

Contents lists available at [ScienceDirect](http://www.sciencedirect.com)

## Journal of Economic Behavior &amp; Organization

journal homepage: [www.elsevier.com/locate/jebo](http://www.elsevier.com/locate/jebo)

# Who cooperates in repeated games: The role of altruism, inequity aversion, and demographics



Anna Dreber<sup>a,\*</sup>, Drew Fudenberg<sup>b</sup>, David G. Rand<sup>c,d</sup>

<sup>a</sup> Department of Economics, Stockholm School of Economics, P.O. Box 6501, 113 83 Stockholm, Sweden

<sup>b</sup> Department of Economics, Harvard University, Littauer Center 310, 1805 Cambridge Street, Cambridge, MA 02138, USA

<sup>c</sup> Department of Psychology, Yale University, 2 Hillhouse Avenue, New Haven, CT 06511, USA

<sup>d</sup> Department of Economics, Yale University, 2 Hillhouse Avenue, New Haven, CT 06511, USA

## ARTICLE INFO

### Article history:

Received 7 June 2012

Received in revised form

10 December 2013

Accepted 13 December 2013

Available online 24 December 2013

### JEL classification:

C72

C91

D03

### Keywords:

Cooperation

Prisoner's dilemma

Altruism

Social preferences

Dictator game

Inequity aversion

Survey

## ABSTRACT

We explore the extent to which altruism, as measured by giving in a dictator game (DG), accounts for play in a noisy version of the repeated prisoner's dilemma. We find that DG giving 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 better explained by inequity aversion or efficiency concerns than money maximization. Various survey questions provide additional evidence for the relative unimportance of social preferences. We conclude that cooperation in repeated games is primarily motivated by long-term payoff maximization and that even though some subjects may have other goals, this does not seem to be the key determinant of how play varies with the parameters of the repeated game. In particular, altruism does not seem to be a major source of the observed diversity of play.

© 2013 Elsevier B.V. All rights reserved.

## 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., 2003](#); [Gächter 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 once they have some experience with the game depends on the payoff parameters, with cooperation being much more prevalent when the returns to cooperation are higher and the future looms larger (e.g., [Dal Bó and Frechette, 2013](#); [Rand and Nowak, 2013](#)). 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, if they

\* Corresponding author. Tel.: +46 87369640; fax: +46 8313207.

E-mail addresses: [anna.dreber@hhs.se](mailto:anna.dreber@hhs.se) (A. Dreber), [dfudenberg@harvard.edu](mailto:dfudenberg@harvard.edu) (D. Fudenberg), [david.rand@yale.edu](mailto:david.rand@yale.edu) (D.G. Rand).

differ in other measurable characteristics from the subjects who do not cooperate, and how such differences vary depending on the gains from cooperation.

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.

As a first step toward understanding the sources of heterogeneous play and the way subjects respond to changes in game parameters, we combine data on play in an infinitely repeated noisy prisoner’s dilemma or “RPD” that was previously analyzed in [Fudenberg et al. \(2012\)](#) with data from an additional dictator game played by the same subjects, and also with survey responses and demographic data.<sup>1</sup> First, we relate each subject’s play in the RPD to their generosity in a dictator game (DG). Next, we investigate whether accounting for inequity aversion ([Fehr and Schmidt, 1999](#)) or pure altruism does a better job of explaining the observed distribution of strategies than money maximization. In addition, we use responses to survey questions to explore the motivations underlying cooperative play in the RPD, as well as to explore whether 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 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. To reduce the influence of RPD play on the DG, we specified that the recipient would be a subject in a later experiment; this was easy to do with the DG but would have been more difficult to implement with a sequential-move game such as the ultimatum or trust games. Behavior in the DG is known to be affected by factors such as double blindness, adding third players, random moves, or expanded choice sets (e.g., [Hoffman et al., 1994](#); [List, 2007](#); [Bardsley, 2008](#); [Cooper and Kagel, forthcoming](#)). Nonetheless, DG giving has been shown to correlate with charitable giving (e.g., [Benz and Meier, 2008](#)) and returning money mailed to subjects in misaddressed envelopes months or years after the DG ([Franzen and Pointner, 2013](#)), suggesting that the DG does provide relevant information about altruistic preferences. Moreover, the DG is not the only game where behavior is sensitive to strategically incidental factors: behavior in other games commonly used to measure social preferences, such as the ultimatum game, the one-shot prisoner’s dilemma and related public goods games, can react to both priming and framing (e.g., [Lieberman et al., 2004](#); [Leliveld et al., 2008](#); [Benjamin et al., 2012](#); [Ellingsen et al., 2012](#); [Rand et al., 2013](#); [Rand and Engel, forthcoming](#)); and, at least in [Dreber et al. \(2013\)](#), DG giving is less influenced by framing effects than the Prisoner’s Dilemma.

The returns to cooperation in the RPD varied, with four different payoff specifications. While the frequency of cooperation varied also with the payoff specification, giving in the DG did not, which suggests that spillovers from the RPD to the DG were minimal. When we predict RPD cooperation with DG play, we find that an individual’s 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 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 giving, 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 altruism as measured by DG giving seems 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 one sort of social preferences for play in our RPD game through the use of the [Fehr and Schmidt](#) inequity aversion model (1999). 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, 1998](#); [Brandts and Solà, 2001](#); [Charness and Rabin, 2002](#); [Cox et al., 2008](#)), and does not allow for a preference for ex-ante equality (e.g., [Bolton et al., 2005](#); [Krawczyk and Le Lec, 2010](#); [Fudenberg and Levine, 2012](#)), it is a parsimonious and widely used specification that is easy to implement, readily yields concrete predictions, and provides a straightforward basis of comparison to monetary payoff maximization.<sup>2</sup>

To apply the FS model, we investigate the expected utility of the various strategies used in the experiment if subjects had utility as described by the inequity aversion model with parameters  $\alpha=2$ ,  $\beta=0.6$ , where  $\alpha$  measures the loss from disadvantageous inequity and  $\beta$  measures the loss from advantageous inequity. We chose these parameters because [Fehr and Schmidt \(2010\)](#) argue that many experiments are well summarized by supposing that some fraction of the population has these payoffs and the rest has “standard” payoffs  $\alpha=\beta=0$ . With these parameters, the highest utility goes to subjects

<sup>1</sup> Outside of the laboratory, actions are often observed with noise: someone who claims they worked hard, or that they were too busy or sick to help, may or may not be telling the truth, so incorporating noisy observations of intended actions brings the lab situation closer to the field. The presence of this noise also facilitates the identification of the subjects’ strategies as e.g. even an agent who intends to always cooperate will sometimes defect “by mistake”.

