

Centre for Decision Research and Experimental Economics

Discussion Paper Series ISSN 1749-3293

CeDEx Discussion Paper No. 2008–02

Framing and Free Riding: Emotional Responses and Punishment in Social Dilemma Games

Robin Cubitt, Michalis Drouvelis and Simon Gaechter

March 2008





The Centre for Decision Research and Experimental Economics was founded in 2000, and is based in the School of Economics at the University of Nottingham.

The focus for the Centre is research into individual and strategic decision-making using a combination of theoretical and experimental methods. On the theory side, members of the Centre investigate individual choice under uncertainty, cooperative and non-cooperative game theory, as well as theories of psychology, bounded rationality and evolutionary game theory. Members of the Centre have applied experimental methods in the fields of Public Economics, Individual Choice under Risk and Uncertainty, Strategic Interaction, and the performance of auctions, markets and other economic institutions. Much of the Centre's research involves collaborative projects with researchers from other departments in the UK and overseas.

Please visit <u>http://www.nottingham.ac.uk/economics/cedex/</u> for more information about the Centre or contact

Karina Whitehead Centre for Decision Research and Experimental Economics School of Economics University of Nottingham University Park Nottingham NG7 2RD Tel: +44 (0) 115 95 15620 Fax: +44 (0) 115 95 14159 karina.whitehead@nottingham.ac.uk

The full list of CeDEx Discussion Papers is available at

http://www.nottingham.ac.uk/economics/cedex/papers/index.html

Framing and Free Riding: Emotional Responses and Punishment in Social Dilemma Games

Robin P. Cubitt, University of Nottingham

Michalis Drouvelis, University of Nottingham

Simon Gächter, University of Nottingham, CESifo and IZA

3 March 2008

Abstract

In this paper, we report an experimental investigation of the effect of framing on social preferences, as revealed in a one-shot linear public goods game. We use two indicators to measure social preferences: self-reported emotional responses; and, as a behavioural indicator of disapproval, punishment. Our findings are that, for a given pattern of contributions, neither punishment nor emotion depends on the Give versus Take framing that we manipulate. To this extent, they suggest that the social preferences we observe are robust to framing effects.

Keywords: framing effects, punishment, emotions, public goods experiments

JEL codes: C92, D01, H41.

Acknowledgements: We thank participants at the 2006 ESA European Meeting in Nottingham, the GATE Conference of the French Economic Association in Behavioural and Experimental Economics in Lyon, the XII Spring Meeting of Young Economists in Hamburg, the 2007 ESA World Meeting in Rome, the 2005 and 2006 CREED-CeDEx workshops, Joep Sonnemans, and Yuan Ju for useful comments. Financial support from the ESRC (PTA-030-2005-00608) and the University of Nottingham is gratefully acknowledged.

1. Introduction

The experimental literature has used a number of simple games to measure aspects of social preferences (Camerer and Fehr, 2004). Its findings have helped to inspire the development of new models of other-regarding preferences, such as those surveyed in Camerer (2003, ch. 2), which in turn have begun to be applied in other areas of economics (Gintis, Bowles, Boyd and Fehr, 2005). However, recent studies have cast doubt on the robustness of elicited social preferences to framing and contextual changes, especially in dictator games (e.g. Dana, Weber and Kuang, 2005; Haley and Fessler, 2005; Bardsley, 2007; List, 2007). Evidence from these studies suggests that behaviour is sensitive to seemingly irrelevant context-specific cues. If such a finding held more generally for the games used to measure social preferences in the laboratory, this would have serious implications for the external validity of the measures obtained, for the theories motivated by them, and for the explanation of subjects' behaviour.

In this paper, we examine the issue of frame-sensitivity in relation to the darker side of social preferences by studying a different experimental game that has played a major part in the literature: the public goods game with punishment. The framing manipulation we consider is of the contributions stage of the game and has a Give versus Take form, previously studied in the social psychology literature and similar to that introduced to economics by Andreoni (1995).¹ In one frame, subjects decide how much to contribute to a public good; in the other frame, they decide how much to withdraw from it. The setup is such that the two decision problems are objectively equivalent, in terms of the feasible set of allocations available to a subject. However, studies from social psychology found related framing manipulations to have an effect on contributions (e.g. Brewer and Kramer, 1986; McDaniel and Sistrunk, 1991; McCusker and Carnevale, 1995; Sell and Son, 1997; van Dijk and Wilke, 2000). Experimental economists have also identified framing effects, typically in a repeated game context (Andreoni, 1995; Sonnemans, Schram and Offerman, 1998; Willinger and Ziegelmeyer, 1999; Cookson, 2000 and Park, 2000).

Our study is distinctive in two main respects. First, we add a stage to the game in which subjects can punish one another and a further phase in which their self-reported

¹ For a conceptual discussion and classification of framing effects, see Levin, Schneider and Gaeth, 1998.

emotional responses are elicited. Second, unlike most of the economics literature, we study a one-shot game.

Rather than considering only the impact of framing on contributions to the public good, we focus mainly on whether it affects emotions and punishment behaviour. The fact that emotions and punishment have played a central role in the social preference literature provides a general motivation for this aspect of our study (Fehr and Gächter, 2000; Bosman and van Winden, 2002; Andreoni, Harbaugh, and Vesterlund, 2003; Masclet, Noussair, Tucker, and Villeval, 2003).

A more specific motivation stems from the findings of Cubitt, Drouvelis, Gächter and Kabalin (2008). They report an experiment in which subjects, playing the role of impartial observers, pass moral judgments on agents in hypothetical scenarios. The scenarios concerned a simple public goods game. Some were described using a Give frame and others using a Take frame; and subjects' judgments of the agents proved to be sensitive to this difference. On average, subjects tended to condemn a failure to contribute more strongly than a withdrawal of support with the same net consequences. However, as there may be a difference between judgments made in the role of impartial observer and responses to agents with whom one is interacting, it is important to investigate whether a parallel finding holds for subjects who are materially affected by each other's decisions. For this case, self-reported emotions are analogous to moral judgments to the extent of being non-behavioural indicators of attitudes; whereas punishment provides a behavioural indicator of willingness to act on one's attitudes.

