Contribution and Payoff Feedback in Public Good Experiments with Punishment: Evidence from the U.K.

Nikiforakis N.

Department of Economics, The University of Melbourne, Victoria 3010, Australia
Email: n.nikiforakis@unimelb.edu.au

Keywords: feedback, cooperation, punishment, public goods, experiment.

EXTENDED ABSTRACT

Recently, experimental economists have provided evidence that cooperation in a social dilemma can be sustained if individuals are allowed to punish free riders. Despite the attention these studies have attracted and the important implications the results have for economic theory, little is known about their robustness. The present study is motivated by findings in oligopoly experiments showing that individuals behave more competitively when they receive information about the earnings of their peers (e.g. Offerman et al. 2002). This is surprising as according to theory information about earnings should not affect behavior when individuals are given all information they need to calculate earnings. This experiment investigates whether a similar phenomenon occurs in public good games with punishment opportunities and, in particular, whether this effect is stronger than the disciplinary effect of punishment. I examine the effect of two different types of feedback on punishment behavior and cooperation levels in a public-good experiment. In one treatment, participants are informed of the individual contributions of the other group members (contribution feedback) as in all public good experiments so far (e.g. Fehr and Gächter , 2000). In a second treatment, instead of receiving information regarding contributions, participants are informed about the payoffs of the other people in their group (payoff feedback). Based on this feedback individuals make their punishment decisions. The information in the two treatments is equivalent as individuals can calculate payoffs based on contribution feedback, and contributions based on payoff feedback. Nevertheless, there is a pronounced difference in behavior across treatments. While cooperation is sustained at intermediate levels under contribution feedback, it quickly unravels under payoff feedback. Part of this decay can be explained by the (weakly) less frequent occurrence of punishments under payoff feedback. The result is that efficiency is much lower under payoff feedback. The evidence raise questions about the robustness of decentralized enforcement of cooperation through punishments.

1. Introduction

Despite the rapidly increasing number of studies on decentralized punishment in social dilemmas and their role in motivating new theories (e.g. Fehr and Schmidt, 1999) nothing is known about whether different types of feedback affect punishing behavior and cooperation in general. This is not a trivial question. Recent experimental studies on oligopolistic markets have shown that participants receiving information about individual profits behave more competitively and tend to be less cooperative than subjects who instead receive information about the aggregate group production (Huck et al., 1999; 2000) or the individual levels of production (e.g. Offerman et al., 2002).

In this paper I examine the effect of the type of feedback on the behavior of individuals in a public-good game with punishment opportunities. The experiment consists of two treatments that differ only in the type of information offered to participants after the contribution stage. In the first treatment, upon deciding on how much to contribute to a public account, participants are informed of the individual contributions of the other group members (contribution feedback). In the second treatment, instead of receiving information regarding contributions, participants are informed about the payoffs of the other people in their group (payoff feedback). Based on this feedback individuals make their punishment decisions.

The rest of the paper is organized as follows. Section 2 discusses the experimental design and procedures. Section 3 presents the results from the experiments, while section 4 concludes by summarizing the evidence.

2. The experiment

2.1. Experimental design

For the purposes of our study we use the linear public-good game with groups of 4 players. In every period, each participant is given an endowment of 20 ECU (Experimental Currency Units). Players then decide simultaneously and without communication how much of the endowment to contribute to a public account, c_i, where $0\leq c_i \leq 20$. The rest ($20- c_i$) remains in the player's own account. In addition to the money that player *i* keeps, *i* receives a fixed percentage of the group's total contribution to the public account, , where $0<1<na$. In particular, a=0.4. This implies that the payoff of player *i* in period t is $20- c_i +0.4\sum c_h$.

