

 Open access • Journal Article • DOI:10.4284/SEJ.2010.76.3.638

## **Asymmetric Enforcement of Cooperation in a Social Dilemma** — [Source link](#)

[Nikos Nikiforakis](#), [Hans-Theo Normann](#), [Brian Wallace](#)

**Institutions:** [Max Planck Society](#), [University College London](#)

**Published on:** 01 Jan 2010 - [Southern Economic Journal](#) (Max Planck Institute for Research on Collective Goods)

**Topics:** [Sanctions](#) and [Punishment \(psychology\)](#)

Related papers:

- [Cooperation and Punishment in Public Goods Experiments](#)
- [z-Tree: Zurich toolbox for ready-made economic experiments](#)
- [Altruistic punishment in humans.](#)
- [Punishment and counter-punishment in public good games: Can we really govern ourselves?](#)
- [A Comparative-Statics Analysis of Punishment in Public Good Experiments](#)

Share this paper:    

View more about this paper here: <https://typeset.io/papers/asymmetric-enforcement-of-cooperation-in-a-social-dilemma-8xxl40rpq4>

ISSN 0819-2642  
ISBN 0 7340 2640 4



**THE UNIVERSITY OF MELBOURNE**  
**DEPARTMENT OF ECONOMICS**

RESEARCH PAPER NUMBER 982

JANUARY 2007

**Asymmetric Enforcement of  
Cooperation in a Social Dilemma**

by

Nikos Nikiforakis  
&  
Hans-Theo Normann  
&  
Brian Wallace

Department of Economics  
The University of Melbourne  
Melbourne Victoria 3010  
Australia.

# Asymmetric Enforcement of Cooperation in a Social Dilemma\*

Nikos Nikiforakis,<sup>†</sup> Hans-Theo Normann,<sup>‡</sup> and Brian Wallace<sup>§</sup>

January 2007

## Abstract

The imposition of sanctions is one of the most common means of enforcing cooperation in decentralized interactions. Typically, agents are asymmetric in the sense that each has a different sanctioning power. Using a public-good experiment we analyze such a decentralized punishment institution in which agents are asymmetric. The asymmetric punishment institution prevents the decay of cooperation towards the non-cooperative equilibrium level. Strong agents contribute less to the public good, but punish more than weak agents. At the aggregate level, we observe remarkable similarities between outcomes in asymmetric and symmetric punishment institutions.

JEL Classification: C92, D70, H41

Keywords: asymmetry, decentralized punishment, public good, punishment effectiveness

---

\*The authors gratefully acknowledge funding from ESRC (project RES-000-22-0948). The authors would also like to thank for their comments seminar participants at Royal Holloway, University College London, University of Melbourne, La Trobe University, the American Meeting of the Economic Science Association in Tucson in 2005, the European Meetings of the Economic Science Association in Alessandria in 2005 and Nottingham in 2006.

<sup>†</sup>Department of Economics, The University of Melbourne, Victoria 3010, Australia, tel: +61 383449717, fax: +61 38344 5104, email: *n.nikiforakis@unimelb.edu.au*

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

<sup>§</sup>Department of Economics & ELSE, University College London, Gower Street, London, WC1E 6BT, U.K., email: *brian.wallace@ucl.ac.uk*

# 1 Introduction

With topics such as nuclear proliferation and climate change hotly debated, discussions on how to solve cooperation problems have returned to the center of attention. The two issues have similarities. Firstly, negotiations between nations are, in essence, decentralized. That is, a central authority to enforce law and order does not exist. Secondly, negative externalities are present in both cases. Refusal to decrease the emission of greenhouse gasses is often believed to impose a cost to the whole world by increasing, for example, the number of extreme weather phenomena. Similarly, developing nuclear weapons increases insecurity to other nations and diminishes their relative bargaining power. Generally, “when externalities are present, private self-interest and overall group welfare may be at odds” (Noussair and Tucker, 2005, p. 649) and we have a social dilemma.

A common way of dealing with negative externalities is the imposition of sanctions (punishments) on parties that deviate from a widely accepted norm of behavior. A recent example is the announcement of the United States’ intention to impose financial and military sanctions to North Korea for conducting its first nuclear test. One purpose of sanctions such as these is to lower the return from acting against the social interest. This implies that the efficacy of sanctions in enforcing cooperation depends critically on the sanctioning power of the sanctioning party. The greater the power of the party abiding to a given norm, the less appealing a deviation from that norm will be for other parties.

The purpose of this paper is to study the efficacy of decentralized sanctions in fostering cooperation when players are asymmetric in their sanctioning power and negative externalities exist. To our knowledge this is the first study on the topic. We investigate two contradicting hypotheses about the impact of asymmetric punishment institutions. On the one hand, Olson (1965) and Axelrod (1997) propose that a “dominant power” will have a positive effect on cooperation. A justification for this hypothesis is that punishment of non-cooperators is a public good. All group members benefit from the resulting increase of cooperation due to punishment, but all have an incentive to let others carry out the punishment cost. By allocating punishment power to one strong player, the free-rider problem at the punishment stage may be alleviated. On the other hand, several contributions to oligopoly theory (discussed below) argue that cooperation between asymmetric agents is less likely. The reason is that the strongest of the group faces less of a threat of being punished when failing to cooperate. Therefore, the strong player is more likely to defect, jeopardizing future cooperation of the whole group.

