Who Cooperates in Repeated Games?

Anna Dreber[a], Drew Fudenberg[b] and David G. Rand[c]

[a]Department of Economics, Stockholm School of Economics, P.O. Box 6501, 113 83 Stockholm, Sweden. anna.dreber@hhs.se

[b]Corresponding author. Department of Economics, Harvard University, Littauer Center 310, 1805 Cambridge Street, Cambridge, MA 02138, USA. Tel: +16174965895, fax: +161749557730, dfudenberg@harvard.edu

[c]Program for Evolutionary Dynamics, Berkman Center for Internet and Society, Department of Psychology, Harvard University, 1 Brattle Square ste 6, Cambridge, MA 02138, USA. drand@fas.harvard.edu

We explore the extent to which social preferences account for the observed heterogeneity in play in the repeated prisoner's dilemma with noise. We find that giving in a post-experimental dictator game is correlated with cooperation in the repeated game when no cooperative equilibria exist, but not when cooperation is an equilibrium. Furthermore, none of the commonly observed strategies are explained by inequity aversion or efficiency concerns. Various survey questions provide additional evidence for relative unimportance of social preferences. We conclude that cooperation in repeated games is primarily motivated by long-term payoff maximization; some subjects may well have other goals but this does not seem to be of first order importance in this setting. In particular social preferences do not seem to be a major source of the observed diversity of play.

Key words: cooperation; prisoner's dilemma; social preferences; dictator game; inequity aversion; survey.
JEL codes: C72, C91, D03.

1. Introduction

Understanding when and why people cooperate in social dilemmas is a key issue not just for economics but for all of the social sciences (as noted by e.g. Ahn et al 2001 and Gachter and Herrmann 2009). Here we focus on the infinitely (i.e. indefinitely) repeated prisoner's dilemma, where cooperation can be an equilibrium if future payoffs loom sufficiently large compared to the present. Laboratory experiments have shown that the overall fraction of subjects who cooperate depends on the payoff parameters, with cooperation being more prevalent when the returns to cooperation are higher and the future looms larger. Nonetheless, there is typically some cooperation even when cooperation is not an equilibrium, and some defection when cooperative equilibria exist. Moreover, there is substantial heterogeneity across subjects in a given treatment: Some may cooperate in most periods while others cooperate hardly at all. This raises the question of who these cooperators are, and whether they differ in other measurable characteristics from the subjects who do not cooperate.

Understanding the heterogeneity of play seems useful for understanding when cooperation will arise, and also for the debate about the role of other-regarding or "social" preferences in supporting cooperation. In particular, the data raise the question of whether the cooperators are motivated by more than just maximizing their own monetary payoff. Although other-regarding motivations clearly play an important role in generating cooperative behavior in some interactions, the extent to which they affect play in infinitely repeated games remains largely unknown.

To better understand the sources of heterogeneous play, we combine data on play in an infinitely repeated noisy prisoner's dilemma or "RPD" that was previously analyzed in Fudenberg et al. (forthcoming) with data from an additional dictator game played by the same subjects, and also with survey responses and demographic data.[1] First, we relate each subject's play in the RPD to their generosity in a dictator game (DG). Next, we investigate whether any of the commonly observed strategies can be explained by inequity aversion (Fehr and Schmidt 1999) or pure altruism. In addition, we use responses to survey questions to explore the motivations underlying cooperative play in the RPD, as well as to explore whether

[1] The earlier paper focused on the types of strategies subjects used, and on the effect of noise in the execution of intended actions. As we argued, such noise is typically present in repeated interactions outside of the laboratory, so incorporating it brings the lab situation closer to the field; the noisy execution also facilitates the identification of the subjects' strategies as e.g. even an agent who intends to always cooperate will sometimes defect "by mistake".

self-reported prosocial behavior outside the laboratory is a good indicator of experimental behavior in the RPD and DG. We also examine whether individual characteristics such as age, major, gender, belief in God and risk attitudes are useful in explaining heterogeneity.

In the RPD subjects could either cooperate or defect in each round, with a constant probability of continuing to another round, and a constant probability that each player's decision will be changed to the opposite. At the end of the last repeated game, subjects played a DG; in specific, they split $6 between themselves and a person who was not present at the same time ("a subject in a later experiment").[2] The returns to cooperation were varied across four different payoff specifications. While cooperation varied with the payoffs, giving in the DG did not.[3] Furthermore, we find that giving in the DG is not correlated with either playing C in the first period of the repeated game or the overall frequency of cooperation in the repeated game, except perhaps in the one "non-cooperative" treatment where cooperation is not an equilibrium. In addition, we find no correlation between DG giving and leniency (waiting for multiple defections before punishing) which is substantially more frequent when the returns to cooperation are high, and earns high payoff in these treatments; and we find no correlation between forgiveness (returning to cooperation after punishing) and DG, except in the non-cooperative treatment. We also relate DG giving to the distribution of strategies played, and find that players who are selfish in the DG are more likely to play "Always Defect" in the non-cooperative treatment, while selfish players are marginally significantly *less* likely to play always defect in the "cooperative" treatments where cooperation is an equilibrium. Thus social preferences as measured by DG giving seem to play a role in promoting cooperation only when cooperation is not supported by self-interest. When the monetary payoffs strongly support cooperation, DG giving has little explanatory power, and what power it may have suggests that in these cases selfishness promotes rather than inhibits cooperation.

We also explore the implications of social preferences for play in our RPD game through the use of the Fehr and Schmidt inequity aversion model (1999). In Fudenberg et al.

---

[2] We chose this order to avoid having the DG play influenced by play in the RPD. It does not appear that RPD play influenced DG play, since DG play was about the same in each treatment while play in the PD was not. We chose the DG because it is easy to implement and has been widely studied. As Cooper and Kagel (2009) emphasize, behavior in DG is known to be affected by adding third players, random moves, etc., but the framing we use is standard and has been widely studied, and other commonly used games such as the one-shot prisoner's dilemma and related public goods games are also sensitive to factors like framing, audience effects and communication. Moreover, sequential move games are harder to implement across sessions.

[3] We do not mean to suggest that subjects do not have social preferences. Rather, these results show that social preferences are not an important part of the reason for variation in cooperation across treatment, which seems to be primarily driven by strategic reasoning.

(forthcoming) we identified a number of strategies that performed well given the distribution of play, and also identified some others that were reasonably common despite performing poorly, most notably the strategy "Always Defect" in treatments where most agents appeared to be conditional cooperators. Here we investigate the expected utility of these strategies if subjects had utility as described by the inequity aversion model. We use a set of parameter values most often used by Fehr and Schmidt (see Fehr and Schmidt 2010), and find that the model gives the highest utility to subjects that always defect in the non-cooperative treatment, while giving the highest utility to a very infrequently played exploitive or 'suspicious' strategy in the cooperative treatments. Since maximizing money payoff also predicts "always defect" in the non-cooperative treatment, allowing a fraction of the population to have Fehr-Schmidt preferences does not help explain the data here, and since the "suspicious" strategy was rarely played in the other treatments it is unlikely to have had much impact on play there.[4] Moreover, the FS model gives very little utility to lenient strategies, which are common in the cooperative treatments and earn large monetary payoffs. Thus the FS model does not favor cooperative strategies, and is not successful at predicting the strategies played by subjects in the specifications which support cooperation. We also examine a simple altruistic preference where subjects derive some benefit from their partner's payoff. We find that although altruism can potentially explain the cooperation we observe at low payoff specifications, it too makes incorrect predictions when the returns to cooperation are large. Thus altruism is also not a good predictor of play in the specifications with cooperative equilibria.

Third, we analyze subjects' motivations for cooperating in the RPD. Subjects indicated how well various motivations (both self-interested and other-regarding) explain their cooperation decisions. We analyze the relationship between these motivations and cooperative play in the RPD. At the individual level, we find that across all payoff specifications, a large majority of subjects reported maximizing their long-term payoff as a more important motivator of playing cooperatively than either a desire to increase their partner's payoff, to do the morally right thing or to avoid upsetting their partner. At the aggregate level, we find that the desire to maximize payoff was a more consistent predictor of RPD cooperation than any of the other motivations. We also assess the role of subjects' beliefs about the intentions of others, and find that subjects who are more inclined to attribute

[4] In addition, the best response to suspicious TFT is itself rarely played.

unprovoked defections to error are more cooperative, but that DG giving is not predictive of this tendency to give the opponent the benefit of the doubt.

Fourth, we examine the correlation between behavior that is observed in the experiments and that is self-reported in survey questions related to the domains of benevolence and universalism. Answers to these survey questions have been previously related to both how spouses/partners and peers answer these questions on behalf of the subjects' behaviors, as well as to benevolence and universalism values (Bardi and Schwarz 2003). However, we find that these questions do not predict experimental behavior in the RPD, except for in the non-cooperative treatment where there is some evidence of a negative correlation between cooperation and these measures. There is, however, evidence of a positive correlation between DG giving and benevolence.

Finally, we explore whether specific individual characteristics are correlated with experimental behavior. Both descriptive measures and the MLE estimation suggest that women are less cooperative than men, and economics majors seem to cooperate less than non-majors. We find no gender difference in DG giving. The other individual characteristics explored have no consistent relation to the various measures of cooperation. This suggests that individual characteristics may have some role, but perhaps not a very substantial one, in explaining heterogeneity in RPD play.

As far as we know, this is the first paper that correlates behavior in the RPD and DG while also linking social psychology survey questions with behavior in both games.[5] Harbaugh and Krause (2000) is perhaps the most related previous paper; they had subjects (children) first play a finitely repeated public goods game and then a modified DG, and they find that DG giving is correlated with first-round contributions but not last-round contributions, although their sample in this treatment is less than 30 subjects. Blanco et al (2011) find no correlation between play in the DG and play in a one-shot public goods game (PGG) but do find a positive correlation between the DG and second-mover play in a sequential PD; it is not clear how to extrapolate from their results to the RPD.[6] Our use of

[5] See Camerer and Fehr (2002) for a discussion of the advantages of experiments over surveys to measure social preferences.

[6] Several papers have run the DG before the main treatment instead of afterwards. Cabrales et al. (2010) had subjects first play a symmetric simultaneous-move version of a DG and then participate in a stylized labor market. They find that heterogeneous social preferences are a significant determinant of choice, and that uncertainty-aversion is a stronger determinant of choices than fairness. Ambrus and Pathak (forthcoming) also had subjects first play a DG and then a main treatment with a public goods game, but their design was quite different, as they matched subjects in the second round who made similar DG contributions, and they told subjects that this was the case. As opposed to using a DG in combination with repeated play, Cabral et al. (2010)

survey questions is related to previous studies linking experimental behavior to survey questions, where the focus has been on the trust game and trust attitudes (e.g., Glaeser et al. 2000, Fehr et al. 2003, Sapienza et al. 2007) or on cooperative play in one-shot cooperation games and trust attitudes (Ahn et al. 2003, Gächter et al. 2004). The results thus far are mixed, with some papers finding that attitudinal trust questions are not good at predicting experimental behavior (Glaeser et al. 2000, Ahn et al. 2003) whereas others find that they are (Fehr et al. 2003, Gächter et al. 2004).[7] In the setting of the DG, Carpenter et al. (2008) find that the specific survey questions for altruism used in their study are positively correlated with DG giving.

