: Experimental Evidence from the Prisoner's Dilemma Games," *Games and Economic Behavior*, 1996, 12(2), pp. 187–218.

Croson, R. and U. Gneezy [2009] "Gender Differences in Preferences," *Journal of Economic Literature*, 47, 1-27.

Dal Bó, P. [2005] "Cooperation Under the Shadow of the Future: Experimental Evidence from Infinitely Repeated Games," *American Economic Review*, 95, 1591-1604.

Dal Bó, P. and G. Frechette [2009] "The Evolution of Cooperation in Infinitely Repeated Games: Experimental Evidence," forthcoming in the *American Economic Review*,

Dohmen, T.J., A. Falk, D. Huffman, J. Schupp, U. Sunde, and G.G. Wagner [forthcoming] "Individual Risk Attitudes: New Evidence from a Large, Representative, Experimentally-Validated Survey," *Journal of European Economic Association*.

Dreber A., D.G. Rand, D. Fudenberg and M.A. Nowak [2008] "Winners Don't Punish," *Nature,* 452, 348-351.

Duffy, J. and J. Ochs [2009] "Cooperative Behavior and the Frequency of Social Interaction," *Games and Economic Behavior*, 66, 785-812.

Engle-Warnick, J. and R.L. Slonim [2006] "Inferring Repeated-Game Strategies from Actions: Evidence from Trust Game Experiments," *Economic Theory*, 54, 603-632.

Falk, A., E. Fehr and U. Fischbacher [2005] "Driving Forces Behind Informal Sanctions," *Econometrica*, 73, 2017-2030.

Fehr, E. and S. Gächter [2000] "Cooperation and Punishment in Public Goods Experiments," *American Economic Review*, 90, 980–94.

Feinberg, R.M. and T.A. Husted [1993] "An Experimental Test of Discount-Rate Effects on Collusive Behavior in Duopoly Markets," *Journal of Industrial Economics*, 41, 153–60.

Fischbacher, U. [2007] "z-Tree: Zurich toolbox for ready-made economic experiments," Experimental Economics, 10, 171–178.

Frank, R.H., T. Gilovich, D.T. Regan [1993] "Does Studying Economics Inhibit Cooperation?" *Journal of Economic Perspectives*, 7, 159–71.

Fudenberg, D. and D. K. Levine [1993] "Self-Confirming Equilibrium," *Econometrica*, 61, 523-546.

Fudenberg, D., D.K. Levine and E. Maskin [1994] "The Folk Theorem with Imperfect Public Information," *Econometrica*, 62, 997-1040.

Fudenberg, D. and E. Maskin [1986] "The Folk Theorem in Repeated Games with Discounting or with Incomplete Information," *Econometrica*, 54, 533-556.

Fudenberg, D. and E. Maskin [1990] "Evolution and Cooperation in Noisy Repeated Games," *American Economic Review*, 80, 274–279.

Fudenberg, D. and E. Maskin [1994] "Evolution and Noisy Repeated Games," mimeo.

Imhof, L.A., D. Fudenberg and M.A. Nowak [2007] "Tit-for-Tat or Win-Stay, Lose-Shift?" *Journal of Theoretical Biology,* 247, 574-580.

Murnighan, J.K. and A.E. Roth [1983] "Expecting Continued Play in Prisoner's Dilemma Games," *Journal of Conflict Resolution*, 27, 279–300.

Nowak, M.A. and K. Sigmund [1992] "Tit for Tat in Heterogeneous Populations," *Nature*, 355, 250–253.

Rand, D., A. Dreber, T. Ellingsen, D. Fudenberg and M. Nowak [2009] "Positive Interactions Support Public Cooperation," *Science*, 325, 1272-1275.

Roth, A.E. and J.K. Murnighan [1978] "Equilibrium Behavior and Repeated Play of the Prisoner's Dilemma," *Journal of Mathematical Psychology*, 17,189 –98.

Wedekind, C. and M. 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.