As a starting point for our analysis we use the voluntary contribution mechanism (VCM) - a

game that economists have repeatedly used to analyze the tension between the social good and self interest. In the VCM, individuals have to make a decision concerning how much they wish to contribute to a public account. The higher the contributions to the public account, the higher the group payoff. However, every individual has also an incentive to free-ride and not contribute. The VCM captures, in a simple manner, the conflict between social and self interest due to an externality present in many situations such as the ones described in the first paragraph and, thus, is suitable for the purpose of this paper.

To investigate the impact of asymmetric punishment institutions on cooperation, we study treatments with asymmetric players (where players differ in their ability to sanction each other) and compare the results to treatments with symmetric players (where all players have equal sanctioning ability). In particular, we investigate four asymmetric punishment institutions and compare them to two symmetric institutions, controlling for the average sanctioning power.

The results indicate that asymmetric institutions can be very successful in fostering cooperation, and are as effective and efficient as symmetric institutions. Strong players contribute less to the public account in comparison to their weak counterparts, but are more likely to punish free-riders. Strong players have higher payoffs than weak players.

Our work builds on previous research that examines the *symmetric* enforcement of cooperation in social dilemmas (e.g. Ostrom et al., 1992; Fehr and Gaechter, 2000; 2002; Masclet et al., 2003; Noussair and Tucker, 2005; Sefton et al., 2005). The evidence indicates that, under certain conditions—such as the absence of retaliation opportunities (Denant-Boemont et al., 2005; Nikiforakis, 2005) and the existence of an adequate punishment threat (Nikiforakis and Normann, 2005)—symmetric players can sustain cooperation through the use of decentralized punishment. The present experiment examines asymmetric enforcement of cooperation. We believe the issue to be of interest as, in most real-world scenarios, symmetric players are the exception rather than the norm.

The rest of the paper is organized as follows. Section 2 presents the experimental design and the procedures. Section 3 contains hypotheses for our treatments and Section 4 illustrates the experimental results. Section 5 concludes.

## 2 The Experiment

The experiment is based on the design of Fehr and Gächter (2002) and Noussair and Tucker (2005) who use the voluntary contribution mechanism with  $n$  players. In every period, all participants

are 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 player  $i$  keeps,  $i$  receives a fixed percentage of the group's total contribution to the public account,  $\alpha$ , where  $0 < \alpha < 1 < n\alpha$ . This implies that the payoff of player  $i$  in the first stage of some period is

$$\pi_i^1 = y - c_i + \alpha \sum_{i=1}^n c_i. \quad (1)$$

In the second stage, 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 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  $i, j=1, \dots, n; j \neq i$ ), and  $e_j$  the income reduction that one punishment point from player  $i$  causes to its recipient, that is,  $j$ 's *punishment effectiveness* (we use this term interchangeably with sanctioning power). Player  $i$ 's payoff at the end of a period is accordingly

$$\pi_i^2 = y - c_i + \alpha \sum_{i=1}^n c_i - \sum_{j \neq i} p_{ij} - \sum_{j \neq i} e_j p_{ji}. \quad (2)$$

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  $\alpha=0.4$ .

Table 1 describes the treatments in the experiment. The treatment labels read " $e_s\_e_w$ ", such that the number on the left indicates the effectiveness of the strong player,  $e_s$ , and the number on the right the effectiveness of the weak players,  $e_w$ . So, for example, in treatment "5\_1", one punishment point from the strong player reduces the income of its recipient by 5 ECU, while one punishment point from the weak players reduces the income of its recipient by 1 ECU. In the asymmetric treatments, there was always one "strong" and three "weak" players.

The treatments differ in two dimensions: First, the *asymmetry level* indicates the relative strength of the strong player's punishment and is denoted by  $l \equiv e_s/e_w$ . Second, the *average effectiveness*,  $\bar{e}$ , is the average punishment effectiveness of the group members,  $\bar{e} \equiv (e_s + 3e_w)/4$ . We ran treatments with  $\bar{e} = 2$  and  $\bar{e} = 3$ . Based on the findings of Nikiforakis and Normann (2005) we anticipated that the higher the level of punishment effectiveness the higher the level

of cooperation. For both  $\bar{e} = 2$  and  $\bar{e} = 3$ , we conducted sessions with  $l = 3$  in addition to the symmetric control sessions with  $l = 1$ . We also ran a treatment with  $\bar{e} = 2$  and  $l = 5$  (“5\_1”). However, we were concerned that a treatment where  $\bar{e} = 3$  and  $l = 5$  (“7.5\_1.5”) would be risking losses for the weak players due to the magnitude of the strong players’ punishment effectiveness in this case. This could have caused frustration and have led to erratic behavior. Consequently, we decided to conduct treatment “4\_2.6” with  $\bar{e} = 3$  and  $l = 1.5$ .<sup>1</sup>

*insert Table 1 about here*

Information feedback is as follows. At the beginning of each experimental session, subjects are informed whether they are assigned the role of a strong or a weak player.<sup>2</sup> These roles remained fixed for the duration of the experiment. At the beginning of each period every player is randomly given a number between 1 and 4 to distinguish their actions from those of the others in that period. To prevent the formation of individual reputation, the numbers are randomly reallocated in the beginning of every period. Participants are aware of this. Such a mechanism ensures that, even though the group members remain the same, the participants cannot link the actions of the other subjects across the periods. Once the participants have contributed at stage one, they are informed about their group’s total contribution to the public account, individual contributions and their payoff from the period as given by equation (1). 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 of the other players. They only know how many points they assigned to the other group members, thus, retaliation as in Denant-Boemont et al. (2005) and Nikiforakis (2005) is not possible.