There have been several past studies on the correlation of individual characteristic variables and cooperation. Economics majors have been found to cooperate significantly less in the one-shot (Frank et al. 1993, Dal Bó 2005) and fixed-length (Dal Bó 2005) PD; in the RPD, however, (Dal Bó 2005, Dreber et al. 2008) the effect goes in the opposite direction, with economics majors cooperating more, although whether the relationship is significant depends on details of the regression.[8] Evidence on the importance of gender for cooperation is mixed (surveyed in Croson and Gneezy 2009), as is role of socio-economic variables (e.g., Glaeser et al. 2000 find positive results in a trust game (TG), whereas Gächter et al. 2004 find no correlation with play in a one-shot PGG. A recent meta-analysis of the DG, however, found that women give more and are thus more prosocial than men, and that older individuals give more than younger individuals (Engel forthcoming).[9] There are also a number of studies that explore the relationship between religiosity and cooperation in versions of the PGG and TG, with mostly negative results (Orbell et al. 1992, Anderson and Meller 2009, Anderson et al. 2010 Paciotti et al. forthcoming).

---

and Reuben and Seutens (forthcoming) examine whether subjects are selfish in a repeated game by varying whether the subjects know the current round of an interaction is the last. Cabral et al. test for and reject a specific model of backwards-looking reciprocity; Reuben and Seutens classify subjects as selfish/reputation building, strong reciprocators, unconditional defectors or unconditional cooperators by also letting subjects condition on whether the opponent cooperated or defected. Both Cabral et al. and Reuben and Seutens conclude that the majority of subjects are selfish.

[7] Sapienza et al. (2008) explain part of the mixed results with the fact that trust has both a belief-based component and a preference based one, where survey questions like the one used in the World Value Survey captures mainly the former component and questions about past trusting behavior mainly the latter. Sapienza et al. argue that since the type of subject pool differs between the studies so does variation in the former component and thus the correlation results; Glaeser et al. (2000) and Sapienza et al. (2008) used students (undergraduates vs. MBAs) as their subjects whereas Fehr et al. (2003) used a representative sample in Germany.

[8] Dreber et al. (2008) report that there is no significant difference between economics majors and non-economics majors in each condition separately. A reanalysis of the data combining the observations from all conditions and using a set of condition dummies, clustering on session, does find a significant positive effect of economics major on cooperation (coeff=0.131, p=0.028).

[9] See Camerer (2003), Cárdenas and Carpenter (2008) and Engel (forthcoming) for reviews of the dictator game.

2. Experimental setup

The purpose of the experiment was to explore the motivations for cooperation in the RPD by correlating cooperativeness in the RPD with giving in the DG while varying the returns to cooperation, looking at the predictive power of other-regarding preference models for play in the RPD, and correlating experimental behavior with self-reported motivations for cooperative play and pro-social behaviors outside the lab, as well as individual characteristics.

Subjects were recruited through the Computer Lab for Experimental Research (CLER) at Harvard Business School, to come to the Harvard Decision Science Laboratory in Cambridge, MA. We analyze the behavior of 278 subjects[10], mainly undergraduate students from schools in the Boston metro area, who participated in our experiments between September 2009 and October 2010. In each session, 12-32 subjects interacted anonymously via computer using the software Z-tree (Fischbacher 2007) when playing the RPDs as well as the DG.

Our experimental procedure has five components. First, subjects play a series of RPDs. Second, subjects play a DG. Third, subjects answer questions about their motivation to cooperate in the PD. Fourth, subjects answer attitudinal questions on benevolence and universalism. Finally, subjects answer a questionnaire in order to provide us with information on various individual characteristics, including age, gender, major, belief in God and risk attitudes.

2.1 Prisoner's dilemma

We use the data from an RPD with error originally reported in Fudenberg et al. (forthcoming). In this previous paper, the data was analyzed to determine what strategies people use in repeated games with error. In each round, subjects chose between cooperation (C) and defection (D). We used an 'equal gains from switching' formulation of the PD (as in Dreber et al. 2008), so that cooperation meant paying a cost of c units for the other to gain a benefit of b units, while defection was passive and led to 0 units for both players, where 30 units = \$1.[11] We fixed c=2, and considered 4 treatments in which b/c=1.5, b/c=2, b/c=2.5 and b/c=4. The probability of a subsequent round (i.e. continuation probability) was δ=7/8. When

[10] We focus our analysis on the 278 subjects in the main treatments of Fudenberg et al. (forthcoming). An additional 106 subjects participated in control experiments using different error rates.

[11] Although not all PDs can be described using this formulation, it allows one to easily vary the payoff to cooperation by adjusting a single parameter, the b/c ratio.

an interaction finished, subjects were rematched according to the turnpike protocol.[12] We introduced execution errors, so that with error probability E=1/8, an intended move was changed to the opposite move.[13] Subjects knew when their own move had been changed but not when the move of the other player had been changed, and the error probability, termination probability, and stage game payoffs were public information for the subjects in each session. Subjects were given a show-up fee of $10 plus their earnings from the RPD and a $6 DG (see below). On average subjects made $22 per session, with a range from $14 to $36. Sessions lasted approximately 90 minutes. See the online appendix 0-A1 for the experimental instructions.

In our current analysis, we focus on two different cooperation measures. First, we consider how often the subject cooperated in all rounds, indicating their overall cooperativeness. Second, we look at how often the subject cooperated in the first round of each interaction; this is independent of the cooperativeness of a subject's opponents, and so is an indication of whether the subject was playing a fundamentally cooperative or non-cooperative strategy. In addition to these two main measures of cooperation, we also consider two strategic features: leniency (waiting for multiple defections to punish) and forgiveness (being willing to return to cooperation after a punishment if the opponent cooperates). To minimize learning effects, we focus on decisions made in the last 4 interactions of each session.[14] Finally, we also use maximum likelihood estimation to calculate the probability weight assigned to each of 11 strategies by *a priori* interesting subsets of players, namely "altruistic" versus "selfish" players in the DG, men versus women, and economics majors versus non-economics majors.[15]

2.2 Dictator game

After the 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

---

[12] This matching protocol was proposed by Kamecke (1997) and implemented in the repeated game context by Dal Bó (2005)

[13] As controls, we also conducted two additional treatments where b/c=4 and either E=1/16 or E=0. These treatments are not the focus of this paper, as they did not provide enough data to be conclusive. Nonetheless, we provide an analysis of the E=0 condition in Appendix G to provide preliminary evidence about who cooperates in repeated games without error. We find largely equivalent results to those from the games with noise: DG giving has little predictive power for explaining cooperation in the RPD.

[14] In Fudenberg et al. (forthcoming) we showed that some learning occurred in earlier rounds of the RPD, and so we focused on the last 4 interactions, which is roughly a third of each session. We adopt the same convention here.

[15] The method we use was introduced by Dal Bó and Frechette (2011), and applied to our RPD data in Fudenberg et al. (forthcoming). To perform statistical tests, bootstrapped standard errors are used. For a full description of the maximum likelihood method, see the online appendix.

participant in the RPD 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. In our analysis, we use whether the subject gave or not as our main experimental measure of prosocial preferences ("DG giving"). We use the amount given in the dictator game as an additional measure ("DG transfer").

2.3 Motivations for cooperating in the prisoner's dilemma

To further explore motivations for cooperation, we had subjects complete a series of questions to elicit the motivation behind their play in the prisoner's dilemma. Subjects indicated the extent to which their motivation for cooperating following each outcome of the previous round (CC, CD, DC or 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.[16] See the online appendix 0-B1 for the motivations questions.

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." For each motivation (i) through (iv), the subject indicated a number between 1 and 7, where 1 is "not at all" and 7 is "very much so." This question in particular looks at the motivation for leniency, a strategic feature that was both common and successful in our treatments with cooperative equilibria.

In the current analysis, we first investigate the extent to which the self-interested motivation of (i) "earning the most points in the long run" is the strongest motivator for playing C, comparing (i) with the other motivations (ii)-(iv). We then look at the importance of each specific motivator across the four possible states in the previous round of the RPD, by making composite measures that are 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), and testing their importance in determining overall and first round cooperation.

2.4 Attitudinal questions on benevolence and universalism

After the behavioral experiments, subjects answered questions previously used in Bardi and Schwartz (2003) that concern prosocial behavior and values in the domains of benevolence and universalism. Here benevolence refers to behaviors that represent a

[16] For other ways of measuring the motivation for cooperation in the PD, see e.g., Ahn et al. (2003).

motivation to help and support individuals who are close to the subject, and universalism describes behaviors that represent a prosocial motivation towards others in general (i.e. not only for individuals close to the subjects).[17] In the analysis, we sum the scores that subjects gave to 10 questions for benevolence and 8 questions for universalism separately.

3. Results

See Appendix A for a summary of the variables used in the analysis. Pooling across treatments, 45% of subjects gave non-zero amounts in the dictator game, the modal transfer was \$0, and the mean transfer was \$1.07 out of \$6 (18% of the endowment). Comparing these results with the range of outcomes in the recent dictator game meta-analysis of Engel (forthcoming), our values are within the range of what is typically observed, although on the less generous end of the spectrum (25% of the 616 studies surveyed had mean transfers below 18% of the endowment). This is consistent with the finding of Engel (forthcoming) that experiments using student subject pools find significantly lower transfers.

Comparing across treatments, we find no significant differences in the distribution of DG transfers (Rank-sum, $p>0.10$ for all comparisons). This is in stark contrast to play in the RPD, which varies markedly across treatments. Therefore in our subsequent analysis, we assume that play in the post-experimental dictator game is independent of play in the RPD.