<sup>2</sup> There is some debate about just how widely and accurately the FS model applies; see [Binmore and Shaked \(2010\)](#) and [Fehr and Schmidt \(2010\)](#).

that always defect in the non-cooperative treatment, and 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 the preference that the FS model suggests does not help explain why some subjects continue to cooperate here. And since the "suspicious" strategy was rarely played in the other treatments it is unlikely to have had much impact on play of other subjects.<sup>3</sup> Moreover, the FS model gives relatively little utility to lenient strategies, which are common in the cooperative treatments and earn large monetary payoffs. Thus allowing for some subset of the population to have FS preferences does not yield better predictions in the cooperative treatments either, especially since the main deviation in observed play from money maximization was an excess of players using the strategy always defect. 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 in the low payoff specification, it too makes incorrect predictions (in this case, an excess of cooperation) when the returns to cooperation are large.

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 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 Schwartz, 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 distribution of strategies played suggest that women are less cooperative than men, and provide some evidence that economics majors cooperate less than non-economics majors. We find no gender or major differences 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. 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. There are two recent studies that explore the role of social versus selfish reasons for cooperation in repeated games. Cabral et al. (2010) and Reuben and Suetens (2012) examine whether subjects are selfish 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 Suetens 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 Suetens conclude that the majority of subjects are selfish. These results are in line with what we find.

Our use of 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., 2013) 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). 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 without execution errors, however, Dal Bó (2005) and Dreber et al. (2008) find that the effect goes in the opposite direction, with economics majors cooperating more. Evidence on the importance of gender for cooperation

<sup>3</sup> Consistent with this, the best response to suspicious TFT is itself rarely played.

**Table 1**  
Summary statistics.

	$b/c = 1.5$	$b/c = 2$	$b/c = 2.5$	$b/c = 4$
Sessions per treatment	3	2	3	4
Subjects per treatment	72	52	64	90
Average number of RPD interactions	11	11.5	10.7	11.3
Minimum number of RPD interactions	10	9	8	8
Maximum number of RPD interactions	13	14	13	15
Average number of rounds per RPD interaction	8.4	8.3	8.3	8.1

is mixed (surveyed in Croson and Gneezy, 2009), as is the 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 that older individuals give more than younger individuals (Engel, 2011).

## 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 from the main treatments of Fudenberg et al. (2012), 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 and risk attitudes.

### 2.1. Prisoner's dilemma

We use data from an RPD with execution error originally reported in Fudenberg et al. (2012). In the RPD, each subject played a stochastic number of rounds with a given opponent; when the current interaction ended, subjects were rematched according to the turnpike protocol (proposed by Kamecke, 1997 and implemented in the repeated game context by Dal Bó, 2005).<sup>4</sup> After each round of the interaction, the continuation probability was  $\delta = 7/8$ . 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 led to 0 units for both players, where 30 units = \$1. 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. We fixed  $c = 2$ , and considered 4 treatments in which  $b/c = 1.5$ ,  $b/c = 2$ ,  $b/c = 2.5$  and  $b/c = 4$ . We introduced execution errors, so that with error probability  $E = 1/8$ , an intended move was changed to the opposite move.<sup>5</sup> Subjects were informed of whether their own move had been changed but not whether 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. As shown in Fudenberg et al. (2012), the only Nash equilibrium in the treatment  $b/c = 1.5$  is "Always Defect." Each of the other treatments has equilibria with cooperative play in the first round, but because of the observation errors there are no equilibria in which subjects intend to cooperate after every sequence of observations. 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 min. See Table 1 for summary statistics and the Online Appendix 0–A 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

<sup>4</sup> Note that all subjects in a given session played the same number of interactions and faced the same sequence of game lengths. The number of interactions did vary across sessions, based largely on the number of subjects in session (using the turnpike protocol, the maximum number of interactions is equal to half the number of subjects). Importantly, however, the distribution of number of interactions did not vary significantly across  $b/c$  conditions (chi<sup>2</sup> test,  $p = 0.30$ ), or when comparing  $b/c = 1.5$  to the other  $b/c$  ratios (chi<sup>2</sup> test,  $p = 0.57$ ).

<sup>5</sup> 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, though there too DG giving has little predictive power for cooperation in the RPD.



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 (i.e., the last 4 repeated games played with the last 4 partners of the session).<sup>6</sup> Finally, we also use the “structural frequency estimation method” (SFEM) of Dal Bó and Frechette (2011) 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.

## 2.2. Dictator game

After the RPD, subjects played a dictator game where they were asked to divide \$6 between themselves and an anonymous recipient that was not a participant in the 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.<sup>7</sup> 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 their likelihood of cooperating following each outcome of the previous round (CC, CD, DC or DD), as well as the extent to which their motivation for cooperating 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. See the Online Appendix O-B for the motivations questions.

For example, subjects were given questions such as “Imagine that in the previous round, your action was C, and the other person's action was D. How likely would you be to choose C this round? When you chose to play C in this situation, to what extent was it because (i) You thought it would earn you the most points in the long run (ii) You wanted to help the other person earn more points (iii) It felt like the morally right thing to do (iv) You felt like it would make the other person upset if you didn't.” 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 predicting 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 motivation to help and support individuals who are close to the subject, and universalism describes behaviors that represent a prosocial motivation toward others in general (i.e., not only for individuals close to the subjects).<sup>8</sup> 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

<sup>6</sup> The effect of the game parameters on cooperation rates, which is predicted by equilibrium theory, is clearest when subjects have some experience with the game (Dal Bó and Frechette, 2011) and there is typically substantial learning in early rounds, which does not fit with equilibrium analysis. Consistent with this, in the noisy RPD considered here, some learning occurred in earlier rounds. For these reasons, the modern literature on experimental play of infinitely repeated games has focused on the last part of the session, and we follow this convention.