All treatments last for 10 periods. For the experiment we use fixed (or “partners”) matching. This implies that every group can be regarded as a statistically independent observation. For treatments “2\_2”, “3\_3” and “5\_1” we have six groups and, thus, six independent observations. For

<sup>1</sup>We did not use treatments with  $\bar{e} = 1$  instead of  $\bar{e} = 3$  as this would require  $e_w < 1$  in the asymmetric treatments. If  $e_w < 1$  a punishment carried out by the weak players increases inequality between the victim and the punisher. This would make comparison with treatments where  $\bar{e} = 2$  harder as in these cases punishment either reduces inequality or leaves inequality unaffected.

<sup>2</sup>In the instructions, we used neutral language. The strong agents were “type A” and the weak agents were “type B”. Punishment was termed as “points that reduce another player’s income”. The instructions are available from the authors upon request.

treatments “4\_1.3”, “4\_2.6” and “6\_2” we have five groups. In two cases, we missed the target of six groups due to individuals not showing up. In the third case (“4\_1.3”), we had to discard one group from the analysis due to a bankruptcy problem.<sup>3</sup>

The experiments were conducted at the University of London (Royal Holloway College and University College). The total number of participants was 132 (not counting the discarded group).<sup>4</sup> The subjects were recruited using an e-mail list of voluntary potential (student) 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. The rate of exchange between the experimental currency unit and the British pound was 1 ECU = £0.04. The average earning in the experiment was £10.61 or roughly \$20. The experiments were conducted using z-Tree (Fischbacher, 1999).

### 3 Hypotheses

In this section, we propose informal conjectures about the effect of an asymmetric punishment institution on cooperation. The only rigorous game-theoretic prediction is the subgame perfect Nash equilibrium. In this equilibrium punishments and, therefore, asymmetries do not matter. The reason is that our experiment is finitely repeated and, by backward induction, the prediction for all periods of all treatments is that players do not punish and do not contribute to the public good. On the other hand, we know from previous research that institutions allowing for decentralized sanctions can increase cooperation levels (e.g. Fehr and Gächter, 2000; 2002; Maslet et al., 2003; Noussair and Tucker, 2005). Asymmetries may also matter and we wish to suggest some conjectures about their impact.

Olson (1965) and Axelrod (1997) argue that the presence of a strong player is likely to improve cooperation. The reason is that the punishments at the second stage of the game (aimed at improving cooperation) are a public good itself. Accordingly, free riding by not punishing is the short-run payoff-maximizing behavior. Let us compare treatments with symmetric punishment

---

<sup>3</sup>The group had one subject contributing nothing for several periods and the other three subjects, who contributed the maximum amount, punished her harshly, causing the bankruptcy of the subject in period 5. In order for the experiment to continue we credited her account with £5. Whereas we did continue collecting data beyond that period, we decided not to include the group in the data analysis. In the first five periods, punishment in this group was six times higher than the treatment average which would cause a severe bias if we included this group in our statistical analysis. Moreover, a post-experimental questionnaire indicated that the behavior of the punished person was due to a misunderstanding.

<sup>4</sup>The observations for symmetric treatments “2\_2” and “3\_3” are taken from Nikiforakis and Normann (2005).



effectiveness to asymmetric treatments. Making one player more effective in punishing and the other three players less effective might help overcome the free-rider problem at the second stage. The strong player's punishment is important to enforce cooperation as the weak players might be unable to enforce cooperation by themselves. Knowing this the strong player is more likely to punish. Therefore, punishment at the second stage is more likely to occur and thus cooperation at the first stage will improve.

**Conjecture 1:** *An asymmetric punishment institution leads to higher contribution levels compared to a symmetric institution.*

Some contributions to oligopoly theory suggest just the opposite, that asymmetries are an obstacle to cooperation. Davidson and Deneckere (1984, 1990), Lambson (1995) and Compte et al. (2002) show that introducing asymmetries in the cost parameters or the capacities of oligopolistic firms typically reduces the scope for successful collusion. The intuition is that making one firm bigger or more efficient implies that this firm faces a smaller punishment following a deviation from cooperation. Hence, the big firm is less likely to cooperate and, in turn, the industry is generally less likely to cooperate. Similarly, the strong players in our game are the most difficult to discipline. Therefore, controlling for  $\bar{e}$ , the group as a whole will find cooperation more difficult compared to the symmetric setup.

**Conjecture 2:** *An asymmetric punishment institution leads to lower contribution levels compared to a symmetric institution.*

Obviously, the two conjectures are mutually exclusive. However, it should be apparent why the two above arguments lead to opposite conclusions. The argument that a strong player can help overcome the free-rider problem at the punishment stage is based on the presumption that the strong player is inclined to enforce cooperation. The reasoning based on the reduced threat the strong player faces assumes, by contrast, that the strong player is a selfish player in the sense that she will not contribute to the public good—unless forced to do so by punishment. The contrasting conjectures suggest that the effects of an asymmetric punishment institution depend on the preferences of the strong player. A strong player inclined to support cooperation in the group may have the punishment power to do so, even if the weak players are initially zero contributors. But the opposite might not be possible. If the strong player has no inclination to establish cooperation, the weak players may not have the punishment power to change the strong player's behavior even if they want to.

