Discussion Paper Series

2005 – 07

Department of Economics
Royal Holloway College
University of London
Egham TW20 0EX

A Comparative Statics Analysis of Punishment in Public-Good Experiments

Nikos Nikiforakis*and Hans-Theo Normann†

June 2005

Abstract

This paper provides a comparative statics analysis of punishment in public-good experiments. We vary systematically the effectiveness of punishment, that is, the factor by which punishment reduces the punished player's income, and we find that contributions to the public good increase monotonically in effectiveness. High effectiveness leads to near complete contribution rates and welfare improvements. Below a certain threshold, however, punishment cannot prevent the decay of cooperation found in the public-good game without punishment. In these cases, the possibility to punish may even worsen welfare. Finally, we show that punishment is a normal and inferior good.

*Department of Economics, Royal Holloway, University of London, Egham, Surrey, TW20 0EX, UK, fax +44 1784 439534, email *n.nikiforakis@rhul.ac.uk*

†Department of Economics, Royal Holloway, University of London, Egham, Surrey, TW20 0EX, UK, fax +44 1784 439534, email *hans.normann@rhul.ac.uk*

1 Introduction

The public-good game is one of the most frequently used institutions for analyzing cooperation in experiments. In the public-good game, players decide how much of their endowment they wish to contribute to a public account and how much they wish to keep to themselves. Public and individual interest are at odds. The individual income decreases when the player's contribution to the public good increases but group welfare is maximized when players contribute their entire endowment. A large number of experimental studies has shown that contributions quickly decline over the course of the experiment and there is nearly complete free riding towards the end (see for example Ledyard, 1995).

One approach to solve the free-rider's problem in public-good experiments that has received a lot of attention recently is to allow players to monitor the members of their group and punish each other. Punishment is decentralized in these experiments, that is, it is carried out by individuals which are not governed by a central authority. A punishment mechanism was first used by Ostrom, Walker and Gardner (1992) in a common-pool resource setting and by Fehr and Gächter (2000) in a public-good experiment.[1] In both experiments, subjects can give costly punishment "points" to the other members of their group which reduce the recipient's income. The results show that the ability of players to monitor and punish each other improves the level of cooperation significantly.

The experiment by Fehr and Gächter (2000) has been the starting point for a rapidly increasing number of studies.[2] However, despite the large number of experiments, very little is known about the robustness of the punishment schemes and the requirements on them to lead to higher contribution rates. To our knowledge, there is no systematic comparative statics analysis of punishment schemes. Carpenter (2005a) emphasizes this point and concludes that the existing literature resembles "a series of unconnected islands". We review some of this scattered evidence in the conclusion.

This paper aims at providing a first comprehensive comparative statics analysis of public-good games with punishment. We study four variants of a punishment scheme. We use the simple linear punishment technology employed in several recent papers (Fehr and Gächter, 2002; Sefton et al., 2002; Anderson and Putterman, forthcoming; Carpenter, 2005a). The linear scheme implies that the punishment points chosen by the punishers reduce the punished player's income by a constant factor.[3] Our treatment variable is this very income reduction factor which we term punishment

[1]Beil (1988) is a Cournot experiment with punishment. Punishment is not costly for the punisher here—in contrast to most of the public goods experiments with punishment. Beil (1988) is discussed in detail in Holt (1995).

[2]See, for example, Bowles, Carpenter and Gintis (2001), Fehr and Gaechter (2002), Sefton et al. (2002), Masclet et al. (2003), Denant-Boemont et al. (2005), Nikiforakis (2005), Anderson and Putterman (2005), Bochet et al. (2005), Carpenter (2005a, 2005b).

[3]In some studies, a punishment point reduces the recipient's income by a percentage, typically 10% (Fehr and

effectiveness. More specifically, we vary punishment effectiveness from one to four, keeping all other aspects of the design constant. To have a basis for comparison, we also run the standard public-good game without punishment opportunities.

The results indicate that the experimenter's choice of the punishment effectiveness is of great importance for the experimental outcome. Contributions to the public account increase monotonically in the effectiveness of punishment. We show that a minimum level of punishment effectiveness (three) is required to significantly raise contributions over time. Below this level, contributions remain constant or even decline over time. A clear pattern emerges also in the relation of punishment effectiveness and welfare. In the treatment with low punishment effectiveness (one), welfare is lower compared to the public good game without punishment. A punishment effectiveness of two is required so that the benefits of higher contribution rates offset the costs of punishment. Only punishment effectiveness of three and four leads to significant welfare improvements. Finally, we also find that the subjects' demand for punishment obeys the law of demand, as previously found by Anderson and Putterman (forthcoming) and Carpenter (2005a).

The paper is organized as follows. Section 2 presents the experimental design and the procedures. Section 3 contains the predictions and Section 4 illustrates the experimental results. Section 5 concludes.