We study a one-shot game for two reasons. First, as we explain in Section 2, oneshot games provide cleaner tests for framing effects than repeated games.² Second, and for similar reasons, one-shot games are of particular interest to the study of social preferences, as they prevent other-regarding considerations from being confounded with strategic ones arising from repetition of the game.

An important issue in studying the effect of the framing of the contribution stage on subsequent punishment behaviour and emotional responses is to recognise that the framing might affect punishment and emotions *either* because it affects the levels of contributions *or* because it affects how subjects see a given level of contributions.

 $^{^{2}}$ Conditional on the existence of a framing effect, it may be of interest to study the persistence or otherwise of the effect in repeat play. But, persistence is a separate issue from that of the initial existence of the effect. It is for study of the latter issue that we suggest one-shot games are superior.

We are most interested in the latter mechanism, since the former is already wellunderstood. Our results suggest that, on average, neither punishment nor emotion varies with the framing of the contribution stage, when one controls for the level of contributions. Thus, the social preferences we observe are robust to our framing manipulation.

The remainder of the paper is structured as follows. Section 2 describes the framework, the design and the hypotheses for our experiment. Section 3 discusses the results and Section 4 concludes.

2. Laboratory experiment with punishment stage

2.1 Framework

The basic building block of our study is a framing manipulation of a voluntary contributions game played by the members of a group. Our framings follow a similar formulation of the payoff function as in Dufwenberg, Gächter and Hennig-Schmidt (2006). In the Give frame, each player is endowed with 20 tokens and has to decide how many of them he keeps for himself and how many he contributes to the public good (described as a "project" to subjects). Each token kept for himself increases his own monetary payoff by one Guilder (our experimental currency). Each token contributed to the public good increases the payoff of every group member by 0.5 Guilders. The payoff function is given by equation (1).

$$\pi_i^1 = 20 - g_i + 0.5 \cdot \sum_{j=1}^n g_j \tag{1}$$

where π_i^1 denotes group member *i*'s payoff, g_i the number of tokens contributed to the public good by group member *i*, and *n* the number of group members.³

In the Take frame, there are initially $20 \cdot n$ tokens in the public good (described again as a "project" to subjects) for each group. Each group member has control of 20 tokens and has to decide how many of them he withdraws from the project and how many he leaves in it. The payoff function for this framing treatment is now given by equation (2).

$$\pi_i^1 = t_i + 0.5 \cdot (20 \cdot n - \sum_{j=1}^n t_j)$$
⁽²⁾

³ The reason for the superscript in group member i 's payoff emerges below.

where t_i indicates the number of tokens withdrawn from the public good by group member *i*.

In each frame, subjects have the same opportunities, regarding the final allocation of tokens. Equations (1) and (2) are equivalent using, for each player *i*, the relationship $g_i = 20 - t_i$. Thus, from a consequentialist perspective, the framing manipulation has no strategic significance.

2.2 Design and procedures

Our experimental design centres on a game with two stages. The first stage is a standard linear public goods game presented in two different framings, as described in Section 2.1; and the second is a punishment stage. We refer to the two resulting treatments as Give-P (Give frame for contributions, with punishment opportunities) and Take-P (Take frame for contributions, with punishment opportunities). In the second stage of the voluntary contributions game, players are allowed to punish each other. Each subject can assign up to five punishment points to each of the other group members. Punishment is costly both for the punishing and the punished parties. Adapting Fehr and Gächter (2002), we choose a punishment technology in which each punishment point assigned costs the punished player two Guilders and the punishing player one Guilder.

The material payoff function from the whole experiment for a given subject i is given by equation (3).

$$\pi_{i} = \pi_{i}^{1} - \sum_{j=1}^{n} p_{ij} - 2 \cdot \sum_{j=1}^{n} p_{ji}$$
(3)

where π_i^1 denotes group member *i*'s payoff from the contribution stage, and p_{ij} the punishment points group member *i* assigns to group member *j*. Conditional on each subject *i* being motivated to maximise equation (3), the unique subgame perfect equilibrium requires that subjects free ride completely in the first stage and refrain completely from punishing in the second stage.

We asked subjects to state their beliefs about contributions of the other group members after they had made their own contribution decision (but before the punishment stage). We also elicited beliefs about how much punishment a subject expected to receive. We elicited these beliefs after subjects had had the opportunity to punish. Elicitation of beliefs was non-incentivised in order to exclude potential income effects in the punishment treatments. We also wanted to avoid punishment being motivated by disappointment about low payoffs resulting from inaccurate beliefs.

In each treatment, subjects were asked at the end of the game to indicate the intensity of emotions they felt about the actual contribution behaviour of each member of their group. The procedure we used to elicit self-reports on perceived emotions is due to Bosman and van Winden (2002). In particular, subjects were given a list of thirteen emotions, and were then asked to indicate the intensity with which they felt each emotion when they saw the contribution of each other group member. The intensity for each emotion was recorded on a 7-point scale (1 = "not at all", ..., 7 = "very much"). Appendix I provides a screenshot of the interface we used for eliciting self-reports on emotions.

Our design is completed by two treatments without punishment opportunities. In these treatments, the second stage consisted only of the elicitation of beliefs and emotions. We refer to the no-punishment treatments as Give-N (Give frame for contributions, no punishment opportunities) and Take-N (Take frame for contributions, no punishment opportunities). The reason for including the no-punishment treatments is to check for two possibilities: (i) if it turns out that there are differences in emotions across frames, we want to be able to check whether emotions are responses to contributions themselves or *ex post* rationalisations of punishment behaviour; and (ii) if it turns out that there is no difference in emotions across frames, we want to be able to check whether this is because emotional response to contributions is the same or because the act of punishment expunges (or arouses) emotional response.