Realized payoffs				**Expected payoffs**		

b/c = 1.5 (Realized payoffs)

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

b/c = 1.5 (Expected payoffs)

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

b/c = 2 (Realized payoffs)

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

b/c = 2 (Expected payoffs)

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

b/c = 2.5 (Realized payoffs)

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

b/c = 2.5 (Expected payoffs)

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

b/c = 4 (Realized payoffs)

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

b/c = 4 (Expected payoffs)

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

Table 1. *Payoff matrices for each b/c. Payoffs are denoted in points.*

	b/c=1.5	**b/c=2**	**b/c=2.5**	**b/c=4**
Sessions per treatment	3	2	3	4
Subjects per treatment	72	52	64	90
% women	50.0	44.2	50.8	57.5
Average number of interactions	11	11.5	10.7	11.25
Average number of rounds per interaction	8.4	8.3	8.3	8.1

Table 2. *Summary statistics for each payoff specification.*

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 subsequent 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 subsequent rounds occur in which either player played D, then play D forever
Lenient Grim 3	Grim3	Play C until 3 subsequent 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 period. 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

Table 3. *Descriptions of the 20 strategies considered.*

	b/c=1.5	b/c=2	b/c=2.5	b/c=4
TFT	0.22**	0.09†	0.09*	0.09**
	(0.05)	(0.05)	(0.04)	(0.03)
TF2T	0.05†	0	0.17*	0.18**
	(0.03)	(0)	(0.07)	(0.06)
TF3T	0.01	0.07	0.05	0.15**
	(0.01)	(0.05)	(0.04)	(0.05)
2TF2T	0	0.10†	0.12†	0.12*
	(0.02)	(0.05)	(0.06)	(0.06)
Grim	0.16**	0.10*	0.12*	0.04†
	(0.05)	(0.05)	(0.05)	(0.02)
Grim2	0.06†	0.21**	0.02	0.06*
	(0.03)	(0.07)	(0.02)	(0.03)
Grim3	0.06†	0.26**	0.24**	0.12**
	(0.03)	(0.08)	(0.07)	(0.04)
ALLD	0.29**	0.17**	0.14**	0.23**
	(0.06)	(0.05)	(0.04)	(0.04)
D-TFT	0.15**	0	0.05	0
	(0.04)	(0)	(0.03)	(0)
Gamma	0.47**	0.51**	0.5**	0.44**
	(0.02)	(0.03)	(0.03)	(0.03)

Table 4. *Maximum likelihood estimates using the last 4 interactions of each session. 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
Descriptive statistics				
%C 1st Round	54%	75%	79%	76%
%C All Rounds	32%	49%	61%	59%
Leniency	29%	63%	67%	66%
Forgiveness	15%	18%	33%	32%
MLE aggregation				
Non-cooperative strategies	44%	17%	19%	23%
Cooperative strategies	56%	83%	81%	77%
Non-lenient strategies	38%	19%	21%	13%
Lenient strategies	18%	64%	60%	64%
Non-forgiving strategies	28%	57%	38%	22%
Forgiving strategies	27%	16%	31%	43%

Table 5. *Descriptive statistics of aggregate behavior, as well as aggregated MLE frequencies from Table 4. The descriptive statistics for leniency and forgiveness are defined in the text below. For MLE aggregation, non-cooperative strategies are ALLD and D-TFT, all others are cooperative; non-lenient strategies are TFT and Grim; lenient strategies are TF2T, TF3T, 2TF2T, Grim2 and Grim3; non-forgiving strategies are Grim, Grim2 and Grim3; forgiving strategies are TFT, TF2T, TF3T and 2TF2T.*