3.1 Prisoner's dilemma and dictator game correlations

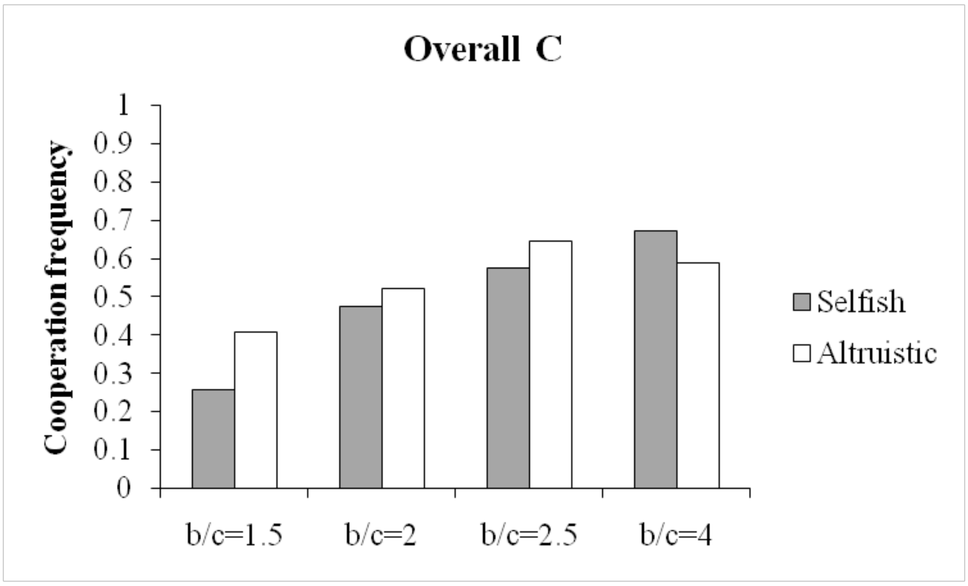
To test for correlations between RPD cooperation and social preferences as measured by giving in the DG, we run censored Tobit regressions on the frequency of cooperation, with a dummy variable for DG giving (a binary variable indicating whether the subject gave anything away or not) as independent variable, using robust standard errors clustered on session. These results are reported in Table 1. We also test the robustness of our DG results reported in Table 1 by using DG transfer (scalar number of dollars transferred to recipient) instead of the binary DG giving variable. These results are reported in Appendix Table B1. In

[17] These terms are commonly used in the psychological literature in connection with pro-sociality (e.g, Luk and Bond 1993, Kasser and Ahuvia 2002). Subjects used a Likert scale from 0-4 to indicate how often they have engaged in a number of behaviors in the last six months relative to their opportunities to do so, where 0 indicates "Never" and 4 indicates "All the Time". For example, one component of the benevolence scale is the frequency with which one "Help[s] out a colleague at work or school who made a mistake," while a component of the universalism scale is the frequency of "Donat[ing] money to alleviate suffering in foreign countries (e.g., hunger relief, refugee assistance)." See online appendix 0-C1 for all questions used to construct the benevolence and universalism scales.

the cases where using DG transfer give different results from using DG giving (i.e. comparing Table 1 and Appendix Table B1), we report this in a footnote.[18]

We begin by considering the most straightforward measure of play in the RPD, namely the frequency of overall cooperation across all rounds. There is reason to expect the relationship between overall cooperation and DG giving to be different in the b/c=1.5 treatment since this treatment has no cooperative equilibria. This expectation is correct, as seen in Figure 1. In regression analysis we thus analyze the relationship between overall cooperation and DG giving in the non-cooperative versus cooperative treatments separately (see Table 1). Consistent with the visual results, we find a significant positive relationship between overall cooperation and DG giving in the non-cooperative treatment (p=0.045).[19] Conversely, there is no significant relationship between these two variables in the cooperative treatments. In addition to a significant result in the non-cooperative treatment, the coefficient is substantially larger (almost 4 times the size), and the lack of significance in the cooperative treatments is not due to lack of power since the sample size is substantially larger than in the non-cooperative treatment. Thus it seems that when no cooperative equilibria exist, social preferences may play a role in the decision about whether or not to cooperate, but that at higher b/c ratios, DG giving is not predictive of overall cooperation.

Figure 1. Overall cooperation and DG giving.



[18] We do not report p-values greater than 0.10 in the text. See the regression tables in the Appendix for all coefficients and standard errors.

[19] This positive relationship is only marginally significant when looking at DG transfer (p=0.086).

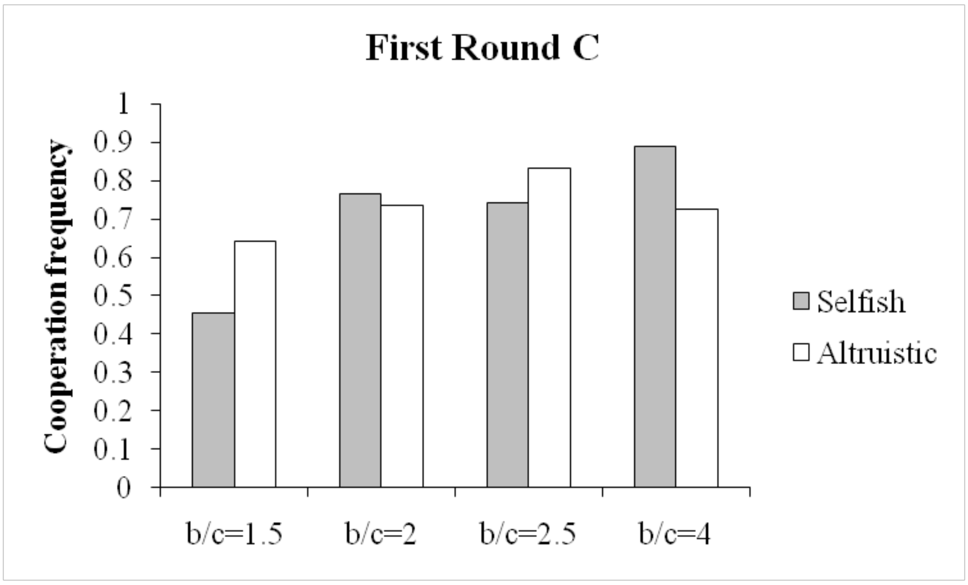
Table 1. Cooperation and DG giving.

	Overall C		First round C	
	b/c=1.5	b/c>1.5	b/c=1.5	b/c>1.5
DG giving	0.160**	0.0415	0.549	-0.00492
	(0.0785)	(0.0512)	(0.460)	(0.418)
Constant	0.242***	0.552***	0.421**	2.029***
	(0.0124)	(0.0454)	(0.178)	(0.556)
Observations	72	168	72	168

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

A subject's level of overall cooperation reflects the strategies of her partners as well as her own strategy. Cooperation in the first round of an interaction, however, depends only on the subject's strategy. Thus we next consider cooperation in the first round of each interaction. Figure 2 again indicates that the relationship between first round cooperation and DG giving may be different for the non-cooperative treatment. While we find no significant relationship between DG giving and first round cooperation in either the non-cooperative treatment or the cooperative treatments (see Table 1), the relationship between DG giving and cooperation in the non-cooperative treatment becomes significant (p=0.028) when including controls for the individual characteristics considered in the last section of the paper. It thus once again appears as if social preferences may play some role in choosing a cooperative strategy at the lowest b/c ratio, where cooperation is not an equilibrium and cooperative strategies do not earn high payoffs, but that in the payoff specifications where cooperation is payoff maximizing, social preferences are not predictive of playing a cooperative strategy.

Figure 2. First round cooperation and DG giving.



Fudenberg et al. (forthcoming) showed that "leniency"- the tendency for players to wait for multiple defections by their partner before retaliating- is common in the noisy RPD but rare when noise is completely absent.[20] There is considerable variation in the amount of leniency shown by different subjects, and it might be related to some forms of social preferences. However, in histories where cooperating this period corresponds to leniency (because the opponent played D in the previous round, and no previous D moves had occurred) there is no significant relationship between DG giving and cooperation, either considering non-cooperative and cooperative treatments separately or jointly. We also investigate forgiveness (returning to cooperation after punishing).[21] In histories with the possibility of forgiveness, we find a significant positive relationship between DG giving and cooperation in the non-cooperative treatment (p=0.043), but no relationship in the cooperative treatments (see Appendix Table B2).[22]

We now ask how the distribution of strategies employed differs based on DG giving. To do so, we use maximum likelihood estimation to calculate the probability weight for each

[20] Here we measure leniency as a conditional probability, by considering all histories s.t. both subjects played C in all but previous round, while in the previous round the other subject played D. For example: (C,C), (C,D), what does "C player" do next?

[21] To measure forgiveness, we examine all histories s.t. (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. We then ask how frequently this formerly cooperative subject showed forgiveness by returning to C.

[22] When considering DG transfer instead of the binary DG giving measure, there is no significant relationship with forgiveness in the non-cooperative treatment, and a significant positive relationship with forgiveness in the cooperative treatments (p=0.026) (see Appendix Table B3).

of the 11 strategies analyzed in Fudenberg et al. (forthcoming) for subjects who gave nothing in the DG compared to those who gave a non-zero amount. These 11 strategies are described in Table 2. Consistent with our previous analyses, Figure 3 shows that in the non-cooperative treatment, selfish players are more likely to play ALLD (p=0.016), but not in the cooperative treatments. Interestingly, we see the opposite pattern in the cooperative treatments: selfish players are marginally significantly *less* likely to play ALLD than players who make non-zero transfers in the DG (p=0.059)! This suggests that the selfish players (correctly) believed that cooperation was payoff maximizing in these treatments. Additionally, we see that selfish players are more likely to play the lenient and forgiving strategy TF2T than altruistic players in the cooperative treatments, although the difference is not statistically significant (p=0.109).

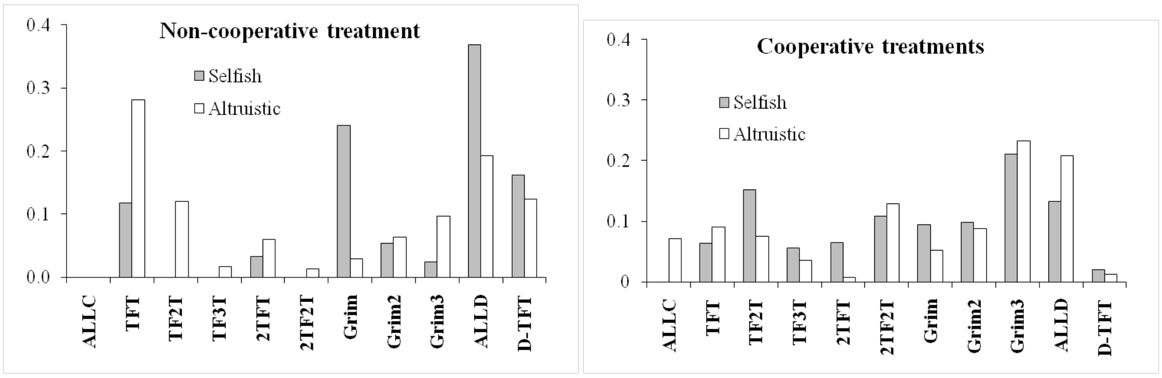
We can also use maximum likelihood estimation to calculate which strategy (or strategies) are most likely for each player by separately analyzing each individual's history of play.[23] We find a significant positive correlation between DG giving and playing ALLD in the non-cooperative treatment (p=0.022), and no significant relationship in the cooperative treatments (p=0.127), although the relationship is trending positive.

Table 2. Strategy descriptions.

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)
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
Always Defect	ALLD	Always play D
Exploitive Tit-for-Tat	D-TFT	Play D in the first round, then play TFT

[23] See the online Appendix for a description.

Figure 3. Strategy frequencies by DG giving.



Taken together, this analysis shows that DG behavior is important for explaining heterogeneous play in the non-cooperative treatment, that is the payoff specification in which no cooperative equilibria exist (and the least cooperative play occurs), but has little explanatory power in the treatments where cooperative equilibria exist.[24] To the extent that DG giving captures social preferences, we conclude that these preferences are neither necessary nor sufficient for explaining why we find high levels of cooperation in the treatments with cooperative equilibria.