After all participants decide on their contribution to the public account, groups enter a second stage where individuals can punish each other. The two experimental treatments differ only with regard to the information participants receive as feedback in the beginning of the punishment stage. In the baseline treatment, which we will refer to as CONTRIB, as in all previous public-good experiments, individuals are informed about how much the other individuals in their group contributed to a public account. In treatment, PAYOFF, participants are informed of the individual payoff of each group member. Based on these information, individuals can purchase punishment points, if they wish, to reduce the income of one or more other participants.

Punishment is costly for the punisher as every point reduces his income by 1 ECU (experimental currency unit). The income of the recipient of the punishment is reduced by 2 ECU. Let p_{ij} denote the number of punishment points that player i assigns to j (where i,j={1, ..., 4}; $j\neq i$). Player i's payoff at the end of period t is accordingly equal to $20- c_i +0.4\sum c_h -\sum p_{ij}-2\sum p_{ji}$.

The maximum number of points a participant can distribute to others is equal to his payoff from the first stage. As in stage one, punishment decisions are made simultaneously and without communication.

All treatments last for T=10 periods. The total payoff of a subject is equal to the sum of payoffs over the ten periods plus a one-off lump-sum payment of 25 ECU. This 25 ECU were given in the beginning of the experiment to pay for any negative payoffs the participant might have incurred in the duration of the experiment.

For the experiment we use fixed (or "partners") matching, that is, the group composition is the same in all periods. For each treatment, we have six groups with 4 subjects, giving us six statistically entirely independent observations. The reason for choosing fixed rather than random matching, where the group composition changes in every period, is that previous public-good experiments with punishment have shown that cooperation is sustained at higher levels when groups remain unchanged (e.g. Fehr and Gächter, 2000; Keser and van Winden, 2000). That is, under fixed matching both the punishment and cooperation levels are higher. This provides a greater challenge for our hypothesis.

Information feedback is as follows. Once the participants have contributed in stage one, they are informed about their group's total contribution to the public account and their payoff from the period as given by equation 1. In addition, in the beginning of stage two in CONTRIB (PAYOFF) individuals are informed of the individual contributions (payoffs). To prevent the formation of individual reputation, every player is randomly given a number between 1 and 4 at the beginning of each period to distinguish their actions (payoffs) in CONTRIB (PAYOFF) from those of the others in that period. Such a mechanism ensures that, even though the group members remain the same, the participants can not create a link between the actions of the other subjects across the periods. At the end of each period, participants are informed about the punishment points they received, the associated income reduction and their payoff from the period as given by equation 2. Subjects are not informed about the individual punishment decisions. They only know how many points they assigned to the other group members and the total number of points they received from the other group members. Retaliation, as in Nikiforakis (in press), is not possible.

The standard game-theoretic predictions are straightforward. As punishment is costly it should never occur in the subgame perfect equilibrium of a finitely repeated game. As a result, in the contribution stage individuals who maximize their own payoff should contribute zero to the public account irrespective of the nature of the feedback. Of course, as is well known, these predictions are violated. Based on previous experimental results, we hypothesize that payoff feedback will reduce cooperative behavior. This might lead to lower levels of cooperation in PAYOFF.

The experiments were conducted in the experimental laboratory of Royal Holloway, University of London, U.K. in February 2005 and March 2006. We ran four sessions, two for each treatment, with a total of 48 subjects. The subjects were recruited using an e-mail list of voluntary potential candidates. Participants were students from various fields. None of the participants had participated previously in a public-good experiment. Sessions lasted approximately fifty minutes and the average payment was £9.60 (£10.60 for CONTRIB, £8.60 for PAYOFF) or roughly $18.10. No show-up fee was given apart of the 25 ECU and the exchange rate between the experimental currency units and the British pound was 1 ECU = £0.04. The experiments were designed and conducted using z-Tree (Fischbacher, 2007).

3. Results

3.1. Punishment behavior

Figure 1 depicts the likelihood of being punished in each treatment as a function of the deviation from the contribution of the other group members, respectively. The figure provides some first evidence about what drives punishments in the experiment. In line with previous studies, it appears that the likelihood of being punished increases (almost linearly) with one's negative deviation from the average contribution of his peers. Individuals contributing more than their peers on average are also sanctioned.