Each subject was assigned at random to a group of three members and played a one-shot voluntary contributions game under one of the treatments just described. (Sessions were allocated to treatments at random.) The advantage of a one-shot game is that it eliminates confounding effects that might come from repeated interaction, allowing us to focus on pure framing effects. In contrast, in a repeated game, there is always the possibility that subjects think that other subjects' future behaviour may be influenced by their own current behaviour. This could confound the investigation of framing effects in two ways. First, sufficient and strong repeated game effects could swamp framing effects that would otherwise be present. Alternatively, if subjects' views of the dependence of other subjects' behaviour on their own are framesensitive, repetition could create a framing effect that would not otherwise be present.

In total, 42 subjects took part in the Give-N treatment; 45 in the Take-N treatment; 42 in the Give-P treatment; and 39 in the Take-P treatment. All subjects were recruited at the University of Nottingham. The vast majority were undergraduate students from different academic fields including, but not confined to, economics. The experiment was conducted in the Centre for Decision Research and Experimental Economics (CeDEx) lab. All treatments were computerised and programmed with the software z-Tree (Fischbacher, 2007). The no-punishment treatments lasted about 50 minutes and the punishment treatments about 70 minutes. At the end of each session, guilders were converted to UK pounds at the pre-announced exchange rate of £0.40 per guilder and subjects were paid in cash. On average, subjects earned about £9 in the no-punishment and £12.30 in the punishment treatments. Before subjects played the game, they received the instructions reproduced in Appendix II. As we wanted to ensure that subjects understood the decision situation and the mechanics of payoff calculations, all participants answered several computerised test questions, concerning what the payoffs would be for various hypothetical configurations of behaviour. The experiment did not proceed until every subject had answered these questions correctly.

2.3 Hypotheses

We now state our hypotheses concerning the effect of framing on contributions, and, more importantly, on punishment behaviour, and reported emotions.

For contributions, the null hypothesis is that framing has no effect on their level, at least in no-punishment treatments. An alternative hypothesis for these treatments, suggested by previous evidence from repeated public goods games and our findings on moral judgments described in Section 1, is that contributions will be higher in the Give frame than in the Take frame. Previous evidence suggests that punishment opportunities will tend to raise contributions (Fehr and Gächter, 2000; 2002), but whether this should be expected to induce a difference in contributions between Give and Take frames depends on the first of the two questions to which we now turn.

These questions, which are our main concern, are: whether the propensity to punish differs between frames; and whether emotional responses to contributions do likewise. We investigate them using the "punishment function" and the "emotions' function", respectively, as tools.

The "punishment function" gives the punishment points assigned by the punisher, as a function of the recipient's deviation from the punisher's contribution. Consistent with previous evidence (e.g., Fehr and Gächter, 2000), our expectation is that the punishment function will be downward sloping for the negative part of the deviation (horizontal) axis, implying that a subject punishes his co-player more, the more he negatively deviates from the punisher's contribution.

Having defined the punishment function, we can now state our derived hypotheses in relation to framing. The null hypothesis predicts that the punishment function does not depend on framing, *ceteris paribus*. This hypothesis is implied by any consequentialist theory of behaviour, even by those which allow subjects to contribute and punish to some extent. For instance, although the inequity aversion model of Fehr and Schmidt (1999) or the ERC model of Bolton and Ockenfels (2000) can explain contribution and punishment, these theories cannot account for any framing effect.⁴ Even if framing leads to different contribution levels, punishment may not be different *for a given deviation* by the recipient from the punisher's contribution. This prediction parallels the finding of Fehr and Gächter (2000) that, even if subjects contribute more under a fixed matching protocol than under a random matching one, these treatment manipulations do not affect the level of punishment for given deviations from the average group contribution. An effect of different treatments on contribution levels need not imply any corresponding effect on punishment, for a given deviation from the punisher's contribution.

In principle, the punishment function could be tilted either upwards or downwards by the Take frame relative to the Give frame. However, the evidence of Cubitt, Drouvelis, Gächter and Kabalin (2008), in which subjects morally rate free riding, suggests that failing to contribute is perceived as morally worse than withdrawing from the public good. Motivated by these findings, the alternative hypothesis states that the "punishment function" would be steeper in the Give, compared to the Take frame. We expect there to be no punishment if the deviation from the punisher's

⁴ Consequentialist theories suppose that subjects' actions are determined only by the final consequences of the actions. Almost all economic theories are consequentialist in this sense. However, psychological game theory (Geanakoplos, Pearce and Stacchetti, 1989) includes beliefs in the payoff function and can therefore explain framing effects (see Dufwenberg, Gächter and Hennig-Schmidt, 2006).

contribution is zero. Similar hypotheses can be derived using j's deviation from the punisher's beliefs about j's contributions in place of j's deviation from i's contribution.

The second tool we use to analyse subjects' perception of free riding is the "emotions' function", which gives aggregate emotions in a given category (i.e. positive or negative) as a function of deviations from individual's own contribution. We expect the "emotions' function" to be negatively sloped in the negative deviation interval for the negative emotions, and positively sloped in the negative deviation interval for the positive emotions.⁵ As with the "punishment function", we can also state hypotheses for the slope and level of the "emotions' function". Since punishment and emotions are likely to be closely related, we expect that any differences observed in the "punishment function" will be reflected in the "emotions' function" as well (and in the same direction).

3. Results

3.1 Contributions

Table 1 shows the average absolute level of contribution and the average belief about average contributions across all subjects in each treatment. In absolute levels, the average contribution is largest in the Give-P treatment (i.e. 7.21 tokens) and smallest in the Take-N treatment (i.e. 4.47 tokens), as previous literature on framed public goods experiments would suggest. A non-parametric Kruskal-Wallis test suggests significant differences among treatments ($\chi^2(3) = 10.089$, p = 0.0178). Using a Wilcoxon rank-sum test shows that the only significant differences at the 5% level are between the Give-P and Take-N treatments (p = 0.002) and between the Take-N and Take-P treatments (p = 0.025).