3.2 Social preference models: Inequity aversion and pure altruism

Next we investigate the implications of social preference models for understanding play in the RPD game. Among the various models describing social preferences, we choose to apply the Fehr and Schmidt (henceforth FS) inequity aversion model (1999) to compute the expected utilities for the strategies identified in Fudenberg et al. (forthcoming).[25] In the FS model, subjects get disutility from unequal outcomes, i.e. both from having less as well as more than other subjects. While the FS model does not capture many important aspects of social preferences such as reciprocity, spite and efficiency concerns (e.g. Rabin 1993, Levine

[24] To further test if the DG has a different effect in the non-cooperative treatment, we regressed overall cooperation on DG giving, pooling the data from all 4 treatments together, and adding a dummy variable for the non-cooperative treatment, as well as an interaction between DG giving and that dummy. The interaction between DG giving and the non-cooperative treatment dummy is not significant, but it does become significant (p=0.019) when we also include controls for the individual characteristics considered in the last section of the paper. The results are similar when considering first round cooperation, where again there is no significant interaction without controls, but the interaction becomes significant (p=0.031) when including controls. We thus conclude that there is a real difference in the effect of the DG variable in the non-cooperative versus cooperative treatments.

[25] For other work linking experimental play to the FS model, see for example Bellemare et al. (2008). In a study on a representative Dutch sample playing the DG and the ultimatum game they find that inequity aversion seems to be a more important motivator in the general population than among students.

1998, Brandts and Sola 2001, Charness and Rabin 2002, Cox et al. 2008), and does not allow for a preference for ex-ante equality (see e.g. Bolton et al. 2005, Krawcyck and le Lec 2006, Fudenberg and Levine forthcoming), it is a parsimonious and widely used specification that is easy to implement and provides a straightforward basis of comparison to monetary payoff maximization.[26] Furthermore, the simplest versions of reciprocity seem unlikely to explain leniency (which is common in our data), as when the opponent deviates, reciprocity suggests retaliation and not forbearance.

We compare the FS inequity averse utility for each strategy analyzed in Fudenberg et al. (forthcoming), as described in Table 2, given the observed distribution of play, using $\alpha=2$ and $\beta=0.6$, where α measures the loss from disadvantageous inequity (i.e. when the opponent's money payoff exceeds the subject's) and β measures the loss from advantageous inequity.[27] To apply the FS model to the set of 11 strategies used by subjects in our experiment, we take the 11x11 payoff matrix, and for each payoff entry p(i,j) we calculate the FS payoff

$$p_{FS}(i,j) = p(i,j) - \alpha \max\left[p(j,i) - p(i,j), 0\right] - \beta \max\left[p(i,j) - p(j,i), 0\right].$$

We multiply the vector of observed strategy frequencies with the FS payoff matrix to get the expected payoff of each strategy. The results, as well as the expected payoffs based purely on monetary payoffs and the observed frequencies of each strategy, are displayed in Table 3.[28]

We see that the FS model assigns a low payoff to the lenient strategies, which are versions of Tit-for-tat, 2-tits-for-1-tatand Grim that wait for 2 (TF2T, 2TF2T, Grim2) or 3 (TF3T, Grim3) defections before punishing. Yet these lenient strategies were very common, and also earned high money payoffs. The lenient strategies obtain low FS payoffs because

[26] There is some debate about just how widely and accurately the FS model applies; see Binmore and Shaked (2010) and Fehr and Schmidt (2010).

[27] We focus on this particular pair of values as it seems to be most widely used (see Fehr and Schmidt 2010). We also consider three other parameter sets in Appendix C. We see qualitatively similar results, in that in the payoff specifications with high returns on cooperation, the strategies with highest FS utility are always less lenient and/or forgiving that those favored by monetary payoff maximization.

[28] Note that here we apply the FS preferences to the overall payoffs in the repeated game. This is consistent with past applications of FS preferences to sequential move games, and seems the natural specification for a repeated game. An alternative approach would be to apply the FS preferences to each period's outcome and then take the expectation of the corresponding sum. This has some odd features, such as penalizing "fair" alternation in a battle-of-the-sexes game, but since past referees have asked about it we carried out the corresponding analysis, which is reported in Appendix C. The results are somewhat different, but still add little explanatory power. The strategies favored by FS preferences (calculated by period) as similar to those favored by payoff maximization (the highest FS payoff strategy when calculated by period is ALLD at b/c=1.5 and b/c=2, and Grim2 at b/c=2.5 and b/c=4).

with certain less cooperative partners they were exploited. In terms of own monetary payoffs, these loses were outweighed by high payoffs received when playing other highly cooperative strategies. But because the FS model strongly penalizes disadvantageous inequity, the losses incurred against the exploitive strategies are amplified when calculating FS payoff.

The strategy which does best using FS payoff is very conservative and hesitant to cooperate. In all three treatments with cooperative equilibria, the strategy with the highest FS payoff is D-TFT (or 'suspicious TFT'). This strategy opens with D, and thereafter plays the action the other player used in the previous period. Although this makes some sense in the context of inequity aversion, it does not do a good job of explaining the observed play as this strategy had no more than 5% share in any of the three treatments where cooperation was common. Selfish payoff maximization (against the observed frequencies) does much better in directly explaining observed play, and the small share of D-TFT seen in the data is unlikely to have had much impact on play of other subjects (D-TFT was entirely absent in the cooperative treatments b/c=2 and b/c=4, and only observed at 5% at b/c=2.5).

Table 3. Frequencies, money payoffs and FS payoffs of observed strategies.

	b/c=1.5			b/c=2			b/c=2.5			b/c=4		
	Freq	Money payoff	FS Payoff	Freq	Money Payoff	FS Payoff	Freq	Money Payoff	FS Payoff	Freq	Money Payoff	FS Payoff
ALLC	0.00	-1.25	-28.68	0.03	6.92	-14.30	0.00	13.27	-8.14	0.06	28.13	-6.51
TFT	0.19	2.40	-3.71	0.06	8.71	3.87	0.09	14.64	9.38	0.07	29.01	19.90
TF2T	0.05	1.53	-11.33	0.00	8.69	-0.34	0.17	14.65	5.19	0.20	**29.67**	14.96
TF3T	0.01	0.90	-15.65	0.03	8.44	-3.47	0.05	14.53	2.08	0.09	29.56	9.88
2TF2T	0.00	1.86	-8.85	0.11	8.89	1.68	0.11	**14.72**	7.02	0.12	29.62	17.44
GRIM	0.14	3.02	-0.45	0.07	8.40	4.03	0.11	12.33	7.38	0.04	23.99	14.35
GRIM2	0.06	2.37	-4.12	0.18	9.03	4.42	0.02	13.98	8.69	0.05	27.90	18.21
GRIM3	0.06	1.79	-8.82	0.28	9.02	2.13	0.24	14.67	7.06	0.11	29.23	16.49
2TFT	0.06	2.87	-0.77	0.07	8.59	5.22	0.02	13.58	9.40	0.03	27.08	19.53
ALLD	0.29	**3.73**	**1.00**	0.17	8.53	2.65	0.14	11.33	4.32	0.23	21.04	9.76
D-TFT	0.15	2.89	-0.31	0.00	**9.19**	**5.31**	0.05	14.66	**9.93**	0.00	28.76	**20.90**

Another implication of the FS analysis is that in the non-cooperative treatment (when b/c=1.5), FS and self-interest both favor ALLD. Yet as reported above, we find some indication of a positive relationship between DG giving and cooperation in the non-cooperative treatment. This finding suggests an alternate social preference: simple altruism. To explore this possibility, we calculate the altruistic payoff of each strategy given the observed distribution of play. A strategy *i* earning money payoff p(i,j) against strategy *j* receives an altruistic payoff of

$$p_A(i,j) = p(i,j) + \gamma p(j,i)$$

where γ represents the extent to which the player values the partner's money payoff.[29] We find that a value of γ=0.22 can fairly well predict behavior in the non-cooperative treatment, where the uncooperative strategies ALLD and D-TFT are roughly as common as the cooperative (and non-lenient) strategies TFT, 2TFT and Grim, and all receive similar altruistic utilities. This altruistic preference, however, predicts too much cooperation when the returns to cooperation are high. In a cooperative treatment such as b/c=4, for example, the strategies with the highest altruistic utility are ALLC and TF3T, which only punishes following 3 Ds in a row, neither of which are frequently played. See Appendix D. Thus pure altruism also does not seem to do a good job of describing the data.

Together, these results provide further evidence that the cooperation in general, and the leniency in particular, observed in our data is primarily the result of strategic considerations rather than of social preferences.

3.3 Motivations to cooperate and survey questions

In studying the questions related to motivations for cooperation, we particularly focus on the extent to which the alternative (i) "earning the most points in the long run" is the best predictor of behavior, as opposed to the various other-regarding motivations (ii) through (iv).

We start by exploring the motivation for playing C in the four different states (CC, CD, DC, DD).[30] We find that for all four states, (i) is stronger than all other motivations, both in the non-cooperative treatment and the cooperative ones. (Appendix Tables E1-E4.) Specifically, in the non-cooperative treatment and the cooperative treatment respectively, 78% to 80% and 75% to 83% of subjects rated (i) higher than (ii), 69% to 80% and 72% to

[29] As Engelmann (forthcoming) points out, these "altruistic" preferences are equivalent to a concern for social efficiency.

[30] In each state we exclude those subjects that gave a 0% probability to playing C in that specific state.

80% rated (i) higher than (iii), and 78% to 88% and 74% to 86% rated (i) higher than (iv). Thus earning the most points in the long run thus seems to be the most important motivation for playing C for most players, across treatments and possible states of play.

To look at how each motivator predicts actual cooperation in the RPD, we made four composite measures, namely the sum of (i) over all four states, the sum of (ii), the sum of (iii), and the sum of (iv).[31] We regress overall cooperation and first round cooperation against all these composite cooperation motivations, for the non-cooperative treatment separately from the three cooperative treatments. We find that for the non-cooperative treatment, "earning the most points in the long run" is significantly positively correlated with overall cooperation ($p=0.009$) and first round cooperation ($p=0.028$).[32] In the analysis of the cooperative treatments, we again find that motivation (i) is significantly positively related to overall cooperation ($p=0.002$) and first round cooperation ($p=0.028$). The motivation "help the other person earn more points" is not significantly related to cooperation in the non-cooperative treatment or the cooperative treatments. The motivation "morally right thing to do" is significantly positively related to overall cooperation ($p=0.014$) in the treatments with cooperative equilibria and not related to cooperation in the non-cooperative treatment. Finally, the motivation "not wanting to upset the other person" is a significant positive predictor of overall cooperation in both the non-cooperative treatment ($p=0.013$) and the cooperative treatments ($p<0.001$) and marginally significantly positively related to first round cooperation in the non-cooperative treatment ($p=0.094$). Thus although several motivations appear to play a role, it seems that payoff maximization is the only motivation which is consistently predictive of cooperation across treatments and cooperation measures. Moreover, the effect of this motivation appears to be stronger when the returns to cooperation are higher.

In summary, these self-report measures complement the analysis of DG giving as well as that of FS and altruistic utility versus monetary payoff maximization. Both sets of analyses suggest that the desire to earn the most money is an important motivator of cooperation across payoff specifications.

We also assessed beliefs (albeit in an un-incentivized fashion) by asking subjects the extent to which they interpreted an opponent's D following a round of mutual cooperation as due to error rather than being intentional (using a 7 point Likert scale). In a regression analysis where the self-report measure is the independent variable, we find that this self-report

[31] We exclude those subjects that gave a 0% probability to playing C in any of the four states.
[32] See Appendix Table E5.

measure is significantly positively correlated with overall cooperation (p<0.001) and first round cooperation (p=0.004) in the cooperative treatments; unfortunately we did not include this question in version of the survey given to subjects in the non-cooperative treatment. We also use this measure to ask whether altruists are more inclined to give opponents the benefit of the doubt. Consistent with our previous analyses, we find no significant relationship between DG giving and this measure of attributing defection following mutual cooperation to error rather than intention.