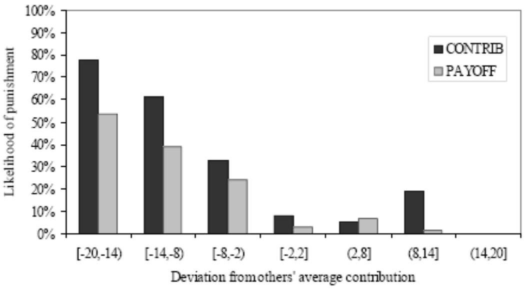

Figure 1: Punishment likelihood as a function of an agent's deviation from the average contribution of his peers]

To analyze formally punishment behavior we use a hurdle model specification with White (1980) robust standard errors. The punishment decision, that is, *i*'s decision whether to punish *j* or not, is captured by a probit regression. The dependent variable, Punish, is a dummy variable taking the value 0 if individual *i* did not punish *j* in a given instance and 1 if she did. The independent variables are PAYOFF - a dummy variable taking the value 1 if the observation comes from treatment PAYOFF and 0 if it comes from treatment CONTRIB; *Negative_Deviation_of_j*:= $\max\{0,(1/3)\sum c_h - c_j\}$, where c_j is the contribution of individual *j* in period *t*; *Positive_Deviation_of_j*:=$\max\{0, c_j - (1/3)\sum c_h\}$; *Total Contribution* - a variable about the total contribution of a group in period t; (*Total Contribution*)² - a variable to capture the non-linear relationship seen in Figure 1; and a variable to control for time effects (period). The punishment-level decision, that is, the decision of how much to punish conditional on the decision to punish, is captured by a truncated Tobit regression.

The results from the regression are presented in Table 1. PAYOFF is negative and significant indicating that individuals are less likely to punish under payoff feedback than under contribution feedback (p=0.55). To the extent that punishing is considered to be a second-order public good this behavior is in line with the hypothesis that information about individual payoffs undermines cooperation. The type of feedback does not appear to affect significantly the punishment-level decision (p=0.650). *Negative_Deviation_of_j* is positive and highly significant in both regressions (p<0.003). That is, the less an individual contributes compared to her peers, the more likely she is to be punished and the higher her punishment will be. *Positive_Deviation_of_j* is not found to affect the punishment decision (p=0.198), but it affects positively the punishment-level decision (p=0.33). Table 1 reveals that there is a concave relationship between total contribution and the likelihood of punishment and a convex relationship between total contribution and the punishment level. Both the likelihood of punishment and the size of the punishment appear to decrease over the course of the experiment. Result 5.1 states the main result.

Result 1: *The likelihood of being punished is significantly lower under payoff feedback.*

	Punishment Decision	Punishment Level
PAYOFF	-0.376*	0.316
	(0.196)	(0.696)
Negative Deviation j	0.114***	0.177***
	(0.014)	(0.057)
Positive Deviation j	-0.034	0.077**
	(0.026)	(0.036)
Total Contribution	0.030*	-0.065**
	(0.017)	(0.033)
$(\textit{Total Contribution})^2$	-0.0004**	0.0009***
	(0.0001)	(0.0003)
Period	-0.109***	-0.179*
	(0.028)	(0.100)
constant	-0.942**	2.581***
	(0.394)	(1.004)
N	1440	204
Wald χ^2	291.38	41.30

Table 1: Maximum likelihood estimates of the hurdle model of punishments

Before we examine contributions to the public account, let us investigate briefly whether contribution feedback might be affecting what Gächter and Herrmann (2005) call *misdirected punishments*, that is, punishments that an individual assigns to other group members who have contributed equal or greater amounts to the public account. This type of punishments constitutes a puzzle; why would individual *i* incur a cost to punish individual *j* who increases *i*'s payoff by contributing to the public account?