That cooperation levels depend on the preferences of the players involved has been proposed by Fehr and Schmidt (1999, section IV). Their argument is based on the existence of “conditionally cooperative enforcers” who may be able to force selfish players into contributing. As for the impact of asymmetries, if chance assigns the role of the strong player to a conditionally cooperative enforcer, the group should cooperate better than in the symmetric treatment. If the strong player is selfish, the opposite should occur. However, whether enforcement of cooperation is more effective in an asymmetric punishment institution than in a symmetric institution remains an open question. The answer to this question depends on intricate details (average punishment effectiveness, distribution of punishment effectiveness, frequency of the conditionally cooperative enforcers) and neither institution generally dominates the other.

**Conjecture 3:** *Whether or not an asymmetric punishment institution improves cooperation depends on the preferences of the strong player.*

## 4 Results

We begin the analysis by taking an overview of the data. We then proceed with a rigorous analysis of the impact of the asymmetric punishment institution on contributions to the public account, the punishment behavior, and on welfare.

### 4.1 Overview

Table 2 presents the summary statistics of our results. The table reports average contributions, average income before and after the punishment stage, and details of the punishment stage. The first thing to note is that in all treatments the punishment institution can sustain cooperation at higher levels than the ones predicted by the subgame perfect Nash equilibrium. The asymmetric institutions appear to be as effective and efficient as the symmetric institution on the aggregate level. Strong players tend to spend more money on punishment than weak players.<sup>5</sup> Treatments with  $\bar{e} = 3$  appear to have higher contributions to the public good than treatments with  $\bar{e} = 2$ .

*insert Table 2 about here*

---

<sup>5</sup>In treatment “4\_2.6” the average number of punishment points assigned by weak agents is slightly higher than of that of the strong agents. This is the product of a single group in which the weak agents punished severely.

## 4.2 Contributions to the public account

The impact of average effectiveness ( $\bar{e}$ ) and asymmetry level ( $l$ ) on contribution rates can be seen more clearly in Figure 1. Contributions in treatments with  $\bar{e} = 3$  are at a high level and, overall, appear to be increasing over time. On the other hand, contributions in treatments with  $\bar{e} = 2$  are at a lower level and remain more or less constant or even decline slightly after period five. The striking fact is that the evolution of contributions is very similar amongst the three treatments with  $\bar{e} = 2$  (the lower three lines) and also amongst the three treatments with  $\bar{e} = 3$  (the upper three lines).

*insert Figure 1 about here*

We now turn to the statistical analysis of the data. The data is a panel as we have repeated observations from the same individuals. Decisions are likely to be strongly correlated within groups. As a result, we include random effects at the group level. As implied by Figure 1, there is a concentration of data at 20 ECU (recall that 20 is the maximum contribution). This implies that we have to use a Tobit model for restricted dependent variables. The independent variables are *three*, a dummy variable which takes the value of one for all treatments with  $\bar{e} = 3$  and zero otherwise; *asymmetry\_level* which is  $l \equiv e_s/e_w$ <sup>6</sup>; and *strong*, a dummy variable which takes the value of one when individual  $i$  is a strong type in an asymmetric treatment and zero otherwise.<sup>7</sup>

*insert Table 3 about here*

Table 3 shows the results of a Tobit regression investigating the effect of the explanatory variables on contributions. As implied by Figure 1, the level of asymmetry does not have a significant impact on contributions, while as average effectiveness increases from 2 to 3 so do contributions on

---

<sup>6</sup>We ran various similar regressions where asymmetry (or different levels of asymmetry) was captured by a dummy variable. The results turned out to be similar.

<sup>7</sup>We ran alternative regressions in which the *strong* dummy is equal to one for strong players in the asymmetric treatments and for all players in the symmetric treatments. Results from these regressions did not differ qualitatively from the ones reported here.

average. The main result from Table 3 is that strong players contribute significantly less than weak players. Adding a time-trend variable in these regressions does not change the results qualitatively.<sup>8</sup>

**Result 1:** *Asymmetric punishment institutions are as successful as symmetric institutions in fostering cooperation.*

**Result 2:** *The average effectiveness of punishment has a significant impact on contributions.*

**Result 3:** *Strong players contribute less to the public good than weak players.*

### 4.3 Punishment behavior

Next, we analyze punishment behavior. The most interesting question is whether strong players punish differently than weak players. According to Olson (1965) and Axelrod (1997), strong players should be punishing more. Figure 2 gives a first answer about whether this is the case in the experiment. On the horizontal axis we plot the deviation of individual  $j$ 's contribution from that of his peers. On the vertical axis we plot the likelihood of  $j$  being punished by strong and weak players. For all levels of deviation, strong players are found to be more likely to punish. The difference exceeds 20 percent in some cases.

*insert Figure 2 about here*

For a regression analysis we need to take into consideration that, of the 3960 possible punishment cases (132 participants times 3 targets per period times 10 periods), punishment was observed in only 613 cases. This implies that the modal behavior in the second stage is to avoid punishing. The appropriate econometric specification to capture the two-stage process is a hurdle-model. The hurdle-model is a parametric generalization of the Tobit model in which the decision to punish and the extent of punishment are determined by two separate stochastic processes.<sup>9</sup> The hurdle is crossed if an individual decides to punish.

The likelihood function for the hurdle model is given by the product of two separate likelihoods. First, the likelihood that a subject punished another group member or not, which is captured by

---

<sup>8</sup>Non-parametric tests yield similar results for the effect of  $\bar{e}$  and  $l$  to the ones reported in this section. However, to study the impact of strong players we need to use parametric tests to isolate the effect of increasing  $\bar{e}$  and  $l$ .