b/c=1.5	ALLC	TFT	TF2T	TF3T	2TF2T	Grim	Grim2	Grim3	ALLD	D-TFT	D-Grim2	D-Grim3	DC-Alt
ALLC	6.99	5.03	6.79	6.98	6.62	-4.17	3.05	6.05	-10.99	2.78	-0.56	3.46	-2.60
TFT	8.31	5.09	7.83	8.22	7.31	2.01	4.16	6.56	-1.81	2.86	0.24	4.39	2.91
TF2T	7.13	5.03	6.90	7.11	6.74	0.57	4.83	6.16	-3.97	2.80	1.87	3.44	-1.59
TF3T	7.01	5.03	6.80	6.99	6.64	-0.55	4.41	6.38	-5.61	2.78	1.30	3.93	-2.14
2TF2T	7.24	5.05	6.92	7.19	6.65	1.08	5.03	6.31	-3.19	2.81	2.13	3.72	-1.18
Grim	14.44	5.15	7.32	8.99	6.55	3.27	4.96	6.49	-0.66	2.94	1.99	4.07	7.64
Grim2	9.63	5.10	6.96	7.59	6.68	2.15	5.69	6.32	-2.23	2.86	2.98	3.82	2.52
Grim3	7.63	5.06	6.89	7.13	6.66	1.13	5.26	6.68	-3.65	2.82	2.40	4.33	1.36
ALLD	**18.99**	**5.22**	8.45	10.91	7.29	**3.48**	5.83	7.97	**1.00**	**2.97**	3.40	5.57	**9.38**
D-TFT	9.81	5.14	9.12	9.69	8.34	1.31	5.25	7.70	-0.32	2.90	1.68	**5.82**	4.41
D-Grim2	12.03	5.18	8.38	9.23	7.96	1.95	**6.73**	7.58	-0.60	2.93	**4.41**	5.30	4.04
D-Grim3	9.35	5.10	8.31	8.65	7.90	0.57	6.16	**8.06**	-2.05	2.84	3.82	5.81	2.94
DC-Alt	13.40	5.13	**11.89**	**12.70**	**11.27**	-1.96	5.72	7.45	-4.59	2.88	3.45	5.09	3.80

b/c=2	ALLC	TFT	TF2T	TF3T	2TF2T	Grim	Grim2	Grim3	ALLD	D-TFT	D-Grim2	D-Grim3	DC-Alt
ALLC	14.00	**11.37**	13.71	13.97	13.50	-0.93	8.73	12.71	-9.99	**8.37**	3.94	9.31	1.19
TFT	15.31	10.18	14.53	15.17	13.70	5.28	8.67	12.51	-0.81	6.64	2.45	9.01	6.71
TF2T	14.13	11.19	13.81	14.10	13.54	3.83	10.52	12.62	-2.97	8.10	6.36	8.84	2.21
TF3T	14.01	11.34	13.71	13.98	13.50	2.73	10.14	13.07	-4.61	8.32	5.78	9.75	1.66
2TF2T	14.23	11.00	13.78	14.17	13.30	4.36	10.69	12.78	-2.20	7.83	6.60	9.11	2.62
Grim	21.47	9.03	11.95	14.18	10.90	**6.53**	8.78	10.83	0.33	5.23	3.97	6.73	11.44
Grim2	16.62	9.83	13.08	13.91	12.68	5.43	11.38	12.20	-1.22	6.69	7.47	8.61	6.30
Grim3	14.63	10.72	13.51	13.96	13.18	4.39	10.95	13.36	-2.65	7.59	6.89	10.12	5.14
ALLD	**25.99**	7.62	11.93	15.23	10.38	5.33	8.43	11.29	**2.00**	4.61	5.20	8.10	**13.20**
D-TFT	16.80	9.36	15.70	16.61	14.48	3.30	9.52	13.44	0.69	5.82	3.84	10.46	8.21
D-Grim2	19.03	8.38	14.17	15.31	13.61	3.92	**11.97**	13.10	0.39	5.35	**8.81**	10.01	7.85
D-Grim3	16.33	9.91	14.66	15.42	14.12	2.53	11.38	**14.62**	-1.05	6.90	8.22	**11.62**	6.73
DC-Alt	20.41	9.37	**18.37**	**19.47**	**17.54**	-0.08	10.19	12.52	-3.60	6.38	7.11	9.32	7.60