Since Ostrom et al. (1992) a number of explanations for the phenomenon of misdirected punishments have been suggested. Fehr and Gächter (2000; p.990) suggest that punishments of this type might be due to (i) errors; (ii) spite, that is, a willingness to increase one's relative advantage by punishing others; (iii) spiteful revenge, that is, *i* who has contributed substantially less than *j* anticipates a punishment from *j* and counter-punishes in advance; and (iv) blind revenge, that is, *i* punishes *j* believing that *j* (who contributed more than *i*) was responsible for a punishment *i* suffered in the previous period. One can think also of additional reasons for the occurrence of these punishments. For example, as the punishers remain anonymous, a group's lowest contributor might strategically punish the group's second lowest contributor in an attempt to force him to raise his contribution in the following period. The punished subject is likely to believe that the sanction came from the cooperators and that the other low contributor was also sanctioned.

Figure 2 shows the number of times that individual *i* punished *j*, when *i* had contributed more than *j*, less than *j* or the same amount as *j*. In CONTRIB, 32.8 percent of the total punishment cases took place when the punisher had contributed less than the victim of the punishment. In PAYOFF this happens only 12.7 percent of the time. However, more punishments occur in PAYOFF than in CONTRIB when the victim has contributed exactly the same amount as the punisher. The overall number of misdirected punishments as a percentage of the total punishments is substantially smaller in PAYOFF than in CONTRIB (21.5 and 33.6 percent respectively). The difference is significant (χ^2=3.44, d.f=1,p<0.1).[1]

Result 2: *The number of the punishment cases where the punisher contributes less or as much as the target of the punishment is significantly lower under payoff feedback.*

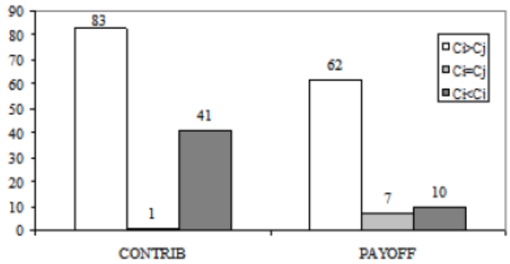

Figure 2: Number of punishments with respect to the relation between the contribution of the punisher, *i*, and the target, *j*

Contributions to the public account

Figure 2 presents the evolution of average contribution over the course of the experiment for each treatment. While average contribution remains roughly stable throughout the experiment at approximately 50 percent of the agents' endowment in CONTRIB, it monotonically declines after period 2 in PAYOFF.

The difference in average contributions between the two treatments is significant after period 5 using groups as independent observations (Mann-

[1] Due to the single observation where c_i=c_j in CONTRIB, we use a 2x2 table for the χ^2 test.

Whitney U test, two-tailed, p<0.1). Individuals appear to be unable to sustain cooperation through the use of punishments under payoff feedback. To emphasize the effect of payoff feedback, in Figure 2 we include the average contribution for the treatment without punishments from Nikiforakis and Normann (in press). It is clear that the decline in contributions after period 2 happens at a very similar rate in VCM and in PAYOFF, although cooperation levels are higher than they were in VCM.

Result 3: *Cooperation breaks down under payoff feedback while it is sustained under contribution feedback.*

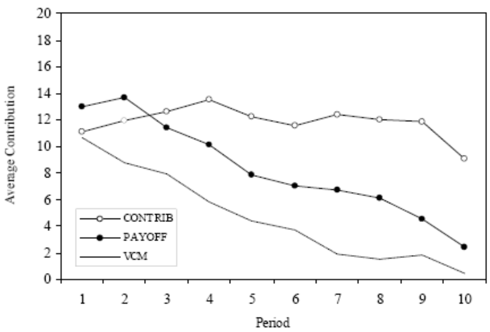

Figure 3: Contributions over time

Result 3 is in line with the hypothesis that information about individual profits would undermine cooperation and with the results of Huck et al. (1999;2000) and Offerman et al. (2002). One possible explanation for these findings is that payoff feedback allows for (or even encourages) payoff comparisons among individuals. These comparisons, in turn, trigger competition.