2 Experimental Design and Procedures

The experiment is based on the linear public-good game with n players also known as the voluntary contribution mechanism. In every period, each participant is given an endowment y. Players then decide simultaneously and without communication how much of the endowment to contribute to a public account, c_i, where $0 \leq c_i \leq y$. The rest $(y - c_i)$ 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, α, where $0 < \alpha < 1 < na$. This implies that the payoff of player i in period t is

$$\pi_i^t = y - c_i + \alpha \sum_{i=1}^{n} c_i. \tag{1}$$

In the treatments with punishment opportunities, a second stage is added. After the participants decide how much to contribute to the public account, they are informed about how much the other individuals in their group contributed. They can then, if they wish, purchase punishment points to reduce the income of one or more other participants. Punishment is costly for the

Gächter, 2000; Masclet et al., 2003; Nikiforakis, 2005). The effectiveness of the punishment points is convex in these studies which makes this technology somewhat less attractive for comparative statics analysis.

punisher as every point reduces his income by 1 ECU (experimental currency unit). Let p_{ij} denote the number of punishment points that player i assigns to j (where i, j=1, ..., n; $j \neq i$), and e the reduction that one punishment point causes to its recipient. Player i's payoff at the end of period t is accordingly equal to

$$\pi_i^t = y - c_i + \alpha \sum_{i=1}^{n} c_i - \sum_{j \neq i} p_{ij} - e \sum_{j \neq i} p_{ji}. \tag{2}$$

The maximum number of points a participant can distribute to others is equal to his payoff from the first stage, that is, $\sum_{j \neq i} p_{ij} \leq y - c_i + \alpha \sum_{i=1}^{n} c_i$. As in stage one, punishment decisions are made simultaneously and without communication.

In all treatments, it is common knowledge that y=20, n=4 and α=0.4. The treatment variable is e, the effectiveness of punishment. We have $e \in \{1, 2, 3, 4\}$ and the treatments are labelled "1", "2", "3" and "4" accordingly. So, for example, one punishment point in treatment "3" reduces the income of its recipient by 3 ECUs. As a control treatment, we use the public good game without punishment. We label this treatment "0".

All treatments last for T=10 periods. The total payoff of a subject is equal to the sum of payoffs over all the ten periods. In the treatments with punishment opportunities, each subject is given a one-off lump-sum payment of 25 ECU in the beginning of the experiment to pay for any negative payoffs the participant might incur 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 of size n=4, 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 participants use punishment more frequently when groups remain unchanged. As a result the effect of punishment on contribution rates has been more pronounced under fixed matching and we expect a clearer and stronger effect as observed, for example, in the "partners" sessions of Fehr and Gächter (2000).[4]

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). 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 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

[4] The main contribution of the public-good experiments with punishment under random (or "strangers") matching is probably to show that costly punishments still occur even though there is random matching. See, for example, Fehr and Gächter (2000).

of the other subjects across the periods. In the treatments where punishment is available, 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 retaliation is not possible.[5]

The experiments were conducted in the experimental laboratory of Royal Holloway, University of London, in January and February 2005.[6] We ran a total of ten sessions (two for each treatment) with a total of 120 subjects. The subjects were recruited using an email list of voluntary potential candidates. Participants were from a variety of backgrounds. None of the participants had participated previously in a public-good experiment. Sessions lasted approximately fifty minutes and the average payment was £11.10 or roughly $20.90, which is more than double than the minimum hourly rate. No show-up fee was given 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 a computer program written zTree (Fischbacher, 1999).

3 Predictions

Consider the treatments with punishments first. Suppose players have reached the final period, T. Punishment at the second stage cannot have an impact on future contributions and it is costly for the punisher. Hence, $p_{ij} = 0$, i, j=1, ..., 4; $j \neq i$, follows. At the contribution stage of T, knowing that punishment threat is non-credible, players will choose $c_i = 0$ for i, j=1, ..., 4. Deviating from $c_i = 0$ is not profitable due to $\partial \pi_i^t / \partial c_i = -1 + \alpha < 0$. Hence, in the subgame perfect Nash equilibrium of period T, we have $c_i = 0$ and $p_{ij} = 0$. By backward induction, the unique subgame perfect Nash equilibrium of the finitely repeated game is the static subgame perfect Nash equilibrium of the last period, implying zero contributions and no punishments in all periods. Since there are no punishments in the subgame perfect Nash equilibrium regardless of punishment effectiveness, treatment variable e does not have an effect in equilibrium.

In addition to the unique subgame perfect Nash equilibrium of the repeated game, there may be other imperfect Nash equilibria in which cooperation can occur and in which punishment and punishment effectiveness may play a role. Ostrom, Walker and Gardner (1992) derive such imper-

[5]Nikiforakis (2005) studies the effect of counter-punishment in a public good game by adding a second punishment stage. He finds that the opportunity to counter-punish reduces cooperation significantly. Furthermore, the experiments with punishment and counter-punishment have efficiency levels lower than the public good game without punishment.

[6]One session for treatment "4"was run in the University College of London due to shortage of subjects who had not participated in public-good experiments. Following the custom of the University College of London these subjects where given a show-up fee of £5.

fect Nash equilibria for a common pool game with punishment possibilities. Intuitively, the higher e, the harsher the threat of punishment and the easier cooperation of this type can be sustained. Such Nash equilibria are imperfect because they are based on non-credible threats.[7]

Punishments and punishment effectiveness are also predicted to have an effect in the model of other-regarding preferences developed by Fehr and Schmidt (1999). In their model, people receive utility from their individual material income as in the standard model but their utility might be reduced from inequitable distributions of income. Fehr and Schmidt (1999, section IV) discuss predictions for a public-good game with punishment and include predictions about the relation between the cost of punishment (that is, the amount of ECUs that a player must pay to reduce one's income by 1 ECU—in our case $1/e$) and punishment behavior. In particular, Fehr and Schmidt (1999) show that, as e increases, higher levels of cooperation should emerge.[8]

We conclude that the unique subgame perfect equilibrium predicts no cooperation and no treatment effects. Imperfect Nash equilibria and the Fehr and Schmidt (1999) model suggest that cooperation is more likely the higher punishment effectiveness e.