<sup>9</sup>For an example of using hurdle-models to analyze decision in public good games see Botelho et al. (2005).

a Probit model, and second, the conditional likelihood that an individual punished with a certain number of points, which is captured by a truncated at zero Tobit model. The two parts of the hurdle-model are estimated separately (McDowell, 2003).

Table 4 presents the estimates of the hurdle-model. The dependent variable in the punishment decision is a dummy taking the value of one if subject  $i$  punished subject  $j$  and zero otherwise. The dependent variable in the punishment-level decision is the number of points that  $i$  assigned to  $j$ . Figure 2 indicates that punishment might be increasing in the negative deviation of  $j$  from his peers. Moreover the reaction seems to be different if  $j$  contributed more or less than his peers, with much more of a reaction to negative deviations. To capture this effect, in addition to our treatment regressors, we include separate *positive [negative] deviation of  $j$*  which measure how much more [less]  $j$  contributed in comparison to the average of the other three group members in period  $t$ . We also add *period* to capture the time trend as punishments appear to decline over time.

*insert Table 4 about here*

The results from the regression lend support to the observations made when looking at Figure 2. Strong players are more likely to punish than their weak counterparts. This is surprising, given that strong players were found to contribute less to the public account. We will return to this issue later in section 4.5. In any event, the tendency of strong players to punish more frequently might explain why, even though strong players contribute less to the public good, they still do not cause the breakdown of cooperation.

In Table 4 we also see that, given a subject decides to sanction another group member, strong and weak players assign the same number of points. This is surprising because, in order to achieve the same level of income reduction of the punish player, weak players need to assign more points than strong players. We conjecture that weak players generally rely on the strong players to enforce cooperation. They are less likely to punish in the first place and, if they punish, the punishment points they assign may well be insufficient to foster cooperation without the punishments from the strong players.

**Result 4:** *Strong players are more likely to punish than weak players. Given a player decides to punish, the strong and weak players assign similar number of points.*

As one would expect, group members contributing more than the average of their peers are significantly less likely to be punished. On the other hand, group members contributing less than the average of their peers are more likely to be punished. In addition, the less they contribute compared to their peers the higher their punishment is. The likelihood of a punishment diminishes over time, but not the punishment intensity.

An interesting finding is that, given participants decided to punish, the level of asymmetry lowers significantly the number of points assigned to an individual. This suggests that the asymmetry helps solve the coordination problem that exists in the punishment stage. In the symmetric treatments individuals observing a free-rider are uncertain whether others will punish her. As a result, more than one group member might punish a defector. In the presence of asymmetries the focal player is the strong player, and the strong player can achieve the income reduction more efficiently. The strong player knows that he has a comparative advantage in punishing and that his weak counter-parts also know this.

**Result 5:** *The higher the level of asymmetry the lower the punishment points assigned.*

Result 5 might be good news for social welfare. If the existence of a focal player reduces punishment points, then the group saves on punishment costs. However, for welfare to benefit from the existence of asymmetries, this effect must outweigh the fact that the punishment carried out by the strong players is more damaging than that of the weak players. This is what we study next.

#### 4.4 Welfare

As mentioned above, the subgame perfect Nash equilibrium predicts that there will be no punishments and no contributions in the experiment. From equation (2), this implies a period payoff of 20 ECU per person. On the other hand, if each member contributes the whole endowment to the public account and abstains from punishments each individual will have a payoff of 32 ECU. These are the benchmarks against which we measure the performance of the asymmetric punishment institution.

Table 2 illustrates how payoffs are shaped in each of the six treatments. As indicated by equation (1), the higher contributions are (column 2) the higher the average income before punishment will be (column 3). Column 7 reports the average income that was lost due to punishment activities and includes the cost to both the punisher and the target. The average payoff (column 8) is found by subtracting column 7 from column 3.

Figure 3 complements Table 2 by illustrating the evolution of average payoffs for each of the treatments separately. The following become apparent. First, payoffs are somewhere between the Pareto optimal payoff of 32 ECU and the one predicted by the subgame perfect Nash equilibrium of 20 ECU. Second, there appear to be no significant differences between treatments. This implies that the positive effect of asymmetries discussed above is offset by the heavier use of punishment by the strong players.

*insert Figure 3 about here*

In columns 9 and 10 of Table 2 we also find the average income of strong and weak players. We see that with the exception of treatment “4\_1.3” strong players tend to have a higher income. A comparison of columns 5 and 6 indicates the greater expenditure on punishment by strong players indicated earlier.

Table 5 presents the results of a random-effects regression where the dependent variable is the payoff of individual  $i$  at the end of a period,  $\pi_i^2$ . We find that neither the level of average effectiveness nor the level of asymmetry have a significant effect on welfare. However, we also find that strong players have higher payoffs than weak players. Average payoffs increase over time due to increasing contributions (in treatments with  $\bar{e} = 3$ ) and falling expenditure on punishment (in all treatments).

**Result 6:** *Group welfare is not affected by average effectiveness or the level of asymmetry.*

**Result 7:** *Strong players have significantly higher payoffs than weak players.*

*insert Table 5 about here*

## 4.5 Discussion

We saw in Result 1 that cooperation levels do not differ significantly between asymmetric and symmetric punishment institutions. Hence, this rejects both Conjecture 1 and Conjecture 2. Neither do we observe improved cooperation rates nor do we see the breakdown of cooperation with

the asymmetric institution. Result 1 is consistent with Conjecture 3, saying that the effect of asymmetries should depend on the attitude of the strong player in the group. If “cooperative” and “non-cooperative” strong players are balanced in the population, the effect of asymmetries might cancel out—which is just what we seem to observe.