<sup>7</sup> These recipients were recruited at a later date using the online labor market Amazon Mechanical Turk (Horton et al., 2011) and given the corresponding money. (The experimental instructions did not specify that the “later subjects” would be in the laboratory, and though it seems likely that subjects presumed this was the case, we do not feel this was deceptive.)

<sup>8</sup> 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 to 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 the Online Appendix O-C for all questions used to construct the benevolence and universalism scales.

**Table 2**

Cooperation: Tobit regressions for DG giving (Panel A) and DG transfer (Panel B).

	Overall C		First round C	
	$b/c = 1.5$	$b/c > 1.5$	$b/c = 1.5$	$b/c > 1.5$
(A) Binary DG measure				
DG giving	0.160** (0.0785)	0.0415 (0.0512)	0.549 (0.460)	−0.00492 (0.418)
Constant	0.242*** (0.0124)	0.552*** (0.0454)	0.421** (0.178)	2.029*** (0.556)
(B) Continuous DG measure				
DG transfer	0.0472* (0.0271)	0.0173 (0.0150)	0.116 (0.133)	0.0154 (0.111)
Constant	0.260*** (0.00971)	0.552*** (0.0387)	0.527*** (0.141)	2.011*** (0.537)
Observations	72	168	72	168

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

endowment). Comparing these results with the range of outcomes in the recent dictator game meta-analysis of Engel (2011), 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).

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. Thus we do not find evidence of treatment-level differences in RPD behavior spilling over into the subsequent DG.<sup>9</sup> However, since the DG takes place after the RPD, an individual's outcome across the series of RPDs could still influence DG play through income effects, introducing a potential confound into our later analyses.<sup>10</sup> Even though subjects were not explicitly told their total payoff until they were paid at the end of the experiment, they could have kept track of it during the RPD (they were provided with the payoff from each individual interaction) and this in turn could have influenced their giving in the DG, as in Houser et al. (2012). In the  $b/c = 1.5$  treatment where there are no cooperative equilibria, we find a negative but non-significant relationship between the amount donated in the DG and total payoff from the RPD. Conversely, in the treatments where  $b/c > 1.5$  and cooperative equilibria do exist, we find a significant positive relationship between the amount donated in the DG and total RPD payoff ( $p = 0.004$ ).<sup>11</sup> We will revisit this relationship below and demonstrate that it does not undermine our findings regarding DG giving and cooperation.

### 3.1. Prisoner's dilemma and dictator game correlations

To test for correlations between RPD cooperation and altruism 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 the independent variable, using robust standard errors clustered on session.<sup>12</sup> These results are reported in Table 2. We also test the robustness of our DG results reported in Table 2 by using DG transfer (scalar number of dollars transferred to recipient) instead of the binary DG giving variable. In the cases where using DG transfer give different results from using DG giving (i.e., comparing Table 2, Panels A and B), we report this in a footnote.<sup>13</sup>

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 Fig. 1. In regression analysis we analyze the relationship between overall cooperation and DG giving in the non-cooperative versus cooperative treatments separately (see Table 2, Panel A). Consistent with the visual results, we find a significant positive relationship between overall cooperation and DG giving in the non-cooperative treatment ( $p = 0.045$ ).<sup>14</sup> Conversely, there is no significant relationship between these two variables in the cooperative treatments, and the coefficient on DG giving is almost 4 times smaller than in the non-cooperative treatment. Thus it seems that when no cooperative equilibria exist, altruism as measured by the DG 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.

<sup>9</sup> Whereas we find no spillover effects, Peysakhovich and Rand (2013) find large spillovers from the RPD to the DG. This may be because differences in long-run cooperation across treatments in our data were much smaller (cooperation in first period of last the 4 interactions: 54% in the treatment with no cooperative equilibria, 75–79% in the treatments with cooperative equilibria) than in Peysakhovich and Rand (2013) (roughly 15% versus 85% in the low versus high cooperation treatments).

<sup>10</sup> Total RPD payoff varies across individuals not only based on the outcome of the PD games, but also based on the total number of games played (which varied substantially across sessions within the same payoff specification).

<sup>11</sup> Tobit regression with DG transfer as the dependent variable and clustering robust standard errors on session.  $b/c = 1.5$ : coeff =  $-0.034$ ,  $p = 0.291$ ;  $b/c > 1.5$ : coeff =  $0.002$ ,  $p = 0.004$ . Using logit regression with the binary DG giving variable as the dependent variable gives qualitatively similar results.

<sup>12</sup> We note that there are only 3 sessions for  $b/c = 1.5$ , potentially limiting our statistical power. Despite this most of our significant results are obtained in this condition.

<sup>13</sup> We do not report  $p$ -values greater than 0.10 in the text.

<sup>14</sup> This positive relationship is only marginally significant when looking at DG transfer ( $p = 0.086$ ).

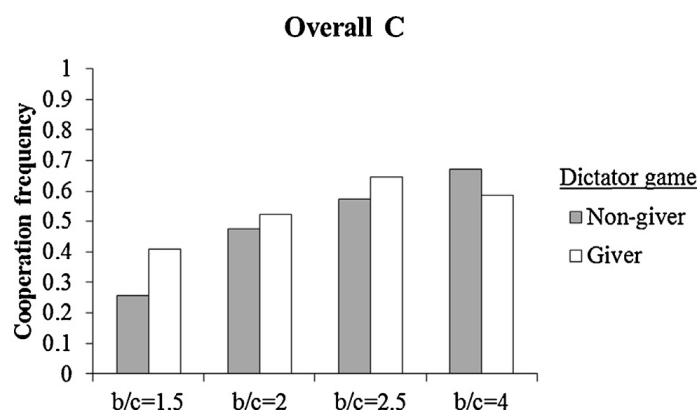


Fig. 1. Overall cooperation and DG giving.

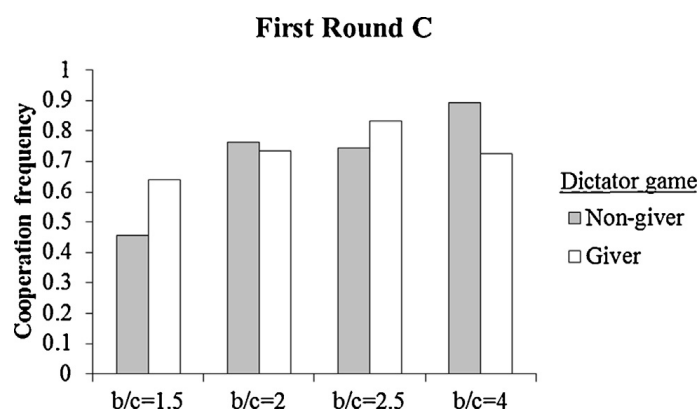


Fig. 2. First round cooperation and DG giving.