Finally, in treatment "0" without punishment, similar arguments as in the first paragraph show that $c_i = 0$ is the unique Nash equilibrium of the stage game. By backward induction, the unique subgame perfect Nash equilibrium of the finitely repeated game has $c_i = 0$ in all periods.

4 Results

We start by reporting results on contributions to the public good, followed by a welfare comparison of the treatments and an investigation of punishment behavior. Unless otherwise noted, the statistical tests below use each group as one observation (so that there are six observations for each treatment). The group data are summarized in Table A1 in the appendix. Since we have $k > 2$ treatments, we report Kruskal-Wallis tests when we compare means, and we proceed to pair wise treatment comparisons with Mann-Whitney U (MWU) tests only if a Kruskal-Wallis tests suggests significant differences.

[7]Ostrom, Walker and Gardner (1992) propose a symmetric strategy which we modify here for our game. In every period except T, contribute $c_i = y$ and choose $p_{ij} = 0$, $j \neq i$. In the event of a deviation, play $c_i = 0$ and punish all players with $p_{ij} = \widetilde{p} > 0$, $j \neq i$, for one period, then resume to $c_i = y$ and $p_{ij} = 0$. In period T, play the static Nash equilibrium unless a deviation occurs in $T - 1$. This strategy is a Nash equilibrium as long as the punishment following a deviation is sufficiently severe. Since the punishments points $\widetilde{p}$ assigned after a deviation must not exceed the minimum stage one income (that is, $\widetilde{p}(n-1) \leq \alpha y$), the harshness of punishments depends primarily on e. Therefore, the higher e, the easier cooperation of this type can be sustained.

[8]Fehr and Schmidt's (1999) prediction for public-goods games with punishment depends on a number of further assumptions like the existence of "conditionally cooperative enforcers" and the distribution of α and β parameters, which are the coefficients of disadvantageous and advantageous inequality, in the subject pool. For details see Fehr and Schmidt (1999, section IV).

Figure 1 presents the mean contributions over all periods of our five treatments. Figure 1 reveals that, as the effectiveness of punishment increases, so does the mean contribution. The relationship between effectiveness and contributions is monotonic; we never find that higher effectiveness leads to lower contributions. Average treatment contribution rates and treatment effectiveness are positively correlated (Spearman correlation, two sided, p=0.01). The differences in mean contributions across the treatments are also statistically significant (Kruskal-Wallis, $d.f.$=4, p=0.01).[9] Further, Table A1 in the appendix reveals that the effectiveness of punishment has an even more pronounced effect on contributions in the second half of the experiment. We summarize

Result 1: *Average contributions monotonically increase in the effectiveness of punishment.*

The evolution of average contributions over time is illustrated in Figure 2. In period one, subjects contribute on average between 40% and 63% of their endowment, consistent with previous public-good experiments with and without punishment. Period one contributions are not significantly different across treatments (Kruskal Wallis, $d.f.$=4, p=0.610).[10] From period two onwards, contributions decrease in treatments "0" and "1", are roughly constant in "2", whereas they increase in treatments "3" and "4". The time trends are significant except for treatment "2" (Spearman correlation, two tailed, $p < 0.05$).[11] Remarkably, Result 1 (monotonicity) holds in every period except for the first. Note also that there is a significant end-game effect in the last period.[12]

Result 2: *Average contributions decrease over time in "0" and "1", they are constant in "2", whereas they increase in "3" and "4". There is a drop in contributions in the last period.*

Table 1 presents the percentage of people who chose not to contribute to the public account in accordance with the subgame perfect Nash equilibrium (that is, $c_i = 0$), as well as the percentage

[9] Most pair wise comparisons confirm the monotonic relationship between effectiveness and contributions and none contradicts it. All treatments have significantly higher contribution rates than "0" (one-sided MWU, p=0.01 for "2", "3" and "4"; p=0.09 for "1"). Treatments "2", "3" and "4" have significantly higher contribution rates than "1" (one-sided MWU, p=0.05 for "3" and "4"; p=0.09 for "2"). Finally, treatment "4" has significantly higher contributions than "2" (one-sided MWU, p=0.07). The results for averages from periods 1 to 10 do not differ much from those reported here which refer to average contribution between periods 6 to 10.

[10] In period one, all players' contributions are statistically independent and we can use 24 observations for each treatment (instead of just six group averages). Period one contributions do still not differ significantly (Kruskal Wallis, $d.f.$=4, p=0.213).

[11] Treatment "0": ρ=−0.988, $\rho < 0.001$ treatment "1": ρ=−0.842, p=0.002; treatment "2": ρ=−0.489, p=0.248; treatment "3": ρ=0.730, p=0.017; treatment "4": ρ=0.679, p=0.025.