Examining subjects' beliefs about contributions, we find that a Kruskal-Wallis test suggests weakly significant differences between treatments ($\chi^2(3) = 6.626$, p = 0.085). Using a Wilcoxon rank-sum test shows that the only significant differences at the 5% level are between the Give-P and Take-N treatments (p = 0.032) and between the Give-P and Take-P treatments (p = 0.023).

⁵ We restrict our attention below to the negative deviation interval, since we are mostly interested in how subjects treat lower contributions by their counterparts.

contributions			
	Average contribution	Beliefs about average	
	levels	contributions	
Give-N	5.88	7.08	
	(6.24)	(5.26)	
Take-N	4.47	6.23	
	(6.04)	(5.93)	
Give-P	7.21	7.70	
	(5.03)	(4.32)	
Take-P	6.41	5.31	
	(5.12)	(3.78)	

 Table 1. Average absolute levels of contribution and beliefs about contributions

Note: Standard deviations in parentheses.

However, our main concern is with punishment behaviour and emotions, as these are our vehicles to answer our initial research question. In the next two sections, we explore whether subjects treat free-riding differently by punishing differently or/and displaying different feelings across framing manipulations, using the "punishment function" and the "emotions' function" as our tools.

3.2 The punishment function

We start our analysis by examining subjects' punishment behaviour for each framing context. Figure 1 below shows the punishment points assigned by the punisher to another player j as a function of j's deviation from the punisher's contribution.⁶ The punishment function is given by the solid line, which indicates the fitted line of the locally weighted regression of punishment assigned on the deviation from the punisher's contribution.

<u>RESULT 1:</u> The punishment function does not depend on framing, ceteris paribus.

Support. In Figure 1, each dot represents a single observation. The punishment functions appear quite similar across frames, having the anticipated negative slope.

⁶ We refer to the punisher as player i, the recipient of punishment as player j, and the third group member as player k.



Figure 1. Punishment as a function of deviation from punisher's contribution

However, the graphs in Figure 1 are simply visual representations of the punishment function and do not control for any factors other than j's deviation from *i*'s contribution. To test econometrically whether the slope of the punishment function differs across frames, we estimated a Tobit regression model. In this regression, the dependent variable is the "punishment assigned by player *i* to player j" and the independent variables comprise "Player k's contribution deviation", "Player j's absolute negative (contribution) deviation", "Player j's positive (contribution) deviation", and the dummy variable "Take". We also included two interaction terms, which indicate whether the slope of the punishment function differs with respect to negative and positive deviations across frames. Note that all deviations are calculated with respect to the punisher's contribution. We include "absolute negative deviation" and "positive deviation" as separate regressors, since Figure 1 suggests that these two different sorts of deviation elicit different punishment responses. The variable "absolute negative deviation" is the absolute value of the actual deviation of subject *j*'s contribution from the punisher's contribution, when subject *j*'s contribution is below the punisher's contribution; and zero otherwise. The variable "positive deviation" is constructed in an analogous way. The variable "Player k 's contribution deviation" is the actual deviation of player k 's contribution from the punisher's contribution. The reason for including such a variable is that player i 's attitude to player j may differ according to the behaviour of player k. The dummy variable "Take" equals 0 for the Give frame, and 1 for the Take frame. The regression results are given in Table 2.

Table 2. The pumsiment function – Kegression results		
	Dependent variable: Punishment assigned by player i to player j	
Player <i>j</i> 's absolute negative deviation	0.790**	
from punisher's contribution	(0.331)	
Player <i>j</i> 's positive deviation from	-0.102	
punisher's contribution	(0.337)	
Player k 's contribution deviation from	0.083	
punisher's contribution	(0.096)	
Take	-1.699	
	(2.739)	
Take \times Player <i>j</i> 's absolute negative	-0.007	
deviation from punisher's contribution	(0.497)	
Take \times Player <i>j</i> 's positive deviation	0.378	
from punisher's contribution	(0.396)	
Constant	-9.025***	
	(3.425)	
Observations	162	

 Table 2. The punishment function – Regression results

Notes: Tobit estimates. Standard errors are presented in parentheses (clustered on groups). Results are corrected for heteroskedasticity. ** denotes significance at the 5-percent level, and *** at the 1-percent level.

The results from Table 2 indicate that the vertical intercept of the punishment function does not differ across treatments, since the coefficient of the dummy variable "Take" is not statistically significant. The coefficients of the interaction terms are also insignificant, implying that the slope of the punishment function with regard both to negative and positive deviations is the same, irrespective of framing. These findings fail to reject the null hypothesis that subjects punish given (negative and positive) deviations from their own contribution equally in different frames. According to our findings, the only variable that has a significant effect on the assignment of punishment is the absolute negative deviation from the punisher's contributions: a subject punishes a co-player more, the less the co-player contributes relative to the punisher.

Performing the same sort of analysis with respect to deviations from punisher's beliefs about contributions, we again find no evidence that the "punishment function" differs between frames. Analogously with the case of deviations from the punisher's actual contribution, the only significant variable that explains punishment is player j's absolute negative deviation from punisher's beliefs about contributions.

Thus far, we have found that subjects do not treat deviations from own behaviour differently across frames, in terms of the punishment they mete out. However, this observation could be attributed either to subjects perceiving deviations differently across frames, but not being prepared to act differently on the basis of this perception; or to subjects not perceiving deviations differently across frames at all. In the next section, we use the elicited emotions to disentangle these two possibilities.

3.3 The emotions' function

In this section, we use as tools the mean positive and mean negative emotions' functions, which are aggregate measures of positive and negative emotions, respectively, expressed as functions of other variables. Put simply, we plot mean positive and mean negative emotions as a function of deviations from player i's contribution. Positive emotions comprise warmth, happiness and joy; whereas, negative emotions comprise anger, fear, envy, sadness, shame, irritation, contempt, guilt and jealousy. To control whether the presence of punishment opportunities has influenced emotions, we also examine emotions for the non-punishment treatments. Our findings are recorded in Result 2.