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, so we next consider cooperation in the first round of each interaction. Fig. 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 2, Panel A), 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 Section 3.4. It thus once again appears as if altruism may play some role in choosing a cooperative strategy (i.e., a strategy that begins with cooperation) at the lowest  $b/c$  ratio, where cooperation is not an equilibrium and cooperative strategies do not earn high payoffs, but that altruism is not predictive of playing a cooperative strategy in the payoff specifications where cooperation is payoff maximizing.

Importantly, these results are robust to controlling for total earnings in the RPD, and so are not confounded by income effects: When rerunning our main regressions from Table 2 including total payoff earned in the RPD, DG giving is positively correlated with overall cooperation ( $p = 0.010$ ) but not first round cooperation when  $b/c = 1.5$ , and is unrelated to either when  $b/c > 1.5$ .<sup>15</sup>

Fudenberg et al. (2012) 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.<sup>16</sup> 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).<sup>17</sup> In histories with the possibility of forgiveness, we find a significant positive relationship

<sup>15</sup> See Online Appendix Table 0-D1. Rerunning the regressions for DG transfer also gives no qualitative changes, see Online Appendix Table 0-D2 for DG transfer.

<sup>16</sup> Here we measure leniency as a conditional probability, by considering all histories s.t. both subjects played C in all but the previous round, while in the previous round the other subject played D. For example: (C, C), (C, D), what does "C player" do next?

<sup>17</sup> 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.

**Table 3**  
Leniency and forgiveness: DG giving (Panel A) and DG transfer (Panel B).

	Leniency		Forgiveness	
	<i>b</i> / <i>c</i> = 1.5	<i>b</i> / <i>c</i> > 1.5	<i>b</i> / <i>c</i> = 1.5	<i>b</i> / <i>c</i> > 1.5
(A) Binary DG measure				
DG giving	0.125 (0.220)	0.155 (0.226)	0.192** (0.0923)	0.0996 (0.0893)
Constant	−0.573** (0.253)	1.010*** (0.180)	−0.0194 (0.0747)	0.314*** (0.0746)
(B) Continuous DG measure				
DG transfer	0.0255 (0.0599)	0.111 (0.0846)	0.0474 (0.0328)	0.0518** (0.0230)
Constant	−0.541*** (0.185)	0.971*** (0.203)	0.0134 (0.0794)	0.305*** (0.0624)
Observations	56	134	49	132

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

**Table 4**  
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

between DG giving and cooperation in the non-cooperative treatment ( $p = 0.043$ ), but no relationship in the cooperative treatments (see Table 3).<sup>18</sup>

We now ask how the distribution of strategies employed differs based on DG giving. To do so, we use the SFEM to calculate the probability weight for each of the 11 strategies analyzed in Fudenberg et al. (2012) for subjects who gave nothing in the DG compared to those who gave a non-zero amount. These 11 strategies are described in Table 4. Consistent with our previous analyses, Fig. 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 the SFEM to calculate which strategy, or strategies, are most likely for each player by separately analyzing each individual's history of play.<sup>19</sup> We find a significant negative 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 in the opposite direction.

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.<sup>20</sup> 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.

<sup>18</sup> 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 Table 3, Panel B).

<sup>19</sup> See the Online Appendix O-E for a description.

<sup>20</sup> 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.



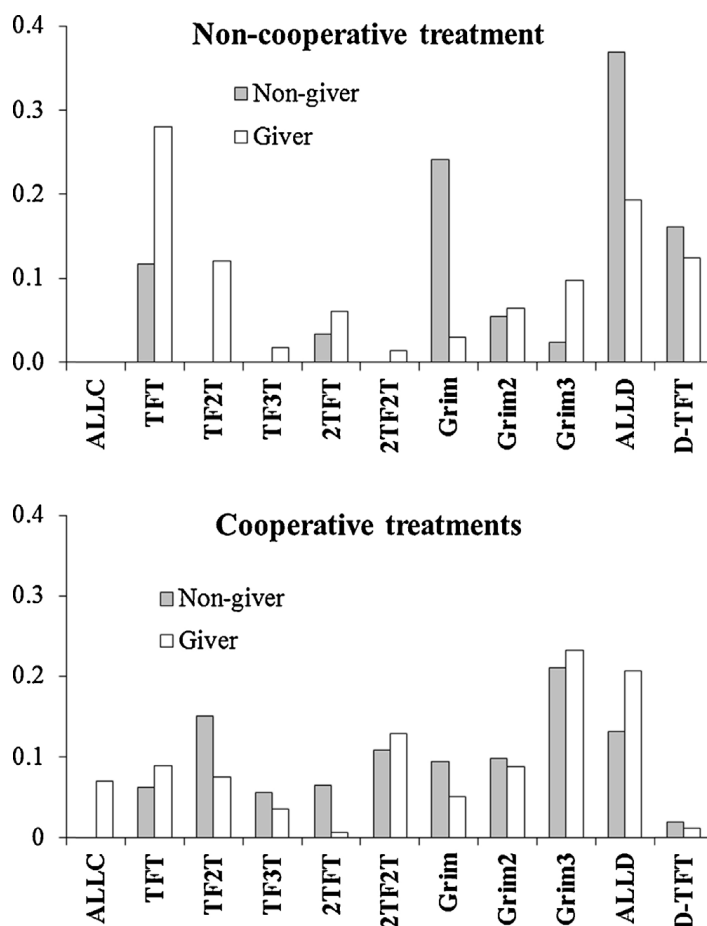


Fig. 3. Strategy frequencies by DG giving.