b/c=2.5	ALLC	TFT	TF2T	TF3T	2TF2T	Grim	Grim2	Grim3	ALLD	D-TFT	D-Grim2	D-Grim3	DC-Alt
ALLC	20.99	**17.71**	20.63	20.96	20.36	2.36	14.41	19.41	-8.99	**13.98**	8.39	15.10	4.99
TFT	22.31	15.27	21.25	22.11	20.09	8.56	13.22	18.48	0.19	10.42	4.66	13.70	10.50
TF2T	21.14	17.35	20.71	21.08	20.36	7.10	16.20	19.05	-1.96	13.42	10.82	14.23	6.00
TF3T	21.00	17.64	20.66	20.96	20.36	6.01	15.79	19.78	-3.61	13.86	10.25	15.57	5.45
2TF2T	21.23	16.93	20.63	21.15	19.95	7.60	16.41	19.21	-1.20	12.85	11.08	14.52	6.42
Grim	28.45	12.94	16.57	19.36	15.25	**9.80**	12.60	15.17	1.33	7.50	5.96	9.41	15.24
Grim2	23.61	14.55	19.21	20.23	18.69	8.69	17.05	18.10	-0.23	10.50	11.94	13.35	10.10
Grim3	21.62	16.38	20.09	20.79	19.72	7.63	16.63	20.04	-1.65	12.35	11.37	15.98	8.96
ALLD	**32.98**	10.04	15.40	19.53	13.49	7.15	11.05	14.60	**3.00**	6.28	7.00	10.62	**16.99**
D-TFT	23.80	13.59	22.28	23.53	20.61	5.26	13.82	19.17	1.69	8.70	6.07	15.11	12.01
D-Grim2	26.03	11.59	19.93	21.37	19.27	5.89	**17.19**	18.63	1.40	7.82	**13.24**	14.74	11.64
D-Grim3	23.35	14.72	21.03	22.19	20.36	4.50	16.61	**21.23**	-0.05	10.94	12.68	**17.42**	10.55
DC-Alt	27.39	13.63	**24.87**	**26.24**	**23.84**	1.80	14.66	17.51	-2.59	9.87	10.77	13.53	11.40

b/c=4	ALLC	TFT	TF2T	TF3T	2TF2T	Grim	Grim2	Grim3	ALLD	D-TFT	D-Grim2	D-Grim3	DC-Alt
ALLC	41.99	**36.74**	41.40	41.91	40.97	12.21	31.44	39.43	-6.02	**30.75**	21.82	32.53	16.43
TFT	43.28	30.54	41.39	42.94	39.32	18.36	26.85	36.29	3.19	21.74	11.31	27.67	21.91
TF2T	42.11	35.82	41.40	42.01	40.78	16.95	33.25	38.47	1.04	29.42	24.24	30.37	17.41
TF3T	42.01	36.58	41.42	41.93	40.96	15.83	32.78	39.79	-0.62	30.52	23.68	32.98	16.83
2TF2T	42.24	34.77	41.16	42.05	39.89	17.43	33.47	38.60	1.80	27.87	24.57	30.67	17.81
Grim	49.44	24.61	30.43	34.91	28.36	**19.63**	24.14	28.21	4.32	14.39	11.85	17.40	26.64
Grim2	44.63	28.74	37.55	39.15	36.79	18.53	**34.13**	35.78	2.76	21.94	25.39	27.70	21.52
Grim3	42.63	33.34	39.90	41.25	39.27	17.49	33.68	40.11	1.35	26.64	24.78	33.44	20.34
ALLD	**53.99**	17.23	25.86	32.45	22.76	12.64	18.89	24.55	**6.00**	11.25	12.41	18.19	**28.40**
D-TFT	44.80	26.28	42.03	44.28	38.96	11.14	26.70	36.40	4.69	17.45	12.60	28.98	23.40
D-Grim2	47.01	21.17	37.25	39.57	36.23	11.78	32.85	35.11	4.38	15.09	**26.53**	28.81	23.02
D-Grim3	44.34	29.11	40.11	42.47	39.04	10.42	32.34	**40.93**	2.96	23.03	25.84	**34.84**	21.92
DC-Alt	48.34	26.33	**44.38**	**46.54**	**42.72**	7.42	27.91	32.58	0.39	20.33	21.84	26.24	22.80

Table 6. *Row player's payoff is shown, averaged over 10^5 randomly simulated games. Best response in each column is shown in bold.*

	b/c=1.5		b/c=2		b/c=2.5		b/c=4	
	Freq	Payoff	Freq	Payoff	Freq	Payoff	Freq	Payoff
TFT	0.22	2.44	0.09	8.84	0.09	14.79	0.09	29.22
TF2T	0.05	1.52	0.00	8.73	0.17	14.74	0.18	29.87
TF3T	0.01	0.84	0.07	8.38	0.05	14.58	0.15	29.65
2TF2T	0.00	1.86	0.10	8.95	0.12	14.82	0.12	29.85
GRIM	0.16	3.02	0.10	8.28	0.12	12.41	0.04	23.29
GRIM2	0.06	2.37	0.21	9.03	0.02	14.07	0.06	27.78
GRIM3	0.06	1.77	0.26	9.03	0.24	14.76	0.12	29.37
ALLD	0.29	3.77	0.17	8.39	0.14	11.44	0.23	20.04
D-TFT	0.15	2.99	0.00	9.42	0.05	14.86	0.00	29.15
D-Grim2		3.06		9.58		13.95		26.95
D-Grim3		2.38		9.66		14.99		29.26
DC-Alt		1.39		8.77		14.49		28.90