[12] We compared average contributions in periods six to nine to the average in period ten. Period ten contributions are significantly lower if we take all treatments together (Wilcoxon, $p < 0.01$) and also for individual treatments "0" (Wilcoxon, p=0.028), "1" (Wilcoxon, p=0.028) and "3" (Wilcoxon, p=0.043).

of people who contributed their whole endowment (that is, $c_i = 20$) in periods 9 and 10 of the experiment. Again, a clear pattern emerges. As the effectiveness of punishment increases, so does the number of subjects who contribute the whole of their endowment, whereas the number of participants not contributing decreases. On one extreme, in the last two periods of the treatment without punishment opportunities a mere 2% contributes the whole of their endowment, while 79% decide not to contribute at all. On the other extreme, in the last two periods of "4" the percentage of participants contributing the whole of their endowment increases to 75%, whereas only 6% contribute nothing. Result 3 summarizes

Result 3: *The effectiveness of punishment has a strong influence on the behavior in the last two periods. In treatments "0" and "1" complete free-riding emerges as the mode whereas in treatments "3" and "4" full contribution is the mode.*

We now turn our attention to subjects' payoffs. As noted by previous research, higher contributions do not necessarily imply higher payoffs in experiments with punishment. The reason is that punishment is costly for both the punisher and the punished which might offset the welfare gain from higher contributions. An interesting question to be answered, therefore, is how the effectiveness of punishment influences the welfare of the groups (the sum of payoffs is equal to social welfare here). To answer this question, we calculate the cumulative relative payoff (CRP) between treatments with punishment and treatment "0". This is done using the following formula

$$CRP_\tau^t = \left(\sum_{s=1}^{t} \bar{\pi}_\tau^s - \sum_{s=1}^{t} \bar{\pi}_0^s \right) \Big/ \sum_{s=1}^{t} \bar{\pi}_0^s \tag{3}$$

where CRP_τ^t is the cumulative relative payoff of treatment $\tau \in \{1, 2, 3, 4\}$ compared to treatment "0" up to period t, and $\sum_{s=1}^{t} \bar{\pi}_\tau^s$ is the average cumulative payoff up to period t of treatment τ.

The evolution over time of the cumulative relative payoff can be seen in Figure 3. In period one, the treatments with punishment opportunities have a lower payoff than the groups in treatment "0". This is the due to the similar contributions across treatments but punishment reduces the payoffs in all treatments except "0". The difference in period one payoffs is significant (Kruskal Wallis, $d.f.$=4, $p = 0.007$). Pair wise comparisons show that treatment "0" has significantly higher period one payoffs than all treatments with punishment except "4" (two sided MWU, p ¡ 0.05).[13] From period two onwards, the cumulative relative payoff in all treatments increases compared to "0". These time trends are significant for all treatments (Spearman correlation, $r > 0.970$, $p < 0.01$). In other words, as the experiment proceeds, the benefits from higher contributions tend to outweigh the costs from punishing.

[13] Payoffs in treatment "4" are also significantly higher than averages in "2" (two sided MWU, p=0.09)

However, as Figure 3 illustrates, the increase in the cumulative average group payoff in treatment "1" in comparison to "0" is not enough to offset the losses of income that took place in the first periods. In treatment "2", the higher contributions manage to offset the costs of punishment only in the last two periods. This holds true for treatment "3" from period seven onwards. Finally, we see treatment "4" is the most efficient as the total average group payoff was higher than in treatment "0" already from period four on. The difference in the total payoff across all ten periods, as observed in t=10 in Figure 3, is statistically significant (Kruskal-Wallis, $d.f.$=4, p=0.02).[14]

Result 4: A punishment effectiveness of "3" or greater is required to have a welfare improvements compared to the public good game without punishment (treatment "0").

Table A1 in the appendix shows that payoffs in all treatments with punishment are higher than "0" in periods 6 to 10. This implies a qualification of Result 4. As noted by Fehr and Gächter (2000), if behavior in the second half of the experiments is the steady state, then experiments with a horizon sufficiently longer than T=10 would unambiguously yield higher welfare for the treatments with punishment.

We now turn to participants' punishment behavior, in particular with respect to punishment effectiveness. Anderson and Putterman (forthcoming) and Carpenter (2005a) have analyzed punishment behavior in great detail.[15] We only state a few basic results here.

Across the four treatments where punishment is allowed, 84.4% (or 81 out of 96 participants) of the participants sanctioned another group member at least once, and 62.5% (60 out of 96 participants) of the subjects punished more than once. In the last period, approximately 16% punished another group member (two out of 24 subjects in "1", three in "2", five in "3", and nine out of 24 subjects in "4"). The income reduction of the punished players was as follows. In periods one to five, the punished players' incomes were reduced by 230 ("1"), 386 ("2"), 489 ("3"), and 280 ("4"). In the second half of the experiment, these figures were reduced to 70 ("1"), 126 ("2"), 117 ("3"), and 128 ("4"). In the last period in treatments "2", "3" and "4" (but not in "1"), players answer the drop in contribution with more punishment than in any other period of the second half of the experiment. Punishments are to a large extent not in accordance with the model of Fehr and Schmidt (1999). In particular, the total number of points in periods 1 to 5 is 230 ("1"), 193 ("2"), 163 ("3"), and 70 ("4"), while in periods 6 to 10 the points purchased were 70 ("1"), 63

[14] Pair wise comparisons reveal that treatments "3" and "4" have significantly higher cumulative group payoff in period ten than both "0" and "1" (two sided MWU, p¡0.05). Further, period ten cumulative payoffs in "4" are significantly higher than "2" (two sided MWU, p=0.07).

[15] Ostrom, Walker and Gardner (1992) also varied the cost or effectiveness of punishment but not many details are available from the paper. They used punishment cost factors (or "fee-to-fine ratio" in their terminology) of 1/2 and 1/4, and they report that "frequency is inversely related to cost".

("2"), 39 ("3"), and 32 ("4"). According to the model of Fehr and Schmidt (1999) subjects should never punish in treatments "1" and "2" and sometimes in treatment "3" as punishing increases average inequality.