The responses to the psychological survey do not suggest that social preferences play a key role in promoting cooperation in repeated games. Neither benevolence nor universalism are related to overall cooperation in either the cooperative or non-cooperative treatments, and moreover, both are significantly *negatively* correlated with first round cooperation in the non-cooperative treatment (p<0.001 and p=0.005 respectively). (Appendix Tables F1 and F2.) This latter result is surprising, since if anything we would have expected a positive correlation. There is however a positive significant correlation between DG giving and benevolence (p=0.021), and a marginally significant positive correlation with universalism (p=0.085).[33] (Appendix Table F3.) We conclude that these questions on self-reported prosocial behavior are not good predictors of experimental behavior in the RPD except for first round cooperation in the non-cooperative treatment, and this correlation with first round cooperation is contrary to what one would expect.[34] Further exploration of this counterintuitive finding is an interesting direction for future research. Interestingly, there is some evidence of correlations between in the psychological measures and DG giving.

3.4 Individual characteristics

In this section we further explore the possible determinants of the heterogeneity in RPD play by examining whether individual characteristics such as being female (0 or 1), being an economics major (0 or 1), age, belief in God (1-7 where a higher number indicates stronger belief) and risk attitudes (0-10 where a higher number indicates more risk taking) can predict cooperative play in the RPD.[35] The self-report general risk taking question used here has previously been explored by e.g. Dohmen et al. (2010), and has found to be a good predictor of a number of risk related activities as well as an incentivized risk task.

---

[33] For DG transfer, the correlation with benevolence is marginally significant (p=0.076) and the correlation with universalism is insignificant (p=0.128).

[34] An interesting extension for future work would be to compare the survey questions used here with the NEO personality inventory used in Carpenter et al. (2008), since that paper finds an association between DG giving and the NEO personality inventory for altruism.

[35] See the online appendix 0-D for the survey questions on individual characteristics.

In Table 4 we analyze the correlation between the individual characteristics and overall cooperation, as well as cooperation in the first round, for the non-cooperative treatment and the cooperative treatments separately. We find that women are significantly less cooperative than men overall in the cooperative treatments (p=0.028), and significantly less cooperative in the first round in both the non-cooperative treatment (p<0.001) and the cooperative treatments (although only marginally, p=0.051). The MLE strategy estimates on the two populations are consistent with this: In the non-cooperative treatment, women were marginally more likely to play D-TFT (p=0.082) and more likely to play ALLD (although the difference was not significant, p=0.184), while men were more likely to play TFT (p=0.004) and Grim (although only marginally, p=0.091), and in the cooperative treatments, women were significantly more likely to play ALLD (p=0.023) and the relatively unforgiving strategies Grim (p=0.018) and 2TF2T (p=0.022), while men were more likely to play ALLC (p=0.008).

There is also some evidence that economics majors cooperate less overall in the cooperative treatments (p=0.033), but are not less likely to cooperate in the first round. This suggests that economics majors are no less likely to choose cooperative strategies. However, even though the coefficient of economics major for first round cooperation in the cooperative treatments is not significant, the size of the coefficient is fairly large (as is the standard error), thus the lack of significance may simply reflect a relatively small sample of economics majors.

Age is not significantly related to cooperation, and there is some evidence that those with stronger believers in God are less likely to cooperate in the first round in the cooperative treatments (p=0.075). Risk attitudes are not uniformly related to cooperation: although there are significant relationships, they go in different directions depending on the treatment. Thus the relationship between risk attitude and cooperation in the RPD remains an open question.

Table 4. Cooperation and individual characteristics.

	Overall C		First round C	
	b/c=1.5	b/c>1.5	b/c=1.5	b/c>1.5
Female	-0.00976	-0.110**	-0.873***	-0.708*
	(0.0907)	(0.0495)	(0.133)	(0.361)
Economics major	-0.0765	-0.168**	-0.0994	-0.838
	(0.136)	(0.0782)	(0.760)	(0.641)
Age	0.0217	-0.00512	0.180	0.000636
	(0.0162)	(0.0125)	(0.108)	(0.0627)
Beliefs in God	-0.00439	-0.0125	0.0113	-0.151*
	(0.00550)	(0.00852)	(0.0568)	(0.0844)
Risk attitudes	0.0291***	-0.0239*	-0.0190	-0.212**
	(0.00166)	(0.0126)	(0.0981)	(0.0868)
Constant	-0.233	0.961***	-2.419	4.447**
	(0.208)	(0.290)	(2.391)	(1.882)
Observations	55	190	55	190

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

We also explore to what extent individual characteristics correlate with DG giving. Engel (forthcoming) in a meta-study finds that women are more altruistic than men in the DG, and there is also evidence suggesting that age is positively related to DG giving. Religious beliefs have been related to DG behavior in several other experiments, albeit with mixed results. Eckel and Grossman (2003) find that giving in a dictator game (DG) is positively correlated with religiosity, whereas Eckel and Grossman (2004) find no relationship. Tan (2006) finds that DG giving and ultimatum game behavior are not correlated with an overall measure of religiosity, but that that religious beliefs (a subset of religiosity) are positively correlated with DG giving. Paciotti et al. (forthcoming) find very little evidence of correlations between religiosity and behavior in the DG, the TG and the PGG.

Appendix Table G1 reports the results from regressing DG giving on individual characteristics. The only significant result is that belief in God is positively related to DG giving (p=0.018).[36]

In sum, we find some evidence that women cooperate less than men in terms of first round cooperation in the RPD, as well as overall cooperation in the cooperative treatments, while there is no gender difference in leniency or in DG giving; and that economics majors cooperate less overall in the cooperative treatments.

4. Discussion

There is typically substantial heterogeneity in play in the RPD. To gain insight into who cooperates in repeated games, we had the same subjects play a repeated prisoner's dilemma and a dictator game, computed payoffs of commonly used strategies under Fehr-Schmit and/or altruistic preferences, and related their play to their responses to a questionnaire on attitudes, motivations and individual characteristics. We find that in most cases, cooperators do not give more in the DG than defectors. We have previously shown that subjects cooperate considerably more in treatments with cooperative equilibria compared to treatments without cooperative equilibria (Fudenberg et al. forthcoming). Though there was substantial heterogeneity in strategies played, the most successful strategies in the former treatments were lenient, in not retaliating for the first defection. One reason for this variation could be that social preferences lead to more lenient play in the treatments with higher b/c. where some subjects cooperate or not for reasons that take other players' payoffs into account. However, we do not find evidence that DG giving is predictive of leniency. There is a positive correlation between DG giving and forgiveness (returning to cooperation after punishing) in the non-cooperative treatment but not in the cooperative treatments. Furthermore, we find that Fehr and Schmidt inequity aversion preferences give very little utility to cooperative, and in particular lenient, strategies, that the strategies favored by such preferences are too rarely played to have had much impact on cooperation by others, and that neither inequity aversion nor pure altruism are successful in predicting the strategies played by subjects in the specifications which support cooperation.

Instead, incomplete learning may be a better explanation of the considerable strategic diversity in our data. Consistent with this, numerous strategies have very close to the maximal

[36] This is however not significant for DG transfer, see Appendix Table G1. Although the coefficient on being female is not significant, it is positive, and in light of Engel's (forthcoming) meta-study we suspect that the coefficient would be significant in a larger sample.

monetary payoff. The main deviation from monetary payoff maximization in the cooperative treatments is the large fraction of subjects playing ALLD. We believe that the reason ALLD persists despite receiving low expected payoffs is 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. [37]

We also explore to what extent individual characteristics such as age, gender, economics major, beliefs in God and risk attitudes can explain the heterogeneity in RPD play. Our results suggest that women cooperate less than men when we look at first round or overall cooperation. 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.[38] 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 as well.[39] We also find that economics majors cooperate less overall than others in the cooperative treatments.

In sum, some subjects have social preferences, and social preferences seem to play a role when the RPD payoffs do not support cooperation. However numerous complementary methods of analysis provide convergent evidence that strategic considerations appear to be more important than social preferences when cooperative equilibrium exist: The observed heterogeneity of play does not correlate well with any of the proxies we used to measure social preferences. In the cooperative treatments, subjects who cooperate are primarily motivated by their own money earnings, and even those who do depart from payoff maximization by *not* cooperating do so for reasons uncorrelated with our social preference proxies.

---

[37] 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. The same logic does not seem to apply to FS payoffs and leniency. Lenient strategies earn low FS payoffs because of exploitation by defectors. Subjects using lenient strategies will observe some opponents who consistently defect despite the lenient player's cooperation. Thus the potential false belief here concerns something that occurs when using the given strategy. This is different from the case of ALLD, where the false belief concerns how opponents would respond if the subject changed their own play to cooperation.

[38] See Charness and Rustichini (forthcoming) for further study of gender and cooperation in different contexts.

[39] See also Eckel and Grossman (2008b) or Bertrand (2010) for a discussion of gender differences in social preferences and possible explanations.

Acknowledgements

We would like to thank Rob Boyd, Armin Falk, Guillaume Frechette, Simon Gächter, Magnus Johannesson, David Laibson, Mihai Manea, Julie Mortimer, and Andy Schotter for helpful comments, and Anat Bardi for the survey questions from Bardi and Schwartz (2003). NSF grants SES-064816 and SES-0951462 provided financial support. DGR is supported by a grant from the John Templeton Foundation.

References

Ahn, T.K., Ostrom, E., Schmidt, D., Walker, J., 2003. Trust in two-person games: game structures and linkages, in: Ostrom, E., Walker, J. (Eds.), Trust and Reciprocity. Interdisciplinary Lessons from Experimental Research. Russell Sage, New York.

Ambrus, A., Pathak, P.A., forthcoming. Cooperation over finite horizons: a theory and experiments. J. Pub. Econ..

Anderson, L.R., Mellor, J., 2009. Religion and cooperation in a public goods experiment. Econ. Letters 105, 58-61.

Anderson, L., Mellor, J., Milyo, J., 2010. Did the devil make them do it? The effects of religion in public goods and trust games. Kyklos 63, 163-175.

Bardi, A., Schwarz, S.H., 2003. Values and behavior: strength and structure of relations. Pers. Soc. Psych. Bull. 29, 1207-1220.

Bellemare, C., Kröger, S., van Soest, A., 2008. Measuring inequity aversion in a heterogeneous population using experimental decisions and subjective probabilities. Econometrica 76, 815-839.

Bertrand, M., 2010. New perspectives on gender, in: Ashenfelter, A., Card, D. (Eds.), Handbook of Labor Economics. Volume 4b., Elsevier, New York.

Binmore, K., Shaked, A., 2010. Experimental economics: where next? J. Econ. Behav. Org. 73, 87-100.

Blanco, M., Engelmann, D., Normann, H.T., 2011. A within-subject analysis of other regarding preferences. Games Econ. Behav. 72, 321-338.