### 3.2. Inequity aversion and pure altruism

Next we investigate the implications of the Fehr and Schmidt (henceforth FS) inequity aversion model (1999) by computing the expected utilities for the strategies identified in Fudenberg et al. (2012).<sup>21</sup> In the FS model, subjects maximize their utility under correct beliefs about the distribution of opponents' strategies, but instead of caring only about their money payoff, subjects get disutility from unequal outcomes. As we noted earlier, the FS model does not capture many important aspects of social preferences; in particular, FS utility is based solely on the realized outcomes and does not depend on what outcomes might have occurred instead. Reciprocity-based models allow for this dependence, but they do not seem to yield crisp predictions in dynamic games. 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, as described in Table 4, given the observed distribution of play, using the parameters  $\alpha = 2$  and  $\beta = 0.6$  favored by Fehr and Schmidt (2010), where  $\alpha$  measures the loss from disadvantageous inequity (i.e., when the opponent's money payoff exceeds the subject's) and  $\beta$  measures the loss from advantageous inequity.<sup>22</sup> Fehr and Schmidt argue that in many studies, the data is fit relatively well by supposing that some subjects have the above preferences, the others are concerned only with their money payoffs ( $\alpha = \beta = 0$ ), and all of them have correct beliefs about the distribution of opponents' play; we check whether the same is true here.

To do this, we take the  $11 \times 11$  payoff matrix corresponding to the strategies used by our subjects, and for each payoff entry  $p(i, j)$  we calculate the FS payoff

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

We multiply the vector of observed strategy frequencies with the FS payoff matrix to get the expected payoff of each strategy. This corresponds to the equilibrium assumption that the subjects have correct beliefs about the strategies used

<sup>21</sup> 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.

<sup>22</sup> We also consider three other parameter sets in the Online Appendix O-F. 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 or forgiving than those favored by monetary payoff maximization.

**Table 5**

Frequencies, money payoffs and FS payoffs of observed strategies. Bootstrapped standard errors shown in italics. Best performing strategy is underlined. Strategies with payoffs not significantly different from the best (at the  $p < 0.05$ ) level shown in bold.

	<i>b/c</i> = 1.5			<i>b/c</i> = 2			<i>b/c</i> = 2.5			<i>b/c</i> = 4		
	Freq	Money payoff	FS payoff	Freq	Money payoff	FS payoff	Freq	Money payoff	FS payoff	Freq	Money payoff	FS payoff
ALLC	0.00 <i>0.00</i>	-1.25 <i>0.87</i>	-28.68 <i>3.78</i>	0.03 <i>0.03</i>	6.92 <i>1.21</i>	-14.30 <i>4.84</i>	0.00 <i>0.02</i>	<b>13.27</b> <i>1.35</i>	-8.14 <i>5.15</i>	0.06 <i>0.03</i>	<b>28.13</b> <i>2.04</i>	-6.51 <i>7.16</i>
TFT	0.19 <i>0.05</i>	2.40 <i>0.38</i>	-3.71 <i>1.02</i>	0.06 <i>0.04</i>	<b>8.71</b> <i>0.72</i>	<b>3.87</b> <i>1.53</i>	0.09 <i>0.04</i>	<b>14.64</b> <i>0.94</i>	<b>9.38</b> <i>1.70</i>	0.07 <i>0.03</i>	<b>29.01</b> <i>1.67</i>	<b>19.90</b> <i>2.71</i>
TF2T	0.05 <i>0.03</i>	1.53 <i>0.48</i>	-11.33 <i>1.52</i>	0.00 <i>0.00</i>	<b>8.69</b> <i>0.84</i>	-0.34 <i>2.32</i>	0.17 <i>0.06</i>	<b>14.65</b> <i>1.02</i>	5.19 <i>2.57</i>	0.20 <i>0.07</i>	<b>29.67</b> <i>1.73</i>	14.96 <i>3.84</i>
TF3T	0.01 <i>0.01</i>	0.90 <i>0.57</i>	-15.65 <i>2.01</i>	0.03 <i>0.03</i>	<b>8.44</b> <i>0.93</i>	-3.47 <i>2.92</i>	0.05 <i>0.05</i>	<b>14.53</b> <i>1.11</i>	2.08 <i>3.17</i>	0.09 <i>0.04</i>	<b>29.56</b> <i>1.81</i>	9.88 <i>4.62</i>
2TFT	0.06 <i>0.04</i>	2.87 <i>0.33</i>	-0.77 <i>0.75</i>	0.07 <i>0.04</i>	<b>8.59</b> <i>0.62</i>	<b>5.22</b> <i>1.06</i>	0.02 <i>0.02</i>	<b>13.58</b> <i>0.83</i>	<b>9.40</b> <i>1.17</i>	0.03 <i>0.02</i>	<b>27.08</b> <i>1.54</i>	<b>19.53</b> <i>2.01</i>
2TF2T	0.00 <i>0.02</i>	1.86 <i>0.44</i>	-8.85 <i>1.33</i>	0.11 <i>0.05</i>	<b>8.89</b> <i>0.80</i>	1.68 <i>2.07</i>	0.11 <i>0.06</i>	<b>14.72</b> <i>0.98</i>	<b>7.02</b> <i>2.30</i>	0.12 <i>0.05</i>	<b>29.62</b> <i>1.70</i>	<b>17.44</b> <i>3.47</i>
GRIM	0.14 <i>0.04</i>	3.02 <i>0.33</i>	-0.45 <i>0.68</i>	0.07 <i>0.05</i>	<b>8.40</b> <i>0.63</i>	<b>4.03</b> <i>0.91</i>	0.11 <i>0.04</i>	12.33 <i>0.71</i>	7.38 <i>0.93</i>	0.04 <i>0.02</i>	23.99 <i>1.43</i>	14.35 <i>1.53</i>
GRIM2	0.06 <i>0.03</i>	2.37 <i>0.40</i>	-4.12 <i>1.17</i>	0.18 <i>0.06</i>	<b>9.03</b> <i>0.74</i>	<b>4.42</b> <i>1.71</i>	0.02 <i>0.03</i>	<b>13.98</b> <i>0.88</i>	<b>8.69</b> <i>1.78</i>	0.05 <i>0.03</i>	<b>27.90</b> <i>1.58</i>	<b>18.21</b> <i>2.77</i>
GRIM3	0.06 <i>0.03</i>	1.79 <i>0.47</i>	-8.82 <i>1.54</i>	0.28 <i>0.08</i>	<b>9.02</b> <i>0.84</i>	<b>2.13</b> <i>2.27</i>	0.24 <i>0.07</i>	<b>14.67</b> <i>1.00</i>	<b>7.06</b> <i>2.44</i>	0.11 <i>0.04</i>	<b>29.23</b> <i>1.71</i>	<b>16.49</b> <i>3.65</i>
ALLD	0.29 <i>0.06</i>	<b>3.73</b> <i>0.29</i>	<b>1.00</b> <i>0.00</i>	0.17 <i>0.06</i>	<b>8.53</b> <i>0.67</i>	2.65 <i>0.07</i>	0.14 <i>0.04</i>	11.33 <i>0.70</i>	4.32 <i>0.11</i>	0.23 <i>0.04</i>	21.04 <i>1.49</i>	9.76 <i>0.37</i>
D-TFT	0.15 <i>0.05</i>	2.89 <i>0.35</i>	-0.31 <i>0.58</i>	0.00 <i>0.00</i>	<b>9.19</b> <i>0.77</i>	<b>5.31</b> <i>0.94</i>	0.05 <i>0.03</i>	<b>14.66</b> <i>1.02</i>	<b>9.93</b> <i>1.13</i>	0.00 <i>0.00</i>	<b>28.76</b> <i>1.74</i>	<b>20.90</b> <i>1.90</i>