To analyze the determinants of punishment in detail, we run a regression with the punishment inflicted on the punished player as the dependent variable. The independent variables include the effectiveness of punishment (e), the punisher's income at the beginning of the second stage, the negative and positive deviation of the punished player's contribution from the average contribution of the other three group members, the group's average contribution in the same period and a variable for the period to capture any time trends. The results in Carpenter (forthcoming) suggest a non-linear relationship between the inflicted punishment and the punishment effectiveness, so, we also include a squared term for e.

Table 2 presents the results from an OLS regression with robust standard errors (White, 1980). As it can be seen from Table 2 there is a positive and concave relationship between the effectiveness of punishment and the punishment that i inflicts on j. The parameters suggest that increases in e up to three lead to an increase in the inflicted punishment and that further increases of e decrease the punishment inflicted on the subjects. Given the fact that the cost of punishment is $1/e$, we can verify previous findings of the literature that punishment is a normal good. Also in line with the existing literature, income is a significant determinant of the dependent variable. The negative sign implying that punishment is an inferior good is easily explained by the fact that the subjects with high income at the beginning of the second stage are typically free-riders who have less reason to punish.[16] Table 2 also illustrates that contributing more than the other members of one's group appears to marginally decrease punishments though this effect is not significant. On the other hand, contributing less than the other members of one's group increases the punishment inflicted. Finally, the higher the contribution of a group the lower the punishment, which also appears to be decreasing over time.

Result 5: *Punishment is a normal and inferior good.*

5 Conclusion

This paper provides a comparative statics analysis of punishment possibilities in public-good experiments. Using the simple linear punishment technology proposed by Fehr and Gächter (2002), we systematically vary punishment effectiveness—the factor by which one punishment point reduces

[16]Typically in public-good experiments, a small fraction of participants tend to punish high contributors although they contributed less than the average of the group. This phenomenon is extensively analyzed in Cinyabuguma et al. (2004).

the punished player's income. We find that punishment has to be sufficiently effective (or, in other words, cheap) for the punisher in order to significantly raise contributions. With a punishment effectiveness of two or less, contributions remain constant at best or decline over time and welfare is not improved compared to the public-good game without punishment. Only a punishment effectiveness of three and four leads to high contribution rates and significant welfare improvements.

The impact punishment effectiveness has on cooperation is not consistent with the unique subgame perfect Nash equilibrium of the game where players do not punish and keep the entire endowment. It is however, consistent with imperfect cooperative Nash equilibria as suggested by Ostrom, Walker and Gardner (1992) and the model proposed by Fehr and Schmidt (1999).

Our results confirm several previous studies. Sefton et al. (2002) used a punishment effectiveness of one, and Fehr and Gächter (2002) employ a punishment effectiveness of three. Consistent with our results, Sefton et al. (2002) find no significant differences between the standard public good treatment and the treatment with sanctions whereas Fehr and Gächter (2002) find that the opportunity to punish increases cooperation significantly. Note that the data in Sefton et al. (2002) and Fehr and Gächter (2002) cannot be compared directly as there are numerous other differences in the design. Finally, the high contributions in the "partners" sessions of Fehr and Gächter (2000) is also consistent with our findings.[17]

Two further related papers are Anderson and Putterman (forthcoming) and Carpenter (2005a). These papers vary the cost of punishment, but, in contrast to our approach, they mainly vary the cost of punishment *within* treatments and keep average cost constant. Carpenter (2005a) varies the cost of punishment across time. In one treatment the cost is increasing over the course of the experiment, whereas in the other it is decreasing. Anderson and Putterman (forthcoming) assign an individually drawn random cost parameter to each player in every period, controlling for expected average cost. Both papers find that subjects' demand for punishment obeys the law of demand, and we were able to confirm this with our data. Anderson and Putterman (forthcoming) additionally vary the expected average cost across treatments; high, medium or low. Whereas average contribution rates do not differ much across treatments, interestingly, efficiency and earnings differ considerably.[18] Both results are consistent with our findings. They find that the treatment with the highest contributions has got the lowest efficiency level.

[17] Fehr and Gächter (2000) have convex cost of punishment and a percentage reduction of the income of the punished player. However, Fehr and Schmidt (1999) note that average punishment effectiveness in Fehr and Gächter (2000) was four.

[18] Anderson and Putterman (forthcoming) report that contributions in the low expected average cost treatment are significantly higher compared to the other two treatments, and efficiency are also significantly different.

References

[1] Anderson, C. and Putterman, L. (forthcoming). Do non-strategic sanctions obey the law of demand? The Demand for Punishment in the Voluntary Contribution Mechanism, Games and Economic Behavior.

[2] Beil, R.O. (1988). Collusive Behavior in Experimental Oligopoly Markets, Working Paper, Texas A&M.

[3] Bochet O., Page T., and Putterman L. (2005). Communication and Punishment in Voluntary Contribution Experiments, Journal of Economic Behavior and Organization.

[4] Bowles, S., Carpenter J. and. Gintis, H. (2001). Mutual Monitoring in Teams: Theory and Evidence on the Importance of Residual Claimancy and Reciprocity, mimeo.

[5] Carpenter, J. (2005a). The Demand for Punishment, forthcoming in: Journal of Economic Behavior and Organization.

[6] Carpenter, J. (2005b) Punishing Free-Riders: How Group Size Affects Mutual Monitoring and the Provision of Public Goods, forthcoming in: Games and Economic Behavior.

[7] Cinyabuguma M., Page T. and Putterman L. (2004). On Perverse and Second-Order Punishment in Public Goods Experiments with Decentralized Sanctions, mimeo.

[8] Denant-Boemont L., Masclet D., and Noussair C. (2005). Anonymity in Punishment, Revenge and Cooperation: A Public Good Experiment, mimeo.