Bolton, G., Brandts, J., Ockenfels, A., 2005. Fair procedures: evidence from games involving lotteries. Econ. J. 115, 1054-1076.

Brandts, J., Solà, C., 2001. Negative reciprocity in simple sequential games. Games Econ. Behav. 36, 136-157.

Cabral, L., Ozbay, E.Y., Schotter, A., 2010. Intrinsic and instrumental reciprocity: an experimental study. Working Paper.

Cabrales, A., Miniaci, R., Piovesan, M., Ponti, G., 2010. Social preferences and strategic uncertainty: an experiment on markets and contracts. Am. Econ. Rev. 100, 2261-2278.

Camerer, C.F., 2003. Behavioral Game Theory. Experiments in Strategic Interaction. Russell Sage Foundation, New York.

Camerer, C.F., Fehr, E., 2002. Measuring social norms and preferences using experimental games: a guide for social scientists. IEER Working Paper No. 97.

Cárdenas, J.-C., Carpenter, J., 2008. Behavioural development economics: lessons from field labs in the developing world. J. Dev. Stud. 44, 311-338.

Carpenter, J., Connolly, C., Myers, C.K., 2008. Altruistic behavior in a representative dictator experiment. Exp. Econ. 11, 282-298.

Charness, G., Rabin, M., 2002. Understanding social preferences with simple tests. Quart. J. Econ. 117, 817–869.

Charness, G., Rustichini, A., forthcoming. Gender differences in cooperation with group membership. Games Econ. Behav.

Luk, C.L., Bond, M.H., 1993. Personality variation and values endorsement in Chinese university students. Pers. Ind. Diff. 14, 429-437.

Cooper, D., Kagel, J.H., 2010. Other-regarding preferences: a selective survey of experimental results, in: Roth, A., Kagel, J.H., (Eds.) Handbook of experimental economics. Princeton University Press, Princeton.

Cox, J.C., Friedman, J.C., Sadiraj, V., 2008. Revealed altruism. Econometrica 76, 31-69.

Croson, R., Gneezy, U., 2009. Gender differences in preferences. J. Econ. Lit. 47, 1-27.

Dal Bó, P., 2005. Cooperation under the shadow of the future: experimental evidence from infinitely repeated game. Am. Econ. Rev. 95, 1591-1604.

Dohmen, T.J., Falk, A., Huffman, D., Schupp, J., Sunde, J., Wagner, G.G., 2010. Individual risk attitudes: new evidence from a large, representative, experimentally-validated survey. J. Euro. Econ. Assoc. 9, 522-550

Dreber, A., Rand, D.G., Fudenberg, D., Nowak, M.A., 2008. Winners don't punish. Nature 452, 348-351.

Eckel, C.C., Grossman, P.J., 2003. Rebates versus matching: does how we subsidize charitable contributions matter? J. Pub. Econ. 87, 681-701.

Eckel, C.C., Grossman, P.J., 2004. Giving to secular causes by the religious and nonreligious: an experimental test of the responsiveness of giving to subsidies. Nonprof. Volunt. Sec. Q. 33, 271-289.

Eckel, C.C., Grossman, P.J., 2008a. Sex and risk: experimental evidence, in: Plott, C., Smith, V., (Eds.) Handbook of Experimental Economics Results. Vol. 1, Elsevier, New York.

Eckel, C.C., Grossman, P.J., 2008b. Differences in the economic decisions of men and women: experimental evidence, in: Plott, C., Smith, V., (Eds.) Handbook of Experimental Economics Results. Vol. 1, Elsevier, New York.

Engel, C., forthcoming. Dictator games: a meta study. Exp.Econ..

Engelmann, D., forthcoming. How not to extend models of inequality aversion. J. Econ. Behav. Org.

Fehr, E., Schmidt, K., 1999. A theory of fairness, competition, and cooperation. Quart. J. Econ. 114, 817-868.

Fehr, E., Schmidt, K., 2010. On inequity aversion: a reply to Binmore and Shaked. J. Econ. Behav. Org. 73, 101-108.

Fehr, E., Fischbacher, U., von Rosenbladt, B., Schupp, J., Wagner, G.G., 2003. A nation-wide laboratory. Examining trust and trustworthiness by integrating behavioral experiments into representative surveys. Institute for Empirical Research in Economic, University of Zurich, Working Paper No. 141.

Fischbacher, U., 2007. z-Tree: Zurich toolbox for ready-made economic experiments. Exp. Econ. 10, 171–178.

Frank, R.H., Gilovich, T., Regan, D.T., 1993. Does studying economics inhibit cooperation? J. Econ. Persp. 7, 159–71.

Fudenberg, D., Levine, D.K., 1993. Self-confirming equilibrium. Econometrica 61, 523-546.

Fudenberg, D., Levine, D.K., forthcoming. Fairness and independence: an impossibility theorem. J. Econ. Behav. Org.

Fudenberg, D., Rand, D.G., Dreber, A., forthcoming. Slow to anger and fast to forget: leniency and forgiveness in an uncertain world. Am. Econ. Rev.

Glaeser, E.L., Laibson, D.I., Scheinkman, J.A., Soutter, C.L., 2000. Measuring trust. Quart. J. Econ. 115, 811-846.

Gächter, S., Herrmann, B., 2009. Reciprocity, culture and human cooperation: previous insights and a new cross-cultural experiment. Phil. Trans. R. Soc. B 364, 791-806.

Gächter, S., Herrmann, B., Thöni, B., 2004. Trust, voluntary cooperation, and socio-economic background: survey and experimental evidence. J. Econ. Behav. Org. 55, 505-531.

Kasser, T., Ahuvia, A., 2002. Materialistic values and well-being in business students. Euro. J. Soc. Psych. 32, 137-146.

Kamecke, U., 1997. Rotations: matching schemes that efficiently preserve the best reply structure of an one shot game. Internat. J. Game Theory 26, 409-417.

Krawcyck, M., Le Lec, F., 2006. Social decisions under risk: evidence from the probabilistic dictator game. Working Paper.

Levine, D.K., 1998. Modeling altruism and spite in experiments. Rev. Econ. Dyn. 1, 593-622.

Orbell, J., Goldman, M., Mulford, M., Dawes, R., 1992. Religion, context and constraint towards strangers. Rat. Soc. 4, 291–307.

Paciotti, B., Richerson, P., Baum, B., Lubell, M., Waring, T., McElreath, R., Efferson, C., Edsten, E., forthcoming. Are religious individuals more generous, trusting, and cooperative? An experimental test of the effect of religion on prosociality. Res. Econ. Anthro.

Rabin, M., 1993. Incorporating fairness into game theory and economics. Am. Econ. Rev. 83, 1281-1302.

Reuben, E., Seutens, S., forthcoming. Revisiting strategic versus non-strategic cooperation. Exp. Econ.

Sapienza, P., Simats, A.T., Zingales, L., 2008. Understanding trust., CEPR Discussion Paper No. 6462.

Tan, J.H.W., 2006. Religion and social preferences: an experimental study. Econ. Letters 90, 60–67.

Appendix A – Summary table

Table A. Summary table of means. Standard deviations in parenthesis.

	b/c = 1.5	b/c>1.5
First round C	0.54 (0.44)	0.76 (0.39)
Overall C	0.32 (0.24)	0.57 (0.30)
Leniency	0.28 (0.40)	0.63 (0.42)
Forgiveness	0.15 (0.18)	0.38 (0.33)
DG giving (fraction of subjects that gave)	0.44 (0.50)	0.45 (0.50)
DG transfer ($)	1.1 (1.64)	1.04 (1.36)
Benevolence	28.87 (4.43)	27.56 (4.74)
Universalism	15.20 (4.51)	15.39 (5.13)
Max payoff*	19.96 (6.68)	22.72 (6.45)
Help*	9.19 (5.37)	11.20 (6.66)
Moral*	10.48 (6.36)	12.07 (7.22)
Upset*	7.94 (4.73)	10.75 (6.49)
Female^	0.5 (0.50)	0.53 (0.50)
Economics major^	0.17 (0.38)	0.14 (0.35)
Age (years old)	20.55 (2.37)	21.00 (2.84)
Beliefs in God	5.68 (2.92)	5.50 (2.88)
Risk attitudes	5.68 (2.17)	5.92 (2.17)

*Motivations.
^Female=1 if female, 0 if male. Economics major=1 if economics major, 0 otherwise.

Appendix B – Correlations between cooperation measures and DG transfer

Table B1. Cooperation and DG transfer.

	Overall C		First round C	
	b/c=1.5	b/c>1.5	b/c=1.5	b/c=1.5
DG transfer	0.0472*	0.0173	0.116	0.0154
	(0.0271)	(0.0150)	(0.133)	(0.111)
Constant	0.260***	0.552***	0.527***	2.011***
	(0.00971)	(0.0387)	(0.141)	(0.537)
Observations	72	168	72	168

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table B2. Leniency, forgiveness and DG giving.

	Leniency		Forgiveness	
	b/c=1.5	b/c>1.5	b/c=1.5	b/c>1.5
DG giving	0.125	0.155	0.192**	0.0996
	(0.220)	(0.226)	(0.0923)	(0.0893)
Constant	-0.573**	1.010***	-0.0194	0.314***
	(0.253)	(0.180)	(0.0747)	(0.0746)
Observations	56	134	49	132

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table B3. Leniency, forgiveness and DG transfer.

	Leniency		Forgiveness	
	b/c=1.5	b/c>1.5	b/c=1.5	b/c>1.5
DG transfer	0.0255	0.111	0.0474	0.0518**
	(0.0599)	(0.0846)	(0.0328)	(0.0230)
Constant	-0.541***	0.971***	0.0134	0.305***
	(0.185)	(0.203)	(0.0794)	(0.0624)
Observations	56	134	49	132

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Appendix C – Fehr Schmidt payoffs for different parameter values

The original Fehr Schmidt (1999) paper compared the self-interested α=0, β=0 parameter set to three different inequity averse parameter sets (α=0.5, β=0.25; α=1, β=0.6; α=4, β=0.6), while subsequent papers considered just the parameter set α=2, β=0.6. For parsimony our main analysis uses the latter parameter set. Here we show the FS payoffs for each strategy in our data using the other three parameter sets. Although the results differ across parameter sets, they are qualitatively similar in that in the specifications with large returns on cooperation, the strategies with highest FS payoff are less lenient and/or forgiving than the strategies with the highest monetary payoffs.