Could there be another reason for the breakdown of cooperation under payoff feedback? Table 2 provides information on each group separately. One explanation could be that contributions under payoff feedback where lower in period 1. This is clearly not the case. In period 1 contributions are slightly higher in PAYOFF than in CONTRIB. The difference, however, is not significant (Mann-Whitney U, p=0.266). Note that since in period 1 there has been no interaction between the participants the test is conducted regarding each of the 48 participants as an independent observation.

Treatment	Group	Mean contribution in period 1	Mean contribution all periods	Mean contribution periods 6-10	Mean payoff all periods	Mean payoff periods 6-10
CONTRIB	1	12.5	18.08	19	28.15	29.60
		(3.80)	(4.10)	(4.40)	(5.47)	(5.01)
CONTRIB	2	9.5	10.53	11.45	23.99	25.97
		(6.02)	(3.68)	(2.08)	(2.97)	(1.51)
CONTRIB	3	12.25	10.53	9.15	19.12	20.84
		(8.43)	(7.33)	(7.39)	(6.16)	(4.67)
CONTRIB	4	10	7.23	4.9	21.34	21.14
		(4.61)	(6.86)	(7.73)	(5.21)	(6.97)
CONTRIB	5	14.5	19.33	20	29.05	32.00
		(5.79)	(2.51)	(0)	(6.89)	(0)
CONTRIB	6	8	5.33	3.75	21.77	21.95
		(4.84)	(3.61)	(2.59)	(3.02)	(2.37)
CONTRIB	**average**	**11.13**	**11.83**	**11.38**	**23.90**	**25.25**
		(5.98)	**(7.20)**	**(7.97)**	**(6.30)**	**(5.98)**
PAYOFF	1	18.75	18.38	17	30.05	29.3
		(2.26)	(5.31)	(7.20)	(4.73)	(5.99)
PAYOFF	2	10	7.95	4.4	23.8	21.74
		(0)	(6.59)	(5.65)	(5.21)	(5.07)
PAYOFF	3	15	9.58	6.30	23.35	22.58
		(5.22)	(6.38)	(5.40)	(4.26)	(3.54)
PAYOFF	4	12.5	3.88	0.25	16.7	19.4
		(8.66)	(6.97)	(1.20)	(8.03)	(2.56)
PAYOFF	5	7.25	3.03	2.4	20.84	20.24
		(4.52)	(3.43)	(2.29)	(2.53)	(1.86)
PAYOFF	6	14.25	6.88	1.8	19.25	20.48
		(4.86)	(7.80)	(3.11)	(6.19)	(2.66)
PAYOFF	**average**	**12.96**	**8.28**	**5.36**	**22.33**	**22.29**
		(6.09)	**(8.01)**	**(7.22)**	**(6.86)**	**(5.09)**

Table 2: Group data

Messick et al. (1983) and Samuelson et al. (1984) hypothesize that the larger the variance within a group the weaker the cooperation norm in the group and as a result cooperation is more likely to break down. However, the variance in period 1 does not significantly differ across treatments in period 1 using groups as independent observations (Mann-Whitney U, p=0.394). Moreover the correlation between a group's average standard deviation and its average contribution is not significant (Spearman's , p=0.505).

The different effect of the two feedback types can be best understood by comparing group 5 from CONTRIB to group 6 from PAYOFF. The two groups have similar contribution rates in the first period and similar standard deviations. Moreover, the average number of punishment points assigned in the first two periods is very similar (1.42 in group 5 and 1.63 in group 6). However, while average contribution for group 5 in period 3 is already 20, average contribution for group 6 in period 3 is 12.25. In the final period, all members of group 5 (CONTRIB) contribute their whole endowment to the public account, while the average contribution in group 6 (PAYOFF) is at the other end of the action space at 0.5. The evolution of contributions in each group can be seen in Figure 4.