However, underlying Conjecture 3 was the notion that players are either (conditionally) cooperative enforcers who both contribute and punish, or selfish players who do neither. This is inconsistent with the results that strong players, on average, contribute less and punish more. If we make a clear distinction between “cooperative” and “non-cooperative” types of players, players (including the strong ones) should either contribute and punish if they are cooperative, or not contribute and refrain from punishing if they are non-cooperative. What players should not do, according to this view, is reduce cooperation at the contribution stage but then punish in order to encourage contributions by others.

Nevertheless, contributing less and punishing more is the modal behavior of the strong players. Table 6 categorizes strong players according to whether they contribute more or less than their peers’ average and whether they punished more or less than their peers on average.<sup>10</sup> Out of 21 strong players, there are 4 players who contribute more and punish more, and there are also 4 players who contribute less and punish less. They are consistent with the categorization of purely “cooperative” and “non-cooperative” types of players. But the majority of players, 10, contribute less and punish more. In the opposite category, contribute more and punish less, only one player can be found. The remaining two subjects are exactly equal to the average in at least one dimension.

We conduct two simple non-parametric tests regarding the data in Table 6. First, the hypothesis that the behavior of the strong players does not differ from that of the weak players is rejected. If strong and weak players behaved alike, the entries in Table 6 should be equally distributed across the four cells. However, a (one sample) chi-square test indicates significant differences ( $\chi^2 = 9.00$ ,  $d.f.=3$ ,  $p=0.029$ ). Second, if we had mainly “cooperative” players who contribute and punish more, and “non-cooperative” players who contribute and punish less, the upper left and the lower right cell should be the two modes. If so, there should be significant differences between the proportions of Table 6. This is, however, not the case as the  $p$ -value for the same or a stronger association of Fisher’s exact test is  $p = 0.603$ .

*insert Table 6 about here*

---

<sup>10</sup>We measure punishment behavior here by the points given to other players, that is, the income the players sacrificed.



The behavioral mode we observe is surprising as lower than average contributions will discourage other players from cooperation while harsher punishments will encourage cooperation. One interpretation is that, understanding their comparative advantage in punishing, strong players simply “save” in the first stage to contribute to the second-order public good, that is, the punishment of free-riders. In that sense, there are two public goods and players seem to treat the two public goods as substitutes. Apparently, strong players shirk (at least partially) at the contribution level but contribute more at punishing. The strong players might “justify” the lower contributions to the public good because they are more active at contributing at the punishment stage. Of course, they can do so only because they do not face much of a threat from weak players. An interesting extension would, therefore, be to see what happens when there are two strong players. Given that the strong players do not cooperate perfectly at both levels, the pattern we observe is efficient. The strong players have a comparative advantage in punishing and, in a way, this is what they “specialize” in.

## 5 Conclusion

In this paper we present the results from an experiment investigating the efficacy of an asymmetric punishment institution in enforcing cooperation in the presence of externalities. We find that the asymmetric enforcement of cooperation is as effective at sustaining cooperation and as efficient as the symmetric institutions at the aggregate level. Asymmetries neither foster cooperation nor do they constitute an obstacle to it. At the individual level, strong players benefit from the asymmetry. They contribute less to the public account and punish more than their weak counterparts. This results to higher payoffs for the strong players.

Our experiment is motivated by the imposition of real sanctions by players asymmetric in their sanctioning power. What can our results, therefore, say about decentralized enforcement of cooperation by asymmetric players? Keeping in mind that one needs to be careful generalizing from experiments where, similar to theoretical models, a number of forces present in the real world have been isolated, a number of points can be made. First, it seems that the existence of asymmetries might limit the punishment expenditure by improving coordination. While this effect alone improves welfare, the higher damage inflicted by the sanctions of strong players can reduce welfare. Second, weak players might not be able to force a non-cooperative strong player to cooperate and, therefore, the need arises to control the strong player. Our data indicates that strong players might be willing to enforce cooperation knowing that they also benefit from it.

Whether the strong players in our experiments had truly cooperative preferences or whether they adopted the role of the enforcing “policeman” (realizing that cooperation depended on them) is a question to be answered in future research. Third, the payoff difference between strong and weak players could indicate that, given the option, weak players will have an incentive to invest a part of their endowment in enhancing their punishment effectiveness, and such incentives may give rise to an “arms race” for punishment effectiveness.

## References

- [1] Axelrod R., (1997). *The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration*, Princeton: Princeton University Press.
- [2] Botelho, A., Harrison, G. W., Pinto, L. M. C. and Rutstrom, E. E., (2005). *Social Norms and Social Choice*, Working Paper 05-23, Department of Economics, College of Business Administration, University of Central Florida.
- [3] Compte, O., Jenny, F., Rey, P. (2002). Capacity constraint, merger and collusion, *European Economic Review*, 46:1-29.
- [4] Davidson, D., and Deneckere, R. (1984). Horizontal Mergers and Collusive Behavior, *International Journal of Industrial Organization*, 2 (2), 117-132
- [5] Davidson, D., and Deneckere, R. (1990). Excess Capacity and Collusion, *International Economic Review*, 31 (3), 521-541.
- [6] Denant-Boemont, L., Masclet, D., Noussair, C. (2006). Punishment, Counterpunishment and Sanction Enforcement in a Social Dilemma Experiment, HALSHS 00009664 (1).
- [7] Fehr, E. and Gächter, S. (2002). Altruistic Punishment in Humans, *Nature* 415, 137-140.
- [8] Fehr, E. and Gächter, S. (2000). Cooperation and Punishment in Public Goods Experiments, *American Economic Review* 90, 980-994.
- [9] Fehr, E. and Schmidt, K. (1999). A Theory of Fairness, Competition and Co-operation, *Quarterly Journal of Economics*, 114, 817-868.
- [10] Fischbacher, U. (1999). *z-Tree: Zurich Toolbox for Readymade Economic Experiments - Experimenter’s Manual*, mimeo, Institute for Empirical Research in Economics, University of Zurich.