Table C1. Low inequity aversion: α=0.5, β=0.25.

	b/c=1.5			b/c=2			b/c=2.5			b/c=4		
	Freq	Money Payoff	FS Payoff	Freq	Money Payoff	FS Payoff	Freq	Money Payoff	FS Payoff	Freq	Money Payoff	FS Payoff
ALLC	0.00	-1.25	-8.11	0.03	6.92	1.61	0.00	13.27	7.92	0.06	28.13	19.47
TFT	0.19	2.40	0.84	0.06	8.71	7.40	0.09	14.64	**13.15**	0.07	29.01	26.44
TF2T	0.05	1.53	-1.69	0.00	8.69	6.43	0.17	14.65	12.28	0.20	**29.67**	25.98
TF3T	0.01	0.90	-3.24	0.03	8.44	5.47	0.05	14.53	11.42	0.09	29.56	24.64
2TF2T	0.00	1.86	-0.82	0.11	8.89	7.09	0.11	**14.72**	12.79	0.12	29.62	**26.55**
GRIM	0.14	3.02	1.97	0.07	8.40	6.87	0.11	12.33	10.54	0.04	23.99	20.60
GRIM2	0.06	2.37	0.71	0.18	9.03	7.78	0.02	13.98	12.51	0.05	27.90	25.19
GRIM3	0.06	1.79	-0.86	0.28	9.02	7.29	0.24	14.67	12.74	0.11	29.23	25.98
2TFT	0.06	2.87	1.85	0.07	8.59	7.54	0.02	13.58	12.21	0.03	27.08	24.66
ALLD	0.29	**3.73**	**2.59**	0.17	8.53	6.08	0.14	11.33	8.41	0.23	21.04	16.34
D-TFT	0.15	2.89	1.97	0.00	**9.19**	**7.87**	0.05	14.66	13.00	0.00	28.76	26.05

Table C2. Moderate inequity aversion: α=1, β=0.6.

	b/c=1.5			b/c=2			b/c=2.5			b/c=4		
	Freq	Money Payoff	FS Payoff	Freq	Money Payoff	FS Payoff	Freq	Money Payoff	FS Payoff	Freq	Money Payoff	FS Payoff
ALLC	0.00	-1.25	-14.97	0.03	6.92	-3.69	0.00	13.27	2.57	0.06	28.13	10.81
TFT	0.19	2.40	-0.73	0.06	8.71	5.99	0.09	14.64	**11.49**	0.07	29.01	**23.58**
TF2T	0.05	1.53	-4.90	0.00	8.69	4.17	0.17	14.65	9.91	0.20	**29.67**	22.29
TF3T	0.01	0.90	-7.38	0.03	8.44	2.49	0.05	14.53	8.31	0.09	29.56	19.72
2TF2T	0.00	1.86	-3.50	0.11	8.89	5.28	0.11	**14.72**	10.85	0.12	29.62	23.46
GRIM	0.14	3.02	0.75	0.07	8.40	4.90	0.11	12.33	8.20	0.04	23.99	16.23
GRIM2	0.06	2.37	-0.99	0.18	9.03	**6.44**	0.02	13.98	10.90	0.05	27.90	22.18
GRIM3	0.06	1.79	-3.53	0.28	9.02	5.54	0.24	14.67	10.78	0.11	29.23	22.66
2TFT	0.06	2.87	0.73	0.07	8.59	6.28	0.02	13.58	10.51	0.03	27.08	21.69
ALLD	0.29	**3.73**	**1.00**	0.17	8.53	2.65	0.14	11.33	4.32	0.23	21.04	9.76
D-TFT	0.15	2.89	0.92	0.00	**9.19**	6.21	0.05	14.66	10.85	0.00	28.76	22.58

Table C3. Strong inequity aversion: α=4, β=0.6.

	b/c=1.5			b/c=2			b/c=2.5			b/c=4		
	Freq	Money Payoff	FS Payoff	Freq	Money Payoff	FS Payoff	Freq	Money Payoff	FS Payoff	Freq	Money Payoff	FS Payoff
ALLC	0.00	-1.25	-56.11	0.03	6.92	-35.53	0.00	13.27	-29.56	0.06	28.13	-41.14
TFT	0.19	2.40	-9.67	0.06	8.71	-0.37	0.09	14.64	5.16	0.07	29.01	12.53
TF2T	0.05	1.53	-24.18	0.00	8.69	-9.36	0.17	14.65	-4.27	0.20	**29.67**	0.30
TF3T	0.01	0.90	-32.19	0.03	8.44	-15.38	0.05	14.53	-10.36	0.09	29.56	-9.80
2TF2T	0.00	1.86	-19.56	0.11	8.89	-5.50	0.11	**14.72**	-0.63	0.12	29.62	5.41
GRIM	0.14	3.02	-2.83	0.07	8.40	2.31	0.11	12.33	5.74	0.04	23.99	10.57
GRIM2	0.06	2.37	-10.40	0.18	9.03	0.39	0.02	13.98	4.27	0.05	27.90	10.27
GRIM3	0.06	1.79	-19.41	0.28	9.02	-4.69	0.24	14.67	-0.38	0.11	29.23	4.14
2TFT	0.06	2.87	-3.78	0.07	8.59	3.10	0.02	13.58	7.19	0.03	27.08	15.21
ALLD	0.29	**3.73**	**1.00**	0.17	8.53	2.65	0.14	11.33	4.32	0.23	21.04	9.76
D-TFT	0.15	2.89	-2.78	0.00	**9.19**	**3.52**	0.05	14.66	**8.10**	0.00	28.76	**17.53**

As argued in the text, we believe that applying FS inequity aversion to payoffs in the overall game is more appropriate than applying this function round by round. Since some readers have disagreed, we also examine the present value of payoffs where the FS utility function is applied to each round's outcome; using α=2 and β=0.6.

As with the FS payoffs in the text (calculated by game) and with payoff maximization, ALLD does the best at b/c=1.5.

At b/c=2, FS round-by-round favors ALLD whereas FS by-game and money maximization favor D-TFT. Neither of these strategies are cooperative, and neither are particularly common, although ALLD is substantially more common than D-TFT (which is entirely absent at b/c=2). Furthermore, although the highest money payoff strategy is D-TFT (9.13), the next highest are Grim2 (9.03) and Grim3 (9.02), which together account for almost half of the probability weight of observed play. Using FS round-by-round, on the other hand, the 2nd highest scorer is Grim, which receives a payoff only very slightly lower than ALLD, and these two strategies together account for less than one quarter of the observed probability weight. So it seems that money maximization gives a better account of the data.

At the higher b/c ratios, there is a qualitative difference between FS round-by-round and FS by-game, with by-game continuing to favor D-TFT and round-by-round instead favoring Grim2; while money maximization favors 2TF2T at b/c=2.5 and TF2T at b/c=4. In these more cooperative treatments FS round-by-round favors lenient cooperation strategies that also do well under payoff maximization. Thus these alternative payoff specifications are consistent with the data but do not add explanatory power.

Table C4. Fehr Schmidt payoffs calculated round-by-round, using α=4, β=0.6.

	b/c=1.5	b/c=2	b/c=2.5	b/c=4
ALLC	-35.9	-24.1	-20.6	-23.9
TFT	-19.2	-11.4	-9.4	-4.5
TF2T	-21.7	-13.1	-10.2	-6.6
TF3T	-24.7	-15.1	-11.9	-10.0
2TF2T	-19.9	-11.5	-9.1	-5.0
GRIM	-12.0	-7.8	-6.9	-4.6
GRIM2	-15.9	-8.4	-6.8	-2.8
GRIM3	-19.5	-10.6	-8.2	-4.9
2TFT	-14.0	-9.0	-8.2	-4.1
ALLD	-9.0	-7.8	-7.8	-6.8
D-TFT	-17.8	-11.8	-10.4	-5.2

Appendix D – Altruistic utilities

Here we show the frequencies, monetary payoffs and altruistic utilities earned by each strategy in Fudenberg et al. (forthcoming) using γ=0.22 (i.e. people value the other player's payoff 22% as much as their own).

Table D1. Altruistic utilities.

	b/c=1.5			b/c=2			b/c=2.5			b/c=4		
	Freq	Money Payoff	Altruistic Payoff	Freq	Money Payoff	Altruistic Payoff	Freq	Money Payoff	Altruistic Payoff	Freq	Money Payoff	Altruistic Payoff
ALLC	0.00	-1.25	1.49	0.03	6.92	10.77	0.00	13.27	18.55	0.06	28.13	38.12
TFT	0.19	2.40	3.52	0.06	8.71	10.88	0.09	14.64	17.94	0.07	29.01	35.56
TF2T	0.05	1.53	3.27	0.00	8.69	11.59	0.17	14.65	18.91	0.20	**29.67**	37.79
TF3T	0.01	0.90	2.91	0.03	8.44	11.61	0.05	14.53	**19.09**	0.09	29.56	**38.22**
2TF2T	0.00	1.86	3.44	0.11	8.89	11.63	0.11	**14.72**	18.79	0.12	29.62	37.41
GRIM	0.14	3.02	3.55	0.07	8.40	9.47	0.11	12.33	14.01	0.04	23.99	27.53
GRIM2	0.06	2.37	3.50	0.18	9.03	11.25	0.02	13.98	17.22	0.05	27.90	34.27
GRIM3	0.06	1.79	3.34	0.28	9.02	**11.73**	0.24	14.67	18.65	0.11	29.23	36.87
2TFT	0.06	2.87	**3.60**	0.07	8.59	10.26	0.02	13.58	16.09	0.03	27.08	32.33
ALLD	0.29	**3.73**	3.54	0.17	8.53	8.25	0.14	11.33	11.26	0.23	21.04	21.53
D-TFT	0.15	2.89	3.53	0.00	**9.19**	10.64	0.05	14.66	17.03	0.00	28.76	33.81

Appendix E– Motivations to cooperate

Table E1. Motivations for C after CC.

	(i)>(ii)	(i)>(iii)	(i)>(iv)
b/c=1.5	0.80	0.77	0.88
b/c>1.5	0.81	0.77	0.84

Table E2. Motivations for C after CD (leniency).

	(i)>(ii)	(i)>(iii)	(i)>(iv)
b/c=1.5	0.80	0.80	0.85
b/c>1.5	0.80	0.80	0.85

Table E3. Motivations for C after DC.

	(i)>(ii)	(i)>(iii)	(i)>(iv)
b/c=1.5	0.78	0.69	0.78
b/c>1.5	0.75	0.72	0.74

Table E4. Motivations for C after DD.

	(i)>(ii)	(i)>(iii)	(i)>(iv)
b/c=1.5	0.78	0.78	0.83
b/c>1.5	0.83	0.79	0.86

Table E5. Motivations for cooperation.

	Overall C		First round C	
	b/c=1.5	b/c>1.5	b/c=1.5	b/c>1.5
Payoff max	0.0265***	0.0176***	0.135**	0.0786**
	(0.00929)	(0.00571)	(0.0577)	(0.0352)
Help	-0.00405	-0.00217	-0.00540	-0.0531
	(0.00354)	(0.00280)	(0.0506)	(0.0475)
Moral	0.00155	0.00631**	-0.00390	0.0542
	(0.00682)	(0.00252)	(0.0288)	(0.0714)
Upset	0.0221**	0.0107***	0.100*	0.0409
	(0.00829)	(0.00225)	(0.0578)	(0.0339)
Constant	-0.381**	0.0415	-2.981**	-0.371
	(0.170)	(0.128)	(1.294)	(0.766)
Observations	29	126	29	126

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Appendix F– Correlations between cooperation measures, DG, benevolence and universalism

Table F1. Cooperation and benevolence.

	Overall C		First round C	
	b/c=1.5	b/c>1.5	b/c=1.5	b/c>1.5
Benevolence	0.00209	0.000975	-0.0261***	-0.00105
	(0.00284)	(0.00430)	(0.00387)	(0.0235)
Constant	0.252**	0.529***	1.407***	1.926***
	(0.0992)	(0.110)	(0.106)	(0.347)
Observations	72	204	72	204

Table F2. Cooperation and universalism.