[9] Fehr, E. and Gächter, S. (2002). Altruistic Punishment in Humans, Nature 415, 137-140.

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

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

[12] Fischbacher, U. (1999). z-Tree: Zurich Toolbox for Readymade Economic Experiments - Experimenter's Manual, mimeo, Institute for Empirical Research in Economics, University of Zurich.

[13] Holt, C.A. (1995) Industrial Organization: A Survey of Laboratory Results, in Handbook of Experimental Economics, edited by J. Kagel and A. Roth, Princeton, N.J.: Princeton University Press, 349-443.

[14] Ledyard, J. (1995). Public Goods: A Survey of Experimental Research, in: Handbook of Experimental Economics ed. by J. Kagel and A. Roth, Princeton University Press.

[15] Masclet, D., Noussair, C., Tucker, S. and Villeval M.C. (2003). Monetary and Non-monetary Punishment in the Voluntary Contributions Mechanism, American Economic Review 93, 366-380.

[16] Nikiforakis, N. (2005). Punishment and Counter-punishment in Public Good Games-Can we still govern ourselves?, mimeo.

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

[18] Sefton, M., Shupp, R. and Walker J. (2002). The Effect of Rewards and Sanctions in Provision of Public Goods, mimeo.

[19] White, H. (1980). A Heteroskedasticity-Consistent Covariance Matrix Estimator and Direct Test for Heteroskedasticity, Econometrica 48, 817-838.

Appendix

Instructions

These are the instructions from treatment "3". The instructions for the other treatments were appropriately adjusted.

You are now taking part in an economic experiment. If you read the following instructions carefully, you can, depending on your decisions and of those made by the others, earn a considerable amount of money. It is therefore important that you take your time to understand the instructions.

The instructions which we have distributed to you are for your private information. Please do not communicate with the other participants during the experiment. Should you have any questions please ask us.

During the experiment we shall not speak of Pounds, but of Experimental Currency Units (ECU). Your entire earnings will be calculated in ECUs. At the end of the experiment the total amount of ECUs you have earned will be converted to Pounds at the rate of 1 ECU = 4p and will be immediately paid to you in cash.

At the beginning of the experiment the participants will be randomly divided into groups of four. You will therefore always be in a group with 3 other participants. The groups will remain the same throughout the experiment. This means that you will always interact with the same three participants. The experiment lasts 10 periods and each period is divided in two stages.

The first stage:

In every period, at the beginning of the first stage each participant will receive 20 ECUs. In the following, we shall refer to this amount as the "endowment". Your task is to decide how to use your endowment. In particular, you have to decide **how many of the 20 ECUs you want to contribute to a project (from 0 to 20) and how many of them to keep for yourself**. The consequences of your decision are explained in detail below.

Once all the players have decided their contribution to the project you will be informed about, the group's total contribution, your income from the project and your payoff in this stage. Your payoff from the first stage in each period is calculated using the following formula. If you have any difficulties, do not hesitate to ask us.

Income at the end of the stage = Endowment of ECUs - Your contribution to the Project + 0.4 * Total contribution to the Project

This formula shows that your income at the end of the period consists of two parts:

1) The ECUs which you have kept for yourself (endowment - contribution)

2) The income from the project, which equals to the 40% of the group's total contribution.

The income of each group member from the project is calculated in the same way. This means that each group member receives the same income from the project. Suppose the sum of the contributions of all group members is 60 ECUs. In this case, each member of the group receives an income from the project of: 0.4 * 60 = 24 ECUs. If the total contribution to the project is 9 ECUs, then each member of the group receives an income of: 0.4 * 9 = 3.6 ECUs from the project, regardless of how much they individually contributed to the project.

You always have the option of keeping the ECUs for yourself or contributing them to the project. Each ECU that you keep raises your end of period income by 1 ECU. Supposing you contributed this point to the project instead, then the total contribution to the project would rise by 1 ECU. Your income from the project would thus rise by 0.4 * 1 = 0.4 ECU. However, the income of the other group members would also rise by 0.4 ECUs each, so that the total income of the group from the project would be 1.6 points. Your contribution to the project therefore also raises the income of the other group members. On the other hand you also earn an income for each point contributed by the other members to the project. In particular, for each point contributed by any member you earn 0.4 ECUs.

In addition to the 20 ECUs per period, each participant receives a one-off lump sum payment of 25 ECUs at the beginning of this part of the experiment. Note that this lump sum payment should not be used to calculate the "End of period income". It will only be added to your total income from all the periods at the very end.

The second stage:

At the second stage of each period you will be informed how much each group member contributed individually to the project at the first stage. At this stage **you can reduce or leave equal the income of each member of your group**. The other group members can also reduce your income if they wish to.

To reduce another player's income you will have to distribute points. Each point will cost you 1 ECU and will reduce the income of the person you assign it to by 3 ECU. If you choose 0 points for a particular group member, you do not change his or her income.

Example: Supposing you give 2 points to player 1 this costs you 2 ECU and reduces player 1's income by 6 ECU.

Your total income from the two stages is therefore calculated as follows:

Total income (in ECUs) at the end of the period = Income from the 1st stage - 3*Points you receive - Points you give

Please note that your income in ECUs at the end of the period can be negative. If your income becomes zero or negative at the end of the second stage you can simply use your 25 ECUs that we gave you in the beginning in order to pay this off.

If you have any further questions please raise your hand and one of the supervisors will come to help you.