by others, which is the usual way that predictions are obtained from FS preferences. The results, as well as the expected payoffs based purely on monetary payoffs and the observed frequencies of each strategy, are displayed in Table 5 along with bootstrapped standard errors.<sup>23</sup>

Fehr and Schmidt have argued that a population with two types of agents- some maximizing money payoffs and some with  $\alpha = 2$  and  $\beta = 0.6$ , can better explain a number of experimental results than the more parsimonious alternative that all agents maximize money payoffs. In contrast, we see that here FS preferences add little to explaining the experimental data. By and large, there are two ways in which money maximization is not consistent with the observed play. First, at  $b/c = 1.5$ , there is a substantial amount of cooperation, even though ALLD earns far more than any of the cooperative strategies. FS preferences do not help to explain these results, as ALLD also earns by far the highest FS payoff. The second deviation from money maximization involves defection rather than cooperation: when  $b/c > 1.5$ , ALLD earns substantially less than most cooperative strategies. Yet a considerable subset (between 14% and 23%, depending on the payoff specification) of subjects nonetheless consistently defect. FS preference do not help to explain this behavior either: as with money maximization, ALLD earns very a low FS payoff in these specifications, and the strategies which receive higher FS preferences also do well from a money maximization perspective.

We also note that FS preferences assign a low payoff to the lenient strategies, which are versions of Tit-for-Tat, 2-Tits-for-1-Tat and 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 with certain less cooperative partners they were exploited. In terms of own monetary payoffs, these losses were outweighed by high payoffs received when playing other highly cooperative strategies. But because FS preferences strongly penalize disadvantageous inequity, the losses incurred against the exploitive strategies are amplified when calculating FS payoff. Conversely, the strategy which does best using FS payoff is very conservative and hesitant to cooperate. In all three treatments

<sup>23</sup> 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 the Online Appendix O-F. The results are somewhat different, but still add little explanatory power as the strategies that have the highest expected sum of per-period FS preferences are similar to those that have the highest expected money payoff (the highest FS payoff strategy when inequity-aversion is applied period-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 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. This is such a small share that inequity aversion D-TFT players seem 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  $b/c=2.5$ ).

We conclude that selfish payoff maximization against the observed frequency distribution of strategies (that is, supposing all agents have  $\alpha = \beta = 0$ ) does as well as (or better than) the more flexible model that allows a fraction of the agents to have non-trivial inequity aversion.<sup>24</sup>

Another implication of the FS analysis is that in the non-cooperative treatment (when  $b/c=1.5$ ), FS preferences 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  $\gamma$  represents the extent to which the player values the partner's money payoff. As Engelmann (2012) points out, these “altruistic” preferences are equivalent to a concern for social efficiency. We find that a value of  $\gamma = 0.22$  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 the Online Appendix 0-G. 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 cooperative treatments can be parsimoniously explained as the result of strategic considerations.

### 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) (while excluding those subjects that indicated a 0% probability to playing C, as their responses to questions for motivations behind playing C are not meaningful).

We start by exploring the motivation for playing C in the four different states (CC, CD, DC, DD). We find that for all four states, (i) is stronger than all other motivations, both in the non-cooperative treatment and the cooperative ones. (Online Appendix Tables 0-H1-H4.) Specifically, in the non-cooperative treatment and the cooperative treatment respectively, 78–80% and 75–83% of subjects rated (i) higher than (ii), 69–80% and 72–80% rated (i) higher than (iii), and 78–88% and 74–86% rated (i) higher than (iv). Thus earning the most points in the long run 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). 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. (Online Appendix Table 0-H5.) 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 ( $p=0.028$ ). 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.

<sup>24</sup> It seems plausible that some of the agents did not have correct expectations; in particular the subjects who play ALLD in the cooperative treatments may have misperceived the prevalence of conditional cooperators. Allowing for these incorrect beliefs would improve the fit of the model but has no obvious relation to social preferences.

**Table 6**  
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	0.00117 (0.0964)	−0.123** (0.0497)	−0.740*** (0.203)	−0.819** (0.380)
Economics major	−0.108 (0.104)	−0.169** (0.0782)	−0.106 (0.665)	−0.936 (0.674)
Age	0.0214 (0.0167)	−0.00702 (0.0120)	0.140 (0.126)	−0.0197 (0.0626)
Risk attitudes	0.0257*** (0.000872)	−0.0258** (0.0123)	−0.000164 (0.0816)	−0.225*** (0.0850)
Constant	−0.239 (0.237)	0.948*** (0.271)	−1.802 (2.633)	4.206** (1.701)
Observations	59	193	59	193

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

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. Each set of analyses suggests 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 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 is 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). (Online Appendix Tables 0-I 1 and 0-I2.) 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 ( $p = 0.085$ ).<sup>25</sup> (Online Appendix Table 0-I3.) We conclude that these questions on self-reported prosocial behavior are not good predictors of experimental behavior in the RPD.

### 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, and attitudes toward risk (0–10 where a higher number indicates more risk taking) can predict cooperative play in the RPD.<sup>26</sup> 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.