	Overall C		First round C	
	b/c=1.5	b/c>1.5	b/c=1.5	b/c>1.5
Universalism	-0.000422	0.000693	-0.0409***	0.00597
	(0.00101)	(0.00563)	(0.0141)	(0.0370)
Constant	0.319***	0.548***	1.271***	1.815***
	(0.0362)	(0.0877)	(0.273)	(0.407)
Observations	72	205	72	205

Table F3. DG giving logit, DG transfer tobit, benevolence and universalism.

	DG giving		DG transfer	
Benevolence	0.0498**		0.0689*	
	(0.0217)		(0.0387)	
Universalism		0.0240*		0.0458
		(0.0139)		(0.0300)
Constant	-1.614***	-0.579*	-2.185*	-0.939
	(0.603)	(0.315)	(1.166)	(0.679)
Observations	238	239	238	239

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Appendix G – Correlations between DG giving, DG transfer and individual characteristics

Table G1. DG giving, DG transfer and individual characteristics.

	DG giving	DG transfer
Female	0.326	0.482
	(0.250)	(0.513)
Economics major	-0.574	-0.953
	(0.461)	(0.724)
Age	-0.00750	0.0192
	(0.0599)	(0.0754)
Belief in God	0.0829**	0.0916
	(0.0352)	(0.0563)
Risk attitudes	0.00658	0.0485
	(0.0604)	(0.105)
Constant	-0.618	-1.467
	(1.153)	(1.378)
Observations	211	211

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

NOT FOR PUBLICATION

Online Appendix - Who Cooperates in Repeated Games?

Appendix 0-A: Sample instructions for PD game

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	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 0-B – Motivations Questionnaire

In this part of the survey, think back through the decisions you made over the course of the session, and in the following questions try to characterize the way you made your choices.

1. Imagine that in the previous round, your action was A, and the other person's action was also A. How likely would you be to choose A this round (circle one)?

0/10 1/10 2/10 3/10 4/10 5/10 6/10 7/10 8/10 9/10 10/10

When you chose to play A in this situation, to what extent was it because
(circle number, where 1 is not at all and 7 is very much so)

(a) You thought it would earn you the most points in the long run
1 2 3 4 5 6 7
(b) You wanted to help the other person earn more points
1 2 3 4 5 6 7
(c) It felt like the morally right thing to do
1 2 3 4 5 6 7
(d) You felt like it would make the other person upset if you didn't
1 2 3 4 5 6 7

Free response for other motivations we didn't list:

When you chose to play B in this situation, to what extent was it because

(a) you thought it would earn you the most points in the long run
1 2 3 4 5 6 7
(b) You wanted to stop the other person from earning more points
1 2 3 4 5 6 7
(c) You wanted to punish the other person
1 2 3 4 5 6 7
(d) You wanted to earn more points than the other person
1 2 3 4 5 6 7

Free response for other motivations we didn't list:

2. Imagine that in the previous round, your action was A, and the other person's action was B.

How likely would you be to choose A this round (circle one)?

0/10 1/10 2/10 3/10 4/10 5/10 6/10 7/10 8/10 9/10 10/10

When you chose to play A in this situation, to what extent was it because
(circle number, where 1 is not at all and 7 is very much so)

(a) You thought it would earn you the most points in the long run
1 2 3 4 5 6 7
(b) You wanted to help the other person earn more points
1 2 3 4 5 6 7
(c) It felt like the morally right thing to do
1 2 3 4 5 6 7
(d) You felt like it would make the other person upset if you didn't
1 2 3 4 5 6 7

Free response for other motivations we didn't list:

When you chose to play B in this situation, to what extent was it because
(a) you thought it would earn you the most points in the long run
1 2 3 4 5 6 7
(b) You wanted to stop the other person from earning more points
1 2 3 4 5 6 7
(c) You wanted to punish the other person
1 2 3 4 5 6 7
(d) You wanted to earn more points than the other person
1 2 3 4 5 6 7

Free response for other motivations we didn't list:

3. Imagine that in the previous round, your action was B, and the other person's action was A.

How likely would you be to play A this round (circle one)?

0/10 1/10 2/10 3/10 4/10 5/10 6/10 7/10 8/10 9/10 10/10

When you chose to play A in this situation, to what extent was it because
(circle number, where 1 is not at all and 7 is very much so)

(a) You thought it would earn you the most points in the long run
1 2 3 4 5 6 7
(b) You wanted to help the other person earn more points
1 2 3 4 5 6 7
(c) It felt like the morally right thing to do
1 2 3 4 5 6 7
(d) You felt like it would make the other person upset if you didn't
1 2 3 4 5 6 7

Free response for other motivations we didn't list:

When you chose to play B in this situation, to what extent was it because
(a) you thought it would earn you the most points in the long run
1 2 3 4 5 6 7
(b) You wanted to stop the other person from earning more points
1 2 3 4 5 6 7
(c) You wanted to punish the other person
1 2 3 4 5 6 7
(d) You wanted to earn more points than the other person
1 2 3 4 5 6 7

Free response for other motivations we didn't list:

4. Imagine that in the previous round, your action was B, and the other person's action was also B.

How likely would you be to play A this round (circle one)?

0/10 1/10 2/10 3/10 4/10 5/10 6/10 7/10 8/10 9/10 10/10

When you chose to play A in this situation, to what extent was it because
(circle number, where 1 is not at all and 7 is very much so)

(a) You thought it would earn you the most points in the long run
1 2 3 4 5 6 7
(b) You wanted to help the other person earn more points
1 2 3 4 5 6 7
(c) It felt like the morally right thing to do
1 2 3 4 5 6 7
(d) You felt like it would make the other person upset if you didn't
1 2 3 4 5 6 7

Free response for other motivations we didn't list:

When you chose to play B in this situation, to what extent was it because
(a) you thought it would earn you the most points in the long run
1 2 3 4 5 6 7
(b) You wanted to stop the other person from earning more points
1 2 3 4 5 6 7
(c) You wanted to punish the other person
1 2 3 4 5 6 7
(d) You wanted to earn more points than the other person
1 2 3 4 5 6 7

Free response for other motivations we didn't list:

When the other person's action was B after a round when you had both played A, to what extent did you interpret the other person's action as intentional versus due to error?

(Intentional) (Error)
1 2 3 4 5 6 7

Please describe any aspects of your decisions and strategy in the experiment that were not captured by the questions above:

Appendix 0-C – Benevolence and Universalism Questionnaire

Behaviors Questionnaire Instructions:

In this questionnaire we are interested in common behaviors. The following pages list these behaviors. We would like you to estimate **how frequently you have engaged in each behavior** during the past 6 months. Think of how often you have engaged in each behavior **relative to your opportunities** to do so.

For example, consider the behavior described as "Say hello to my neighbours". Estimate how frequently you have said hello to your neighbours relative to the times you have seen your neighbours in the past 6 months.

Please use the following scale:

0	1	2	3	4
Never	Rarely	Sometimes	Frequently	All the Time

0 – I have never engaged in this behavior.
1 – I have engaged in this behavior in about one quarter of the times I had opportunities to do so.
2 – I have engaged in this behavior in about half of the times I had opportunities to do so.
3 – I have engaged in this behavior in more than half of the times I had opportunities to do so.
4 – I have engaged in this behavior every time I had an opportunity to do so.

<u>How frequently do I (fill in a number):</u>

1. Help out a colleague at work or school who made a mistake. _____
2. Donate money to alleviate suffering in foreign countries (e.g., hunger relief, refugee assistance). _____
3. Do my friends and family favors without being asked. _____
4. Use environmentally friendly products (e.g., recycled paper products). _____
5. Lend things to people I know (e.g., class notes, books, milk). _____
6. Make sure everyone I know receives equal treatment, even if I don't personally like him/her. _____
7. Keep promises I have made. _____
8. Take time to understand other people's world views. _____
9. Spend time with my friends when they are down to try to cheer them up. _____
10. Sign petitions to support environmental protection efforts. _____
11. Give small gifts to my friends and family for no reason. _____
12. Show my objections to prejudice (e.g., against racial groups, the homeless). _____
13. Forgive another person when they have hurt my feelings. _____
14. Actively support human rights causes through contributions, demonstrations, etc. _____
15. Emphasize the good qualities of other people when I talk about them. _____
16. Rejoice in the successes of others around me. _____
17. Participate in projects to protect the environment (e.g., beach clean-up). _____
18. Help my friends with school projects, moving, driving to the airport, etc. _____

Appendix 0-D – Individual Characteristics Survey

Your gender (circle one): Female Male

Your age: ______

Your major: ______

Your minor(s): _____________
____
____

Imagine you have just won $250,000 in the lottery. Almost immediately after you collect, you receive the following financial offer from a reputable bank, the conditions of which are as follows:

> You have a chance to double your money within two years. It is equally possible that you could lose half of the amount invested. That is, there is a 50% chance your investment will be doubled and 50% chance of your investment being halved.

What share of your lottery winnings would you be prepared to invest in this financially risky yet potentially lucrative investment? (Circle one)

$0
$25,000
$50,000
$75,000
$100,000
$125,000
$150,000
$175,000
$200,000
$225,000
$250,000

Highest level of education completed:

Less than a high school degree
High School Diploma
Vocational Training
Attended College
Bachelor's Degree
Graduate Degree

Do you believe in God?

(Confident atheist) 1 2 3 4 5 6 7 8 9 (Confident believer)

Are you generally a person who is fully prepared to take risks or do you try to avoid taking risks?

(Unwilling to take risks) 0 1 2 3 4 5 6 7 8 9 10 (Fully prepared to take risk)

Appendix 0-E – Maximum likelihood estimation method[40]

To assess the prevalence of each strategy in our data, we use the maximum likelihood estimate technique from Dal Bó and Frechette (2011) and Fudenberg et al. (forthcoming). Here we reproduce a portion of Fudenberg et al. (forthcoming) that describes the method.

We suppose that each subject chooses a fixed strategy for the last four interactions, and 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.[41] 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

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

where y_{ikr} is 1 if the subject chose C and 0 if the subject chose D.[42]

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.

[40] This subsection is copied verbatim from Fudenberg et al. (forthcoming); we include it here for the reader's convenience.

[41] 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.

[42] Thus the probability of an error in implementing one's strategy is $1/(1+\exp(1/\gamma))$. 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 of Fudenberg et al. (forthcoming) we show that the MLE estimates of strategy shares are robust to allowing each strategy have a different value of γ.

For any given set of strategies S and proportions p, we then derive the likelihood for the entire sample, namely $\sum_{I} \ln\left(\sum_{s \in S} p(s) p_i(s)\right)$. 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. 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.

This approach can also be used to calculate which strategy is most likely for an individual subject *i*. To do so, we evaluate $p_i(s)$ for each strategy *s*, using the value of estimated value of γ for the whole population (i.e. we assume all players in a given session have an equal error rate). We then assign subject *i* the strategy (or strategies) which have the largest value of $p_i(s)$. In our data it is often the case that multiple cooperative strategies are equally likely. However, for subjects where ALLD maximizes $p_i(s)$, ALLD is the unique maximizer (for our data and strategy set). Therefore we focus on using the MLE to indentify ALLD players.