<u>RESULT 2.</u> The (mean positive and mean negative) emotions' function does not depend on framing, ceteris paribus. The elicited emotions are not affected by the presence of punishment, ceteris paribus.

Support. Figures 2 and 3 provide a graphical illustration of the positive and negative emotions' function for each of the four treatments, respectively. In all figures, each dot represents a single observation and the solid line indicates the fitted

line of the locally weighted regression of emotions expressed on the deviation from the punisher's contribution.



Figure 2. Mean positive emotions for each treatment

Figure 3. Mean negative emotions for each treatment



To test whether the emotions' function differs across Give and Take treatments, we estimated an ordered probit model for the average positive and negative emotions. In this model, all data are pooled. We also include a dummy variable called "No-Punishment", which takes on the value '1' for the no-punishment treatments and '0' for the punishment treatments, to test for any difference in the elicited emotions between punishment and no-punishment treatments, for the reasons explained in Section 2.2. The inclusion of other independent variables follows similar reasoning as for the case of the punishment function. Table 3 provides econometric evidence for Result 2.

Table 3. The positive and negative emotions' function – Regression results			
	Dependent variable:	Dependent variable:	
	Mean positive emotions	Mean negative emotions	
Player j's absolute	-0.080**	0.096***	
negative deviation from	(0.032)	(0.021)	
player <i>i</i> 's contribution			
Player j's positive	0.118***	0.018	
deviation from player <i>i</i> 's	(0.024)	(0.019)	
contribution			
Player k's contribution	-0.003	0.012	
deviation from player <i>i</i> 's	(0.009)	(0.010)	
contribution			
Take	0.013	-0.168	
	(0.201)	(0.214)	
Take \times Player <i>j</i> 's absolute	0.000	0.035	
negative deviation from	(0.042)	(0.042)	
player <i>i</i> 's contribution			
Take \times Player <i>j</i> 's positive	-0.026	-0.020	
deviation from player <i>i</i> 's	(0.028)	(0.025)	
contribution			
No-Punishment	0.016	-0.001	
	(0.141)	(0.138)	
Observations	336	336	

Notes: Ordered probit estimates. Standard errors are presented in parentheses (clustered on groups). Results are corrected for heteroskedasticity. ** denotes significance at the 5-percent level, and *** at the 1-percent level. All data are pooled.

It is clear from Table 3 that the coefficients are not statistically different across frames, implying that the aggregate "emotions' function" is not frame-sensitive, for either positive or negative emotions. This strengthens our conclusion that subjects do not consider negative deviations from their own contribution differently across frames.

The coefficient on the variable No Punishment is not significant in either equation reported in Table 3. This finding provides no support for the idea that punishment opportunities expunge emotional responses. To investigate further whether elicited emotions vary with the presence or absence of a punishment option, we also estimated four ordered probit models: two relating to the no-punishment case and two to the punishment case (in each case, taking mean positive and mean negative emotions, separately). The rationale and interpretation of the independent variables in these models is parallel to that described above. Our regression results indicate that the coefficients for the dummy variable 'Take' and its interaction terms are not statistically significant. Elicited emotions do not differ between the Give and Take frames for either the no-punishment or the punishment treatment, taken separately, corroborating our earlier conclusion that the act of punishment does not appear to expunge emotional responses.

Just as with the punishment function, replacing deviations from contributions with deviations from beliefs about contributions in the emotions functions makes little qualitative difference to the results. In particular, econometric analysis finds no framing differences for either positive or negative emotions, whereas the only significant variable is the absolute negative deviation from player i's beliefs about contributions. These results are strikingly similar to those reported in Tables 3 in terms of statistical significance.

4. Conclusion

Motivated by previous findings about altruistic behaviour in dictator games, this paper reports an investigation of the frame-sensitivity of two indicators of negative reciprocity, an equally-important, if darker, source of social preferences. Specifically, we study punishment and emotional responses in a one-shot public goods game. Our findings suggest that neither is sensitive to the Give versus Take framing manipulation we consider, controlling for the level of contributions. We conclude that the main determinant of punishment is the difference between the contributions of

the punisher and the punished; and the same difference is also the main source of emotional responses. In contrast to the frame-sensitivity of altruism in dictator games, our framing has little effect on these indicators of negative reciprocity.

Appendix I – Screenshot for eliciting emotions

[Note: The screenshot for eliciting self-reports on emotions is presented below. The order of emotions was the same in all four treatments.]

eriod	1 out of 1		
	You can now see the number of tokens eac which you feel each emotion when you see	h member of your group has contributed. Please the contribution of the other members. Tokens contributed by you: tokens Your Income: Guilders	e indicate for each emotion the intensity with
		Group Member 1: Tokens contributed: tokens Income from stage 1: Guilders	Group Member 2: Tokens contributed: tokens Income from stage 1: Guilders
	Warmth	not at all CCCCCC very much	not at all CCCCCC very much
	Anger	not at all CCCCCC very much	not at all CCCCCC very much
	Fear	not at all CCCCCC very much	not at all CCCCCC very much
	Envy	not at all CCCCCC very much	not at all CCCCCC very much
	Sadness	notatall CCCCCCC very much	not at all CCCCCC very much
	Happiness	not at all CCCCCC very much	not at all CCCCCC very much
	Shame	not at all CCCCCC very much	not at all CCCCCC very much
-	Irritation	not at all CCCCCC very much	not at all CCCCCC very much
	Contempt	not at all CCCCCC very much	not at all CCCCCC very much
	Guilt	not at all CCCCCC very much	not at all CCCCCC very much
	Joy	not at all CCCCCC corry much	not at all CCCCCC very much
	Jealousy	not at all CCCCCC very much	not at all CCCCCC very much
	Surprise	not at all CCCCCC very much	not at all CCCCCC very much
		1	ОК

Appendix II – Instructions

[Note: These are the written instructions for the contribution stage (Stage 1) and the punishment stage (Stage 2), as presented to subjects facing the Give-P treatment. Amendments to the Stage 1 instructions for the Take and No Punishment treatments are given in square brackets and curly brackets, respectively. Stage 2 instructions were the same across Give-P and Take-P treatments, but were not applicable to the No Punishment treatments. Screenshots are as for the Give treatments; those for the Take treatments were amended accordingly, for example replacing the word 'contribute' with 'withdraw'.]