In [Table 6](#) 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.<sup>27</sup>

First we consider the effect of gender. We find that women are significantly less cooperative than men overall in the cooperative treatments ( $p = 0.014$ ), and significantly less cooperative in the first round in both the non-cooperative treatment ( $p = 0.001$ ) and the cooperative treatments ( $p = 0.032$ ). The SFEM 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 ( $p = 0.032$ ), but are not less likely to cooperate in the first round. This suggests that economics majors are no less likely to choose cooperative strategies (i.e., strategies that open with cooperation), but instead play cooperative strategies which are less lenient or forgiving, or both. This result is consistent with the idea that (i) economic majors learned that cooperation is advantageous

<sup>25</sup> For DG transfer, the correlation with benevolence is marginally significant ( $p = 0.076$ ) and the correlation with universalism is insignificant ( $p = 0.128$ ).

<sup>26</sup> See the Online Appendix 0-J for the survey questions on individual characteristics.

<sup>27</sup> Some subjects did not complete all of the demographic questions, and the DG was not administered in the first experimental sessions. Thus the number of subjects in [Table 2](#) regressions differs from those in [Table 6](#).

in the RPD, and therefore choose cooperative strategies, but (ii) only learned about Grim, and thus predominantly use this non-lenient and non-forgiving strategy. However, we note that the size of the coefficient of economics major for first round cooperation in the cooperative treatments 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 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.

We also explore to what extent individual characteristics correlate with DG giving. Engel (2011) 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. Online Appendix Table O-K1 reports the results from regressing DG giving on individual characteristics. We find that women are marginally significantly more altruistic than men ( $p=0.071$ ), and otherwise nothing is significant.

#### 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–Schmidt and altruistic preferences, and related subjects' 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., 2012). 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. Here we investigated whether one reason for this variation could be that social preferences such as those we measure here 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. Contrary to this account, we do not find evidence that DG giving is predictive of leniency. 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. We restrict our analysis to comparatively simple hypotheses, so we cannot reject more complex sorts of social preferences such as those based on intentions, signaling, and reciprocity (e.g., Charness and Rabin, 2002; Battigalli and Dufwenberg, 2009; Levine, 1998) but then those more complex stories are typically so flexible they have little content in dynamic games. Still, it would be interesting to explore the correlations of play in other games with RPD play to evaluate the role of other sorts of social preferences such as spite.

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 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.<sup>28</sup>

To gain further insight into why some people cooperated “too much” in the non-cooperative treatment, and others cooperated “too little” in the cooperative treatments, future work could elicit players' beliefs about the distribution of others' strategies and motivations. Such beliefs data would directly inform us about which strategies players *thought* were payoff maximizing, and shed light on what assumptions about the preferences of others were driving those self-interest calculations.

In sum, some subjects have social preferences, as for example indicated by our data, and social preferences as measured by DG giving seem to play a role when the RPD payoffs do not support cooperation. However, as also indicated in field data by List (2006), 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 seem to be 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.

<sup>28</sup> 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.



## 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.

## Appendix A.

**Table A1**

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 <sup>a</sup>	19.96 (6.68)	22.72 (6.45)
Help <sup>a</sup>	9.19 (5.37)	11.20 (6.66)
Moral <sup>a</sup>	10.48 (6.36)	12.07 (7.22)
Upset <sup>a</sup>	7.94 (4.73)	10.75 (6.49)
Female <sup>b</sup>	0.5 (0.50)	0.53 (0.50)
Economics major <sup>b</sup>	0.17 (0.38)	0.14 (0.35)
Age (years old)	20.55 (2.37)	21.00 (2.84)
Risk attitudes	5.68 (2.17)	5.92 (2.17)

<sup>a</sup> Motivations.

<sup>b</sup> Female = 1 if female, 0 if male. Economics major = 1 if economics major, 0 otherwise.

## Appendix B. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.jebo.2013.12.007>.

## 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 Foundation, New York, pp. 323–351.
- Bardi, A., Schwartz, S.H., 2003. Values and behavior: strength and structure of relations. *Personality and Social Psychology Bulletin* 29, 1207–1220.
- Bardsley, N., 2008. Dictator game giving: altruism or artefact? *Experimental Economics* 11, 122–133.
- Battigalli, P., Dufwenberg, M., 2009. Dynamic psychological games. *Journal of Economic Theory* 144, 1–35.
- 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.
- Benjamin, D.J., Choi, J.J., Fisher, G., 2012. Religious identity and social behavior. Cornell University Working Paper.
- Benz, M., Meier, S., 2008. Do people behave in experiments as in the field? Evidence from donations. *Experimental Economics* 11, 268–281.
- Binmore, K., Shaked, A., 2010. Experimental economics: where next? *Journal of Economic Behavior and Organization* 73, 87–100.
- Blanco, M., Engelmann, D., Normann, H.T., 2011. A within-subject analysis of other regarding preferences. *Games and Economic Behavior* 72, 321–338.
- Bolton, G.E., Brandts, J., Ockenfels, A., 2005. Fair procedures: evidence from games involving lotteries. *Economic Journal* 115, 1054–1076.
- Brandts, J., Solà, C., 2001. Reference points and negative reciprocity in simple sequential games. *Games and Economic Behavior* 36, 138–157.
- Cabral, L., Ozbay, E.Y., Schotter, A., 2010. Intrinsic and instrumental reciprocity: an experimental study. New York University Working Paper.
- Carpenter, J., Connolly, C., Myers, C.K., 2008. Altruistic behavior in a representative dictator experiment. *Experimental Economics* 11, 282–298.
- Charness, G., Rabin, M., 2002. Understanding social preferences with simple tests. *Quarterly Journal of Economics* 117, 817–869.
- Cooper, D., Kagel, J.H. Other-regarding preferences: a selective survey of experimental results. In: Roth, A., Kagel, J.H. (Eds.), *Handbook of Experimental Economics*, vol. 2. Princeton University Press, Princeton (forthcoming).
- Cox, J.C., Friedman, D., Sadiraj, V., 2008. Revealed altruism. *Econometrica* 76, 31–69.
- Croson, R., Gneezy, U., 2009. Gender differences in preferences. *Journal of Economic Literature* 47, 448–474.
- Dal Bó, P., 2005. Cooperation under the shadow of the future: experimental evidence from infinitely repeated games. *American Economic Review* 95, 1591–1604.
- Dal Bó, P., Frechette, G., 2011. The evolution of cooperation in infinitely repeated games: experimental evidence. *American Economic Review* 101, 411–429.
- Dal Bó, P., Frechette, G., 2013. Strategy choice in the infinitely repeated prisoners dilemma. SSRN Working Paper 2292390.
- 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. *Journal of the European Economic Association* 9, 522–550.
- Dreber, A., Ellingsen, T., Johannesson, M., Rand, D.G., 2013. Do people care about social context? Framing effects in dictator games. *Experimental Economics* 16, 349–371.
- Dreber, A., Rand, D.G., Fudenberg, D., Nowak, M.A., 2008. Winners don't punish. *Nature* 452, 348–351.
- Ellingsen, T., Johannesson, M., Mollerstrom, J., Munkhammar, S., 2012. Social framing effects: preferences or beliefs? *Games and Economic Behavior* 76, 117–130.