Table 7. *Observed frequencies and resulting expected payoffs for each strategy.*

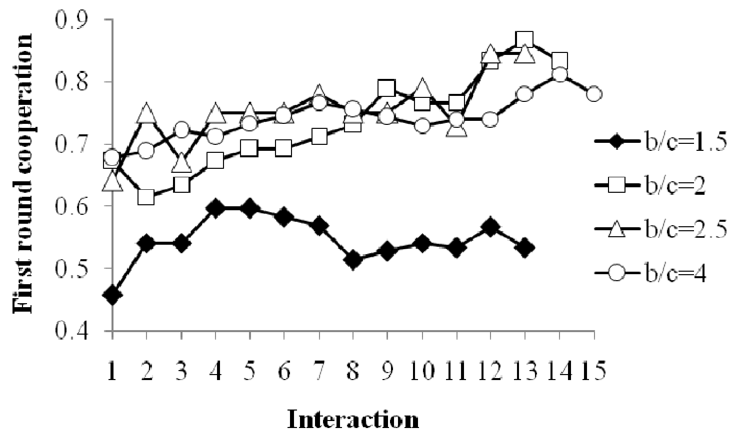

Figure 1. *First round cooperation over the course of the session, by payoff specification.*

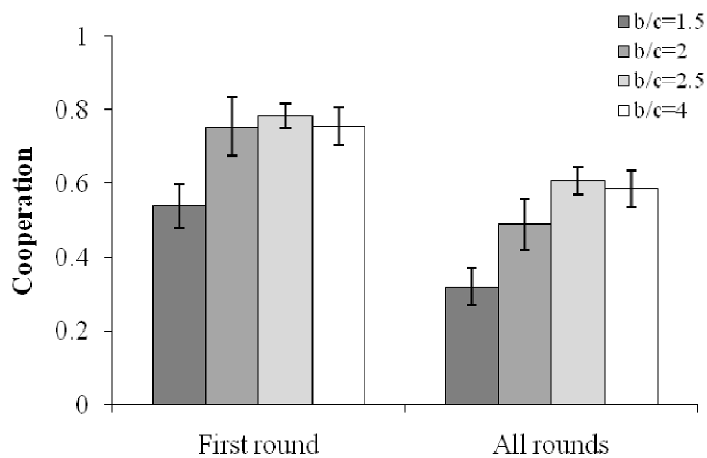

Figure 2. *First round and overall cooperation by payoff specification, averaged over the last 4 interactions of each session. Error bars indicate standard error of the mean, clustered on session and subject.*

Appendix A – MLE strategy frequency estimates using simulated data

	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)

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 2nd 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

Appendix B – MLE strategy frequencies using full 20 strategy set

	b/c=1.5	b/c=2	b/c=2.5	b/c=4
ALLC	0	0.03	0	0.05†
	(0)	(0.03)	(0.02)	(0.03)
TFT	0.19**	0.07	0.09*	0.07*
	(0.05)	(0.04)	(0.04)	(0.03)
TF2T	0.05†	0	0.16*	0.19**
	(0.03)	(0)	(0.07)	(0.06)
TF3T	0.01	0.03	0.05	0.09*
	(0.01)	(0.03)	(0.04)	(0.04)
2TF2T	0	0.11*	0.11†	0.12*
	(0.02)	(0.05)	(0.06)	(0.06)
Grim	0.14**	0.05	0.11*	0.02
	(0.05)	(0.05)	(0.05)	(0.02)
Grim2	0.05†	0.16*	0.02	0.05†
	(0.03)	(0.07)	(0.03)	(0.03)
Grim3	0.06†	0.27**	0.24**	0.11*
	(0.03)	(0.08)	(0.08)	(0.04)
PTFT	0	0	0	0
	(0)	(0)	(0)	(0)
2PTFT	0	0.03	0	0
	(0)	(0.03)	(0)	(0)
2TFT	0.06	0.07†	0.02	0.03
	(0.04)	(0.04)	(0.02)	(0.02)
T2	0	0	0	0
	(0)	(0)	(0)	(0)
ALLD	0.27**	0.17**	0.14**	0.21**
	(0.05)	(0.06)	(0.05)	(0.04)
C-to-ALLD	0	0.01	0	0.01
	(0.01)	(0.02)	(0)	(0.01)
D-TFT	0.1**	0	0.03	0
	(0.04)	(0)	(0.03)	(0)
D-TF2T	0	0	0.01	0
	(0)	(0)	(0.02)	(0)
D-TF3T	0.01	0	0	0
	(0.01)	(0)	(0)	(0)
D-Grim2	0.05†	0.01	0	0.01
	(0.03)	(0.01)	(0.01)	(0.01)
D-Grim3	0	0	0.01	0
	(0)	(0)	(0.01)	(0)
DC-Alt	0	0	0	0.01
	(0)	(0)	(0)	(0.01)
Gamma	0.46**	0.49**	0.49**	0.43**
	(0.02)	(0.03)	(0.03)	(0.02)