Instructions

You are now taking part in an economic experiment financed by the University of Nottingham. You can earn a considerable amount of money depending on the decisions made by you and other participants. It is therefore very important that you read these instructions with care.

These instructions are solely for your private use. It is prohibited to communicate with other participants during the experiment. If you have any questions, please ask me. If you violate this rule, you will be dismissed from the experiment and forfeit all payments.

During the experiment we will not speak in terms of Pounds, but in Guilders. During the experiment your entire earnings will be calculated in Guilders. At the end of the experiment the total amount of Guilders you have earned will be converted to Pounds at the following rate:

1 Guilder = 0.40 Pounds

At the end of the experiment your entire earnings from the experiment will be paid to you **in cash**.

During the experiment, you will be asked to fill in a few questionnaires. The answers you provide in these questionnaires are completely anonymous. They will not be revealed to anyone either during the experiment or after it. Furthermore, your answers to these questionnaires will not affect your earnings during the experiment.

At the beginning of the experiment, all participants will be randomly divided into groups of three. Apart from you, there will be two more members in your group. You will not learn who the other people in your group are at any point.

The experiment consists of two stages {one stage}. In the following pages we describe the experiment in detail. At the end of this introductory information we ask you to do several computerised control exercises, which are designed to check that you have understood the decision situation.

Detailed Information on the Experiment

Stage 1

Each participant receives an endowment of **20 tokens** [There are **60 tokens** in a project for your group]. At stage 1, you have to decide how many of these 20 tokens you contribute to a group project and how many you keep for yourself. [At stage 1, you have to decide how many of these 60 tokens you withdraw from the project for yourself and how many of them you leave in the project]. The two other members of your group have to make the same decision. They can also either contribute tokens to a project or keep tokens for themselves. [They can also either withdraw tokens from the project for themselves or leave tokens in the project]. You and the other members of the group can each choose any amount between 0 and 20 tokens to contribute [withdraw].

Every token that you do not contribute to [withdraw from] the project automatically belongs to you and earns you one Guilder.

For the tokens contributed to [that are not withdrawn from] the project the following happens: **the project's value will be multiplied by 1.5 and this amount will be divided equally among all three members of the group.** For example, if 1 token is contributed to [is not withdrawn from] the project, the project's value increases to 1.5 Guilders. This amount is divided equally among all three members of the group. Thus every group member receives 0.5 Guilders.

Your income from the project rises by 0.5 Guilders if you contribute [withdraw] one token more to [less from] the project. At the same time, the income of the other two members of the group also rises by 0.5 tokens, because they receive the same income from the project as you do. Therefore, if you contribute [withdraw] one token more to [less from] the project, the income from the project received by the whole group together increases by 1.5 Guilders. It is also true that your income rises by 0.5 Guilders if another group member contributes [withdraws] one token more to [less from] the project.

After all three members of the group have made their decisions about the amounts of tokens they contribute to [withdraw from] the project the total income achieved by each participant is determined.

How is your income calculated from your decision?

The income of every member of the group is calculated in the same way. As you can see, your **income** consists of two parts:

(1) The tokens which you have kept [withdrawn] for yourself ('income from tokens kept') ['income from tokens withdrawn'] whereby **1 token = 1 Guilder**.

(2) The 'income from the project' calculated as follows: Your income from the project = $0.5 \text{ times sum of all tokens contributed to the project by members of your group [0.5 times (60 - sum of all tokens withdrawn from the project by members of your group)].$

Your total income in Guilders at stage 1 of the experiment is therefore:

(20 – tokens contributed to the project by you) + 0.5*(sum of all tokens contributed to the project by members of your group) [(Tokens withdrawn from the project by you) + 0.5*(60 – sum of all tokens withdrawn from the project by members of your group)]

If you do not contribute anything to [withdraw all 20 tokens from] the project the income from tokens kept [withdrawn] is 20. If you contribute [leave] for instance 7 tokens to the project your income from tokens kept [withdrawn] is 13. At the same time, the total sum of tokens contributed to [left in] the project increases and so does your 'income from the project'.

In order to explain the income calculation we give some examples. Please read them carefully:

Example 1:

If each of the three members of the group contributes 0 tokens to [withdraws 20 tokens from] the project, all three will receive an 'income from tokens kept [withdrawn]' of 20. Nobody receives anything from the project, because no one contributed [left] anything. Therefore the total income of every member of the group is 20 tokens.

Calculation of the income from stage 1 for every participant: (20-0) + 0.5 * (0) = 20[Calculation of the income from stage 1 for every participant: (20) + 0.5 * (60-60) = 20]

Example 2:

If each of the three members of the group contributes 20 [withdraws 0] tokens, there will be a total of 60 tokens contributed to [left in] the project. The 'income from tokens kept [withdrawn]' is 0 for everyone, but each member receives an income from the project of 0.5 * 60 = 30 tokens.

Calculation of the income from stage 1 for every participant: (20-20) + 0.5 * (60) = 30

[Calculation of the income from stage 1 for every participant: (0) + 0.5 * (60-0) = 30]

Example 3:

If you contribute 20 [withdraw 0] tokens, the second member 10 tokens and the third 0 [20] tokens, the following incomes are calculated.

- Because you and the second member of the group have together contributed [withdrawn] 30 tokens, everyone will receive 0.5 * 30 = 15 Guilders from the project.
- You contributed all your 20 tokens to [withdrew 0 tokens from] the project. You will therefore receive 15 Guilders in total from the project.
- The second member of the group also receives 15 Guilders from the project. In addition, he receives 10 Guilders as the 'income from tokens kept [withdrawn]', because he contributed [withdrew] 10 tokens to [from] the project. Thus, he receives 15 + 10 = 25 Guilders altogether.