- Engel, C., 2011. Dictator games: a meta study. *Experimental Economics* 14, 583–610.
- Engelmann, D., 2012. How not to extend models of inequality aversion. *Journal of Economic Behavior and Organization* 81, 599–605.
- Fehr, E., Schmidt, K.M., 1999. A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics* 114, 817–868.
- Fehr, E., Schmidt, K.M., 2010. On inequity aversion: a reply to Binmore and Shaked. *Journal of Economic Behavior and Organization* 73, 101–108.
- Fehr, E., Fischbacher, U., von Rosenblatt, B., Schupp, J., Wagner, G.G., 2003. A nation-wide laboratory. Examining trust and trustworthiness by integrating behavioral experiments into representative surveys. University of Zurich Working Paper No. 141.
- Fischbacher, U., 2007. z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics* 10, 171–178.
- Frank, R.H., Gilovich, T., Regan, D.T., 1993. Does studying economics inhibit cooperation? *Journal of Economic Perspectives* 7, 159–171.
- Franzen, A., Pointner, S., 2013. The external validity of giving in the dictator game. *Experimental Economics* 16, 155–169.
- Fudenberg, D., Levine, D.K., 1993. Self-confirming equilibrium. *Econometrica* 61, 523–545.
- Fudenberg, D., Levine, D.K., 2012. Fairness and independence: an impossibility theorem. *Journal of Economic Behavior and Organization* 81, 606–612.
- Fudenberg, D., Rand, D.G., Dreber, A., 2012. Slow to anger and fast to forget: leniency and forgiveness in an uncertain world. *American Economic Review* 102, 720–749.
- Glaeser, E.L., Laibson, D.I., Scheinkman, J.A., Soutter, C.L., 2000. Measuring trust. *Quarterly Journal of Economics* 115, 811–846.
- Gächter, S., Herrmann, B., 2009. Reciprocity, culture and human cooperation: previous insights and a new cross-cultural experiment. *Philosophical Transactions of the Royal Society B* 364, 791–806.
- Gächter, S., Herrmann, B., Thöni, B., 2004. Trust, voluntary cooperation, and socio-economic background: survey and experimental evidence. *Journal of Economic Behavior and Organization* 55, 505–531.
- Harbaugh, W.T., Krause, K., 2000. Children's contributions in public goods experiments: the development of altruistic and free-riding behaviours. *Economic Inquiry* 38, 95–109.
- Hoffman, E., McCabe, K.A., Shachat, K., Smith, V.L., 1994. Preferences, property rights, and anonymity in bargaining games. *Games and Economic Behavior* 7, 346–380.
- Horton, J.J., Rand, D.G., Zeckhauser, R.J., 2011. The online laboratory: conducting experiments in a real labor market. *Experimental Economics* 14, 399–425.
- Houser, D., Vetter, S., Winter, J., 2012. Fairness and cheating. *European Economic Review* 56, 1645–1655.
- Kasser, T., Ahuvia, A., 2002. Materialistic values and well-being in business students. *European Journal of Social Psychology* 32, 137–146.
- Kamecke, U., 1997. Rotations: matching schemes that efficiently preserve the best reply structure of a one shot game. *International Journal of Game Theory* 26, 409–417.
- Krawczyk, M., Le Lec, F., 2010. 'Give me a chance!' An experiment in social decision under risk. *Experimental Economics* 13, 500–511.
- Leliveld, M.C., van Dijk, E., van Beest, I., 2008. Initial ownership in bargaining: introducing the giving, splitting, and taking ultimatum bargaining game. *Personality and Social Psychology Bulletin* 34, 1214–1225.
- Levine, D.K., 1998. Modeling altruism and spite in experiments. *Review of Economic Dynamics* 1, 593–622.
- Liberman, V., Samuels, S.M., Ross, L., 2004. The name of the game: predictive power of reputations versus situational labels in determining prisoner's dilemma game moves. *Personality and Social Psychology Bulletin* 30, 1175–1185.
- List, J.A., 2006. The behavioralist meets the market: measuring social preferences and reputation effects in actual transactions. *Journal of Political Economy* 114, 1–37.
- List, J.A., 2007. On the interpretation of giving in dictator games. *Journal of Political Economy* 115, 482–493.
- Luk, C.L., Bond, M.H., 1993. Personality variation and values endorsement in Chinese university students. *Personality and Individual Differences* 14, 429–437.
- Peysakhovich, A., Rand, D.G., 2013. Habits of virtue: creating norms of cooperation and defection in the laboratory. SSRN Working Paper 2294242.
- Rabin, M., 1993. Incorporating fairness into game theory and economics. *American Economic Review* 83, 1281–1302.
- Rand, D.G., Engel, C. What does "clean" really mean? The implicit framing of decontextualized experiments. *Economics Letters* (forthcoming).
- Rand, D.G., Nowak, M.A., 2013. Human cooperation. *Trends in Cognitive Sciences* 17, 413–425.
- Rand, D.G., Dreber, A., Haque, O., Kane, R., Nowak, M.A., Coakley, S., 2013. Religious motivations for cooperation: an experimental investigation using explicit primes. *Religion, Brain and Behavior*, <http://dx.doi.org/10.1080/2153599X.2013.775664>.
- Reuben, E., Suetens, S., 2012. Revisiting strategic versus non-strategic cooperation. *Experimental Economics* 15, 24–43.
- Sapienza, P., Toldra-Simats, A., Zingales, L., 2013. Understanding trust. *Economic Journal*, <http://dx.doi.org/10.1111/eoj.12036>.