Control Questionnaire

1. Each group member has an endowment of 20 ECUs. Nobody (including you) contributes any ECUs to the project. How high is:

a. Your income at the end of the period?

b. The income of the other group members

2. Each group member has an endowment of 20 ECUs. You contribute 20 ECUs to the project. All other group members contribute 20 ECUs each to the project. What is:

a. Your income at the end of the period?

b. The income of the other group members?

3. Each group member has an endowment of 20 ECUs. The other three group members contribute together a total of 30 ECUs to the project. What is:

a. Your income at the end of the period if you contribute 0 ECUs to the project?

b. Your income at the end of the period if you contribute 15 ECUs to the project?

4. Each group member has an endowment of 20 ECUs. You contribute 8 ECUs to the project. What is:

a. Your income at the end of the first stage if the other group members together contribute a further total of 7 ECUs to the project?

b. Your income at the end of the first stage if the other group members together contribute a further total of 22 ECUs to the project?

5. Your income from the first period is 25 ECU. How much will your income at the end of the period be if:

a. You receive 2 points, but do not assign any yourself?

b. You receive 2 points and assign 3 points yourself?

	"0"	"1"	"2"	"3"	"4"
$c_i = 0$	0.79	0.40	0.23	0.02	0.06
$c_i = 20$	0.02	0.15	0.35	0.52	0.75

Table 1 - Percentage of individuals contributing 0 or 20 in periods 9-10

Dependent variable: punishment inflicted from i on j		
	coefficient	standard error
e^2	-0.15**	0.07
e	0.92**	0.39
income	-0.11***	0.03
$\max\{c_i - (1/(n-1))\sum_{j\neq i} c_j,\, 0\}$	-0.02	0.02
$\max\{(1/(n-1))\sum_{j\neq i} c_j - c_i,\, 0\}$	0.37***	0.05
$(1/n)\sum_{i=1}^{n} c_i$	0.07***	0.02
period	-.10***	0.02
constant	1.78***	0.54

Table 2 - Determinants of punishment

Treatment	Group	Mean contribution all periods	Mean contribution periods 6-10	Mean payoff all periods	Mean payoff periods 6-10
0	1	3.7	0.35	22.22	20.21
0	2	2.95	0.5	21.77	20.30
0	3	7.85	3.55	24.71	22.13
0	4	4.38	3.05	22.63	21.83
0	5	2.53	1.55	21.52	20.93
0	6	6.73	2.2	24.04	21.32
0	**average**	**4.69**	**1.87**	**22.81**	**21.12**
1	11	14.45	14	24.72	25.10
1	12	9.2	4.5	19.92	21.50
1	13	3.08	2.3	20.90	20.48
1	14	15.63	15.5	25.63	27.80
1	15	3.88	1.9	22.13	21.04
1	16	2.05	0.85	20.68	20.51
1	**average**	**8.05**	**6.51**	**22.33**	**22.74**
2	21	18.08	19	28.15	29.60
2	22	10.53	11.45	23.99	25.97
2	23	10.53	9.15	19.12	20.84
2	24	7.23	4.9	21.34	21.14
2	25	19.33	20	29.05	32.00
2	26	5.33	3.75	21.77	21.95
2	**average**	**11.83**	**11.38**	**23.90**	**25.25**
3	31	16	18	24.60	28.40
3	32	16.75	18.1	27.95	30.46
3	33	13.73	17.5	22.34	28.10
3	34	19.25	20	28.85	32.00
3	35	17.9	19	26.64	29.20
3	36	11.58	11.7	26.55	26.62
3	**average**	**15.87**	**17.38**	**26.15**	**29.13**
4	41	10.48	12.38	25.54	27.11
4	42	19.40	20.00	29.77	32.00
4	43	18.85	19.25	29.69	29.68
4	44	18.58	18.75	29.52	28.75
4	45	17.63	18.44	26.83	28.25
4	46	17.23	18.75	27.21	28.75
4	**average**	**17.03**	**17.93**	**28.09**	**29.09**

Table A1 – Group data

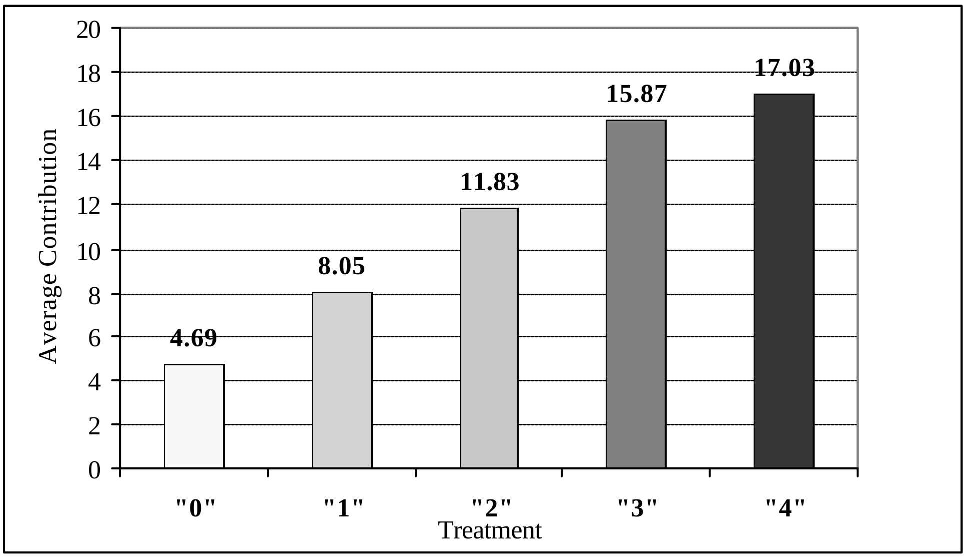

Figure 1 - Average contribution (data from all periods)