- The third member of the group, who did not contribute anything [withdrew all tokens], also receives the 15 Guilders from the project and additionally the 20 Guilders from the 'income from tokens kept [withdrawn]', which means 20 + 15 = 35.

Calculation of your income from stage 1: (20-20) + 0.5 * (30) = 15Calculation of the income from stage 1 for the 2^{nd} group member: (20-10) + 0.5 * (30) = 25

Calculation of the income from stage 1 for the 3^{rd} group member: (20-0) + 0.5 * (30) = 35

[Calculation of your income from stage 1: (0) + 0.5 * (60-30) = 15

Calculation of the income from stage 1 for the 2^{nd} group member: (10) + 0.5 * (60-30) = 25

Calculation of the income from stage 1 for the 3^{rd} group member: (20) + 0.5 * (60-30) = 35]

Example 4:

The two other members of your group contribute 20 [withdraw 0] tokens each to [from] the project. You do not contribute anything [withdraw all tokens]. In this case the income will be calculated as follows:

Calculation of your income from stage 1: (20-0) + 0.5 * (40) = 40

Calculation of the income from stage 1 for the 2^{nd} group member: (20-20) + 0.5 * (40) = 20

Calculation of the income from stage 1 for the 3^{rd} group member: (20-20) + 0.5 * (40) = 20

[Calculation of your income from stage 1: (20) + 0.5 * (60-20) = 40

Calculation of the income from stage 1 for the 2^{nd} group member: (0) + 0.5 * (60-20) = 20

Calculation of the income from stage 1 for the 3^{rd} group member: (0) + 0.5 * (60-20) = 20]

Devied	
reilou	
1 out of 1	
Your endowment 20	
How many tokens do you want to contribute?	
	ОК
HELP	
Please fill in the amount of tokens (between 0 and 20) you want to contribute to the project	
When you are ready, please press the "OK"-button.	

When making your decision, the following input-screen will appear:

As mentioned above, your **endowment in the experiment is 20 tokens** [there are **60 tokens** in a project for your group]. You have to decide how many tokens you contribute to the project by typing a number between 0 and 20 in the input field [You have to decide how many of these 60 tokens you withdraw from the project by typing a number between 0 and 20 in the input field]. This field can be reached by clicking it with the mouse. By deciding how many tokens to contribute to [withdraw from] the project, you automatically decide how many tokens you keep for yourself [you leave in the project]. After entering the amount of tokens you contribute [withdraw] you must press the O.K. button using the mouse. Once you have done this, your decision can no longer be revised.

After that, you will be informed about the amount of tokens contributed to [withdrawn from] the project by you, the sum of tokens contributed to [withdrawn from] the project and your total income in this stage.

1 out of 1	
Tokens contributed by you	
Sum of tokens contributed	
Your income from takens kent	
Your income from the project	
Your total income in this stage	
	_
continue	
HELP	_
You can inspect the results of this stane	
Tou can inspect the results of this stage.	
When you are ready, please press the "Continue"-button.	

Stage 1 is now over and stage 2 commences.

Stage 2

At this stage, you will see how many tokens each of the other two group members has contributed to [withdrawn from] the project and his or her corresponding income from stage 1. Moreover, you can either **decrease** or **leave unchanged** the income of each other group member by assigning **deduction points** to them. The other group members can also decrease **your** income, by assigning deduction points to you, if they wish to do so.

Period			
1 outof 1			
Tokens contributed			
income irom stage i			
Your decision in stage 2	-	Assign no points: 0 Assign deduction points: negative number	Calculation
	Costs of deduction points distribute	d by you	
			ок
HELP			
Please insert your decision. Note the sign of your distribution	n of points. Then press the "Calculation"-	button.	
When you are ready, please press the "OK"-button.			

You will see the following input screen at stage 2:

You must now decide how many deduction points to assign to each of the other two group members. In the first column you can see your contribution [withdrawal] and your income from stage 1. In the other two columns, you can see the same information for each of the two other members of the group.

If you do not wish to change the income of a specific group member then enter 0 in the large box for that group member. If you do wish to reduce a group member's income, enter instead the number of deduction points that you wish to assign to them, preceded by minus sign (without spaces between them). For example, to assign 2 deduction points to a group member, type -2 in the relevant box. You can move from one input field to the other by pressing the tab -key $(\rightarrow |)$ or by using the mouse. You must enter a response in each large box.

You can assign between 0 and 5 deduction points to each other group member.

For each deduction point that you assign, there is a cost to you of one Guilder. Thus, the total cost to you in Guilders of assigning deduction points to other group members is given by the total number of deduction points that you assign. You can check the total cost on the computer, by pressing the 'Calculation' button after you have assigned deduction points. Until you press the OK-button, you can still change your decision. To recalculate the costs after making a change, simply press the cost calculation button again.

The effects of assigning deduction points to other group members are as follows: If you give 0 points to a particular group member, you will not have any effect on his or her income. However, for each deduction point that you assign to a particular group member, you will decrease their income by 2 Guilders (unless their income is already exhausted). For example, if you give a group member 2 deduction points (i.e., enter -2), you will decrease their income by 4 Guilders. And so on.

Your own income will be reduced by 2 Guilders for each deduction point that is assigned to you by the other two group members, except that, if all of your income from the first stage is exhausted as a result of deduction points, your income cannot be reduced any further by other group members. Therefore, your total income from the two stages is calculated as follows:

Total income	(in	Guilders)	after stage 2	
---------------------	-----	-------------------	---------------	--

- = income from stage 1
- -2*(sum of deduction points assigned to you) (2) - costs of deduction points assigned by you

(1)

- if (1) + (2) is greater than or equal to 0;
- =**0** costs of deduction points assigned by you

if (1) + (2) is less than 0

Please note that your income in Guilders after stage 2 can be negative, if the cost of deduction points assigned by you exceeds your income from stage 1 less any reduction in your income caused by deduction points assigned to you by other group members.

However, at the end of the experiment and in addition to the calculation just given, you and the other members of your group will each receive a lump sum payment of **10 Guilders.** This payment is enough to cover any losses that you could incur.