- [11] Lambson, V.E. (1995). Optimal penal codes in nearly symmetric Bertrand supergames with capacity constraints, *Journal of Mathematical Economics*, Volume 24 (1), 1-22.
- [12] 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.
- [13] McDowell, A. (2003). From the Help Desk: Hurdle Models, *The Stata Journal* 3 (2), 178-184.
- [14] Nikiforakis, N. (2005). Punishment and Counter-Punishment in Public Good Games-Can We Really Govern Ourselves?, SSRN (<http://ssrn.com/abstract=764185>).
- [15] Nikiforakis, N. and Normann H.T. (2005). A Comparative Analysis of Punishment in Public-Good Experiments, *Royal Holloway Discussion Paper Series* 05/7.
- [16] Noussair, C., and Tucker, S. (2005). "Combining Monetary and Social Sanctions to Promote Cooperation", *Economic Inquiry*, 43(3),649-660.
- [17] Olson, M. (1965). *The logic of collective action: Public goods and the theory of groups*, Harvard University Press, Cambridge, MA.
- [18] 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.
- [19] Sefton, M., Shupp, R. and Walker J. (2005). The Effect of Rewards and Sanctions in Provision of Public Goods, Ball State University, Department of Economics, Working Paper 200504.

Treatment	Average Effectiveness <sup>1</sup> ( $\bar{e}$ )	Effectiveness <sup>2</sup> of Strong ( $e_s$ )	Effectiveness <sup>2</sup> of Weak ( $e_w$ )	Asymmetry Level <sup>3</sup> ( $l$ )	No. of participants
"2_2"	2	2	2	1	24
"4_1.3"	2	4	1.3	3	20
"5_1"	2	5	1	5	24
"3_3"	3	3	3	1	24
"4_2.6"	3	4	2.6	1.5	20
"6_2"	3	6	2	3	20

<sup>1</sup> "Average Effectiveness is the average punishment effectiveness of the group members, i.e.  $(e_s + 3e_w)/4$ .

<sup>2</sup> "Effectiveness" refers to the income reduction in ECU caused to the recipient by a single punishment point.

<sup>3</sup> "Asymmetry Level" is defined as  $e_s/e_w$ .

Table 1: Experimental Design

Treatment	Average Contribution	Average Income Before Punishment	Average Points Given	Average Points Given by Strong	Average Points Given by Weak	Average Income Lost*	Average Payoff	Average Payoff Strong	Average Payoff Weak
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
"2_2"	11.83	27.10	1.07	-	-	3.20	23.90	-	-
"4_1.3"	12.36	27.42	1.01	1.12	0.97	3.08	24.34	23.24	24.69
"5_1"	12.49	27.49	0.47	0.55	0.44	1.49	26.00	28.40	25.20
"3_3"	15.87	29.52	0.84	-	-	3.37	26.15	-	-
"4_2.6"	15.92	29.55	0.68	0.60	0.70	2.64	26.91	27.85	26.56
"6_2"	15.17	29.10	0.91	1.40	0.74	4.12	24.99	26.17	24.59

\* "Average Income Lost" is the sum of the average punishment cost paid by the punishers and the average income reduction suffered by punishment recipients.

Table 2: Summary Statistics

Independent variables	Dependent variable: contribution ( $c_i$ )
Constant	11.520*** (0.680)
Asymmetry Level	0.194 (0.201)
Strong Player (dummy)	-1.369** (0.587)
Three (dummy)	3.144*** (0.560)
	N = 1320
	Wald $\chi^2 = 38.02^{**}$

*Notes:* Standard errors in parenthesis. \* denotes significance at the 10% level, \*\* at the 5% level, and \*\*\* at the 1% level. Regression is a random effects (at group level) Tobit.

Table 3 – Determinants of Contributions

	<b>Punishment Decision</b>	<b>Punishment-Level Decision</b>
Independent variables	Dependent variable: probability of being punished	Dependent variable: punishment points received (conditional on being punished)
Constant	-0.959*** (0.227)	2.01*** (0.369)
Asymmetry Level	-0.061 (0.062)	-0.216** (0.102)
Strong Player (dummy)	0.222** (0.080)	0.018 (0.165)
Three (dummy)	0.064 (0.182)	-0.492* (0.296)
Positive deviation	-0.052*** (0.013)	0.046 (0.032)
(Absolute) Negative deviation	0.162*** (0.009)	0.100*** (0.013)
Period	-0.075*** (0.010)	0.018 (0.021)
	N = 3960	N = 613
	Wald $\chi^2 = 546.54^{***}$	Wald $\chi^2 = 66.85^{***}$

*Notes:* Standard errors in parenthesis. \* denotes significance at the 10% level, \*\* at the 5% level, and \*\*\* at the 1% level. Punishment Decision regression is a random effects (at group level) Probit. Punishment-Level Decision regression is a random effects (at a group level) linear regression with observation truncated at zero.