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

Appendix C – MLE strategy frequencies for games without noise

	Dal Bo & Frechette 2010						Dreber et al 2008		New data
Payoff	R=32	R=40	R=48	R=32	R=40	R=48	b/c=1.5	b/c=2	b/c=4
δ	0.5	0.5	0.5	0.75	0.75	0.75	0.75	0.75	0.875
TFT	0.07†	0.08†	0.24**	0.23**	0.46**	0.57**	0.15*	0.40**	0.07
	(0.04)	(0.04)	(0.07)	(0.08)	(0.12)	(0.14)	(0.07)	(0.14)	(0.09)
TF2T	0	0.02	0.16*	0.11†	0.12	0	0	0	0
	(0)	(0.02)	(0.07)	(0.06)	(0.07)	(0)	(0)	(0)	(0.03)
TF3T	0	0.01	0.02	0	0	0.07	0	0	0.16
	(0)	(0.01)	(0.02)	(0)	(0.05)	(0.05)	(0)	(0)	(0.13)
2TF2T	0	0	0	0	0.11	0	0	0	0
	(0)	(0)	(0.03)	(0.03)	(0.07)	(0)	(0)	(0)	(0.07)
Grim	0	0.04	0	0	0.14†	0.26*	0.07†	0.21†	0.35*
	(0)	(0.03)	(0)	(0.02)	(0.08)	(0.12)	(0.05)	(0.14)	(0.15)
Grim2	0	0.01	0.02	0	0	0.04	0	0	0.18
	(0)	(0.01)	(0.02)	(0)	(0.03)	(0.04)	(0)	(0)	(0.15)
Grim3	0	0	0.02	0	0	0.07	0	0	0.08
	(0)	(0.01)	(0.02)	(0)	(0.04)	(0.05)	(0)	(0)	(0.14)
ALLD	0.91**	0.76**	0.49**	0.66**	0.11*	0	0.64**	0.3**	0.06
	(0.04)	(0.06)	(0.08)	(0.08)	(0.05)	(0)	(0.09)	(0.09)	(0.05)
D-TFT	0.02	0.08†	0.04	0	0.08†	0	0.14*	0.09†	0.40**
	(0.02)	(0.04)	(0.03)	(0)	(0.04)	(0)	(0.07)	(0.07)	(0.08)
Gamma	0.34**	0.49**	0.4**	0.45**	0.33**	0.28**	0.36**	0.42**	0.19**
	(0.04)	(0.04)	(0.03)	(0.04)	(0.02)	(0.03)	(0.04)	(0.04)	(0.05)

Table A3. Maximum likelihood estimates for Dal Bò & Frechette [2009], 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; δ=3/4, the last 8 interactions; for δ =7/8 (our control), the last 4 interactions. Bootstrapped standard errors in parentheses.
† Significant at p<0.1, * Significant at p<0.05, ** Significant at p<0.01.

Appendix D – Motivations and demographics analysis

b/c	(i) > (ii)	(i) > (iii)	(i) > (iv)
1.5	0.63	0.74	0.78
2	0.87	0.90	0.87
2.5	0.93	0.89	0.87
4	0.88	0.81	0.98

Table A4. Share of subjects (excluding ALLD) who rate motivation (i) higher than (ii), (iii) and (iv) respectively. Motivations: (i) earning the most points (ii) helping the other person earn points; (iii) feeling it's the moral thing to do; (iv) not wanting to upset the other person."