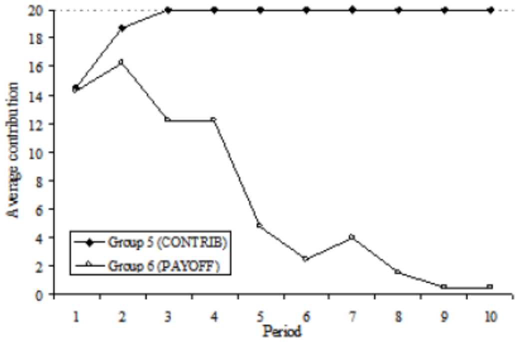

Figure 4: Evolution of average contribution in groups 5 and 6

Conclusion

In this paper we saw that the information given to participants as feedback after the contribution stage in a public-good experiment affects their behavior significantly. Feedback in the form of payoffs (payoff feedback) leads to lower levels of cooperation and a decreased propensity to punish compared to a treatment where feedback regarded information about others' contributions (contribution feedback). These results are consistent with previous studies showing that information about other participants' payoffs undermines cooperation (see Huck et al., 1999; 2000; Offerman et al., 2002). Payoff feedback is also found to reduce significantly the proportion of misdirected punishments, that is punishments where the target had contributed more than the punisher. Finally, we found no significant difference in payoffs between the two treatments.

Our results raise a number of questions. First, the results raise questions about the robustness of pro-social behavior observed in experiments, and, in particular, contributions and `altruistic' punishment in public good games. The results also question the ability of decentralized punishments to foster cooperation. Clearly, more experiments are needed before we are able to provide definitive answers.

References:

Fehr, E., Gächter, S., (2000). Cooperation and Punishment in Public Goods Experiments, *American Economic Review* 90 (4), 980--994.

Fehr, E., Schmidt, K., (1999). A Theory of Fairness, Competition and Co-operation, *Quarterly Journal of Economics* 114, 817--868.

Huck S., Normann, H.T, Oechssler, J., (1999). Learning in Cournot oligopoly: An experiment, *Economic Journal* 109, C80-C95.

Huck S., Normann, H.T, Oechssler, J., (2000). Does information about competitors' actions increase or decrease competition in experimental oligopoly markets?, *International Journal of Industrial Organization* 18, 39-57.

Keser, C., van Winden, F., (2000), Partners Contribute More than Strangers: Conditional Cooperation, *Scandinavian Journal of Economics* 102 (1), 23-39.

Nikiforakis, N. (in press), Punishment and Counter-punishment in Public Good Games: Can we Really Govern Ourselves?, forthcoming: *Journal of Public Economics.*

Nikiforakis, N., Normann, H.T. (in press), A Comparative Statics Analysis of Punishment in Public-Good Experiments, forthcoming in: *Experimental Economics*

Offerman, T., Potters, J., Sonnemans, J., (2002), Imitation and Belief Learning in an Oligopoly Experiment, *The Review of Economic Studies* 69 (4), 973-997.

Ostrom, E., Walker, J., Gardner, R., (1992), Covenants with and without a Sword: Self Governance is Possible, *American Political Science Review* 86, 404--417.

Messick, D. M., Wilke, H. A. M., Brewer, M. B., Kramer, R., Zemke, P. E., and Lui, L., (1983), Individual adaptations and structural changes as solutions to social dilemmas. *Journal of Personality and Social Psychology* 44, 294-309.

Samuelson, C.D, Messick, D. M., Rutte, C., & Wilke, H. A. M. (1984). Individual and structural solutions to resource dilemmas in two cultures. *Journal of Personality and Social Psychology* 47, 94-l04.