Table 4 – Determinants of Punishment

Independent variables	Dependent variable: profit ( $\pi_i^2$ )
Constant	23.966*** (1.702)
Asymmetry Level	0.200 (0.482)
Strong Player (dummy)	1.247*** (0.422)
Three (dummy)	1.499 (1.420)
	N = 1320
	Wald $\chi^2 = 10.03^{**}$

*Notes:* Standard errors in parenthesis. \* denotes significance at the 10% level, \*\* at the 5% level, and \*\*\* at the 1% level. Regression is a random effects (at group level) linear regression.

Table 5 – Determinants of Welfare



		<b>Punish</b>	
		More	Less
<b>Contribute</b>	More	4	1
	Less	10	4

*Note:* The total number of strong players is 21. The actions of two strong players are equal to the average of the weak players in at least one dimension.

Table 6 – Categorization of strong players according to whether they contributed/punished more or less compared to the average of the weak players.

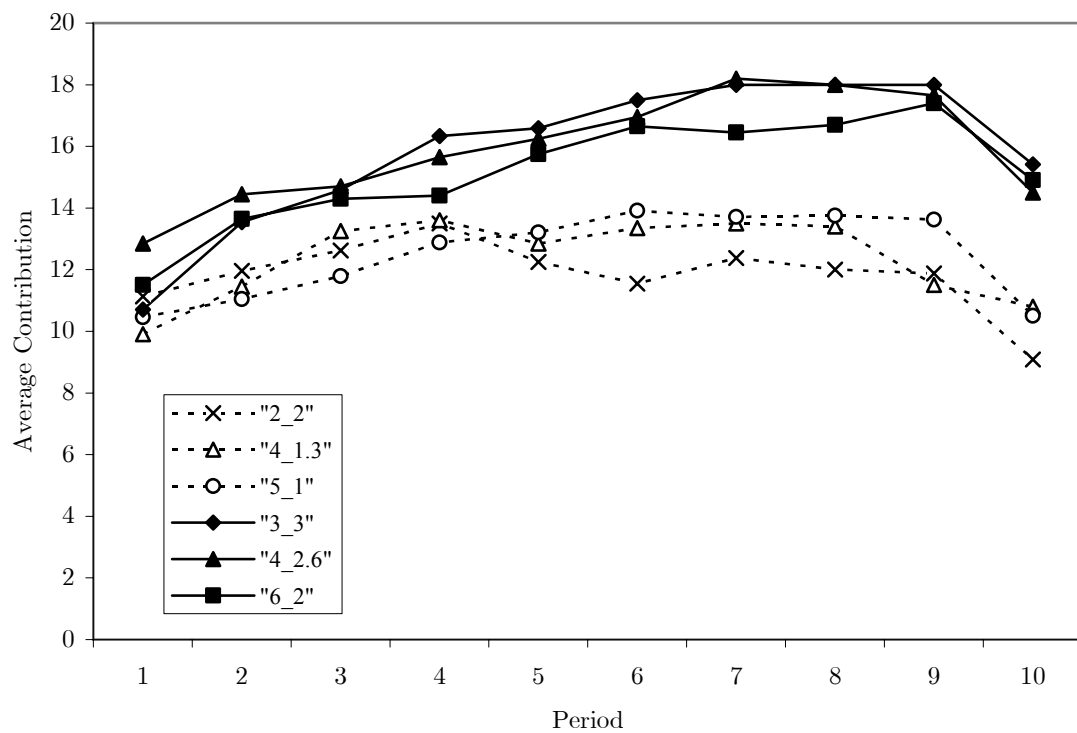


Figure 1: Evolution of Contributions

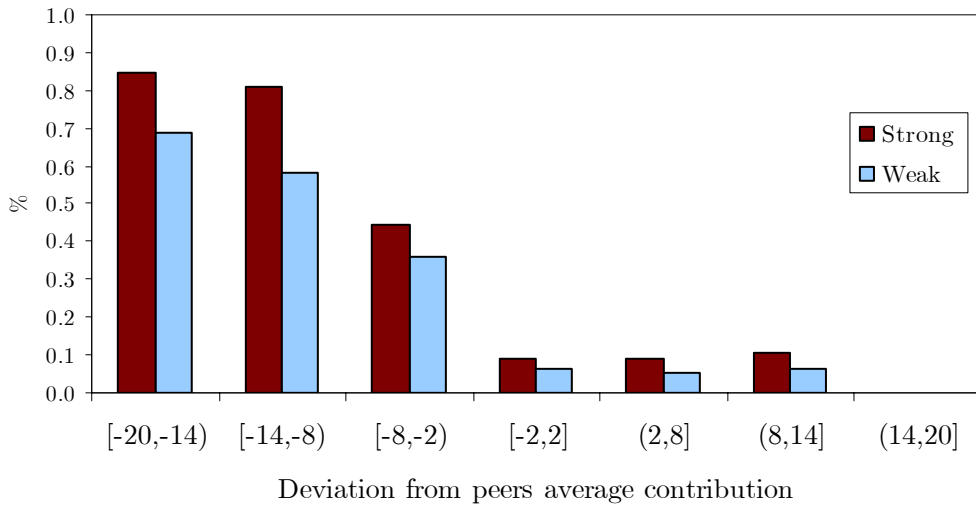


Figure 2: Punishment likelihood as a function of deviation from the contribution of peers