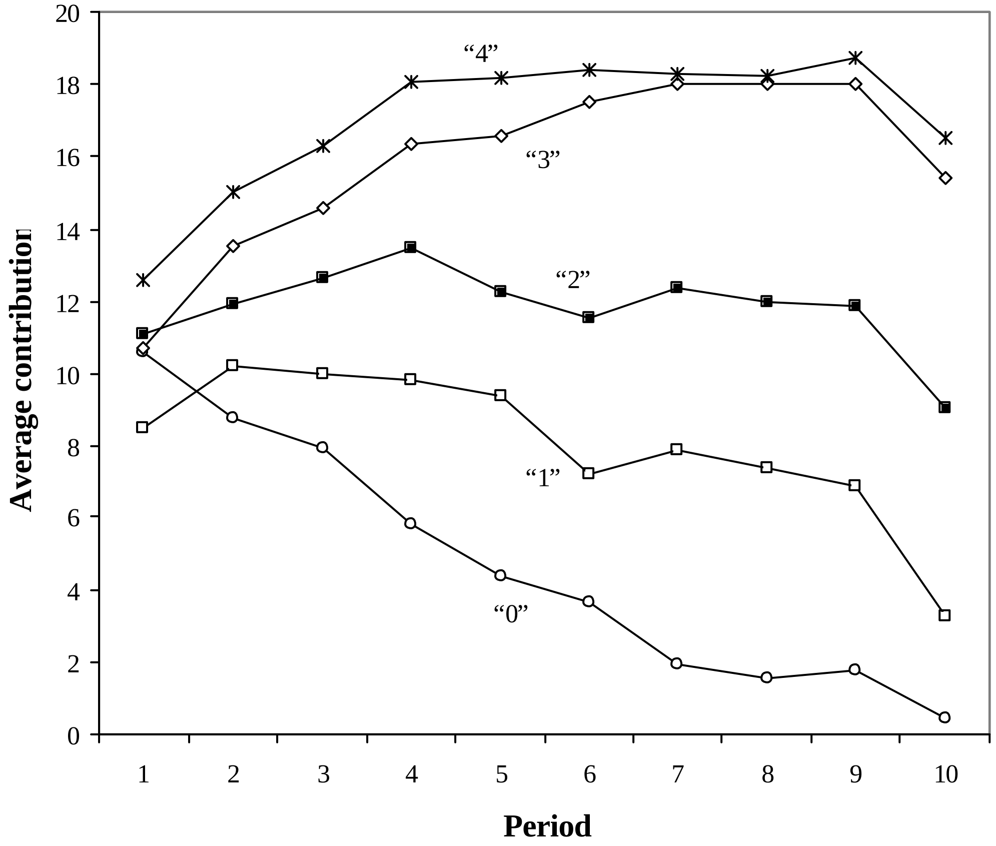

Figure 2 - Average contributions over time

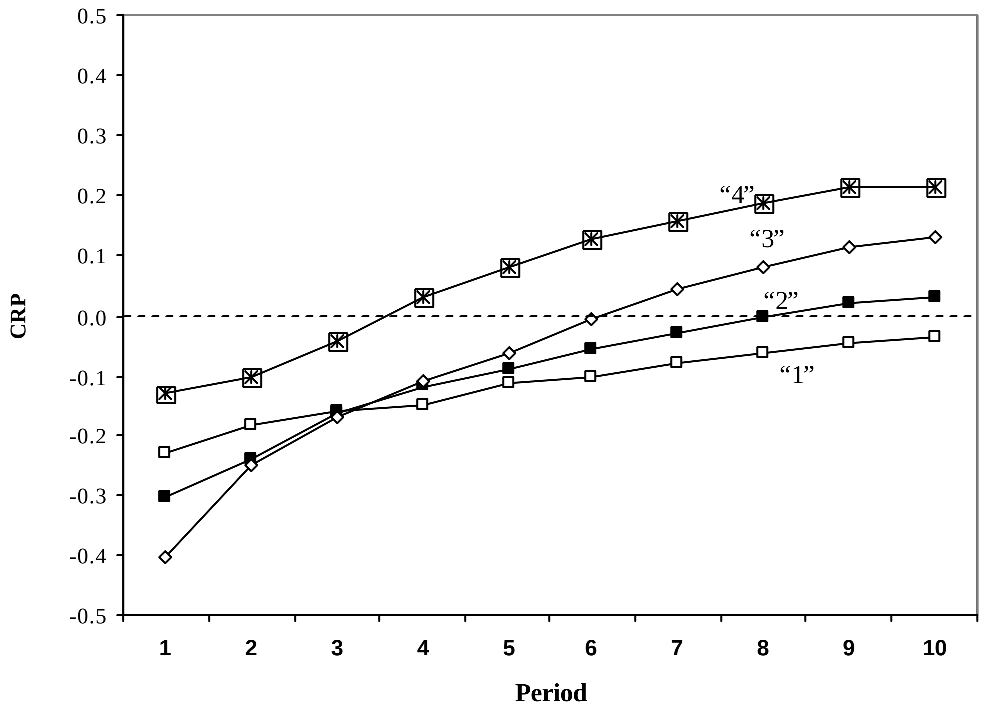

Figure 3 - Relative cumulative payoff over time