Do you have any questions?

References

- Andreoni, J., (1995), 'Warm-Glow versus Cold-Prickle: The Effects of Positive and Negative Framing on Cooperation in Experiments', *Quarterly Journal of Economics*, vol. 110, pp. 1-21
- Andreoni, J., Harbaugh, W., and Vesterlund, L., (2003), 'The Carrot or the Stick: Rewards, Punishments, and Cooperation', *American Economic Review*, vol. 93, pp. 893-902
- Bardsley, N., (2007, forthcoming), 'Dictator game giving: altruism or artefact?', Experimental Economics
- Bolton, G., and Ockenfels, A, (2000), 'ERC: A theory of Equity, Reciprocity, and Competition', *American Economic Review*, vol. 90, pp. 166-193
- Bosman, R., and van Winden, F., (2002), 'Emotional Hazard in a Power-to-Take Experiment', *Economic Journal*, vol. 112, pp. 147-169
- Brewer, M. B. and Kramer, R. M., (1986), 'Choice behavior in social dilemmas: Effects of social identity, group size, and decision framing', *Journal of Personality and Social Psychology*, vol. 50, pp. 543-549
- Camerer, C., (2003), Behavioural game theory: Experiments on strategic interaction, Princeton: Princeton University Press
- Camerer, C., and Fehr, E., (2004), 'Measuring Social Norms and Social Preferences using Experimental Games: A Guide for Social Scientists' in *Foundations of Human Sociality – Economic Experiments and Ethnographic Evidence from Fifteen Small-Scale Societies*, edited by Joseph Henrich, Robert Boyd, Samuel Bowles, Colin Camerer, Ernst Fehr and Herbert Gintis, pp. 55-95
- Cookson, R., (2000), 'Framing effects in Public Goods Experiments', *Experimental Economics*, vol. 3, pp. 55-79
- Cubitt, R. P., Drouvelis, M., Gächter, S., and Kabalin, R., (2008), 'Moral Judgments and Framing in Social Dilemma Games: An Internet Experiment', *CeDEx Discussion Paper, School of Economics, University of Nottingham.* In Preparation
- Dana, J., Weber, R., and Kuang, J. X., (2005), 'Exploiting moral wiggle room: Experiments demonstrating an illusory preference for fairness', Working Paper, Center for Behavioral Decision Research, Carnegie-Mellon University.

- **Dufwenberg, M., Gächter, S., and Hennig-Schmidt, H.,** (2006), 'The Framing of Games and the Psychology of Strategic Choice', *CeDEx Discussion Paper No.* 2006-20
- Fehr, E., and Gächter, S., (2000), 'Cooperation and Punishment in Public Goods Experiments', *American Economic Review*, vol. 90, pp. 980-994
- Fehr, E., and Gächter, S., (2002), 'Altruistic Punishment in Humans', *Nature*, vol. 415, pp. 137-140
- Fehr, E., and Schmidt, K., (1999), 'A Theory of Fairness, Competition and Cooperation, *Quarterly Journal of Economics*, vol. 114, pp. 817-868
- Fischbacher, U., (2007) 'z-Tree: Zurich Toolbox for Ready-made Economic Experiments' *Experimental Economics*, vol. 10, pp.171-178
- Geanakoplos, J., Pearce, D., and Stacchetti, E., (1989), 'Psychological Games and Sequential Rationality', *Games and Economic Behavior*, vol. 1, pp. 60-79
- Gintis, H., Bowles, S., Boyd, R., and Fehr, E., (2005), Moral Sentiments and Material Interests – The Foundation of Cooperation in Economic Life, MIT Press, Cambridge Massachusetts
- Haley, K. J., and Fessler, D. M. T., (2005), 'Nobody's watching? Subtle cues affect generosity in an anonymous economic game', *Evolution and Human Behaviour*, vol. 26, pp. 245-256
- Levin, I., Schneider, S., and Gaeth G., (1998), 'All Frames are Not Created Equal: A Typology and Critical Analysis of Framing Effects', Organizational Behavior and Human Decision Processes, vol. 76, pp. 149-188
- List, J. A., (2007), 'On the interpretation of giving in dictator games', *Journal of Political Economy*, vol. 115, pp.482-93
- McCusker, C., and Carnevale, P. J., (1995), 'Framing in Resource Dilemmas: Loss Aversion and the Moderating Effects on Sanctions', *Organizational Behavior and Human Decision Processes*, vol. 61, pp. 190-201
- McDaniel, W. C., and Sistrunk, F., (1991), 'Management dilemmas and decisions: Impact of framing and anticipated responses', *Journal of Conflict Resolution*, vol. 35, pp. 21-42
- Masclet, D., Noussair, C., Tucker, S., and Villeval, M-C., (2003), 'Monetary and Nonmonetary Punishment in the Voluntary Contributions Mechanism', *American Economic Review*, vol. 93, pp. 366-380

- Park, E. S., (2000), 'Warm-Glow versus Cold Prickle: a further experimental study of framing effects on free-riding', *Journal of Economic Behavior and Organization*, vol. 43, pp. 405-421
- Sell, J., and Son, Y., (1997), 'Comparing public goods and Common Pool Resources: Three Experiments', *Social Psychology Quarterly*, vol. 60, pp. 118-137
- Sonnemans, J., Schram, A., and Offerman, T., (1998), 'Public Good Provision and Public Bad Prevention: The Effect of Framing', *Journal of Economic Behaviour and Organization*, vol. 34, pp. 143-161
- Van Dijk, E., and Wilke, H., (2000), 'Decision-Induced Focusing in Social Dilemmas: Give-Some, Keep-Some, Take-Some, and Leave-Some Dilemmas', *Journal of Personality and Social Psychology*, vol. 78, pp. 92-104
- Willinger, M., and Ziegelmeyer, A., (1999), 'Framing and cooperation in public good games: an experiment with an interior solution', *Economics Letters*, vol. 65 (3), pp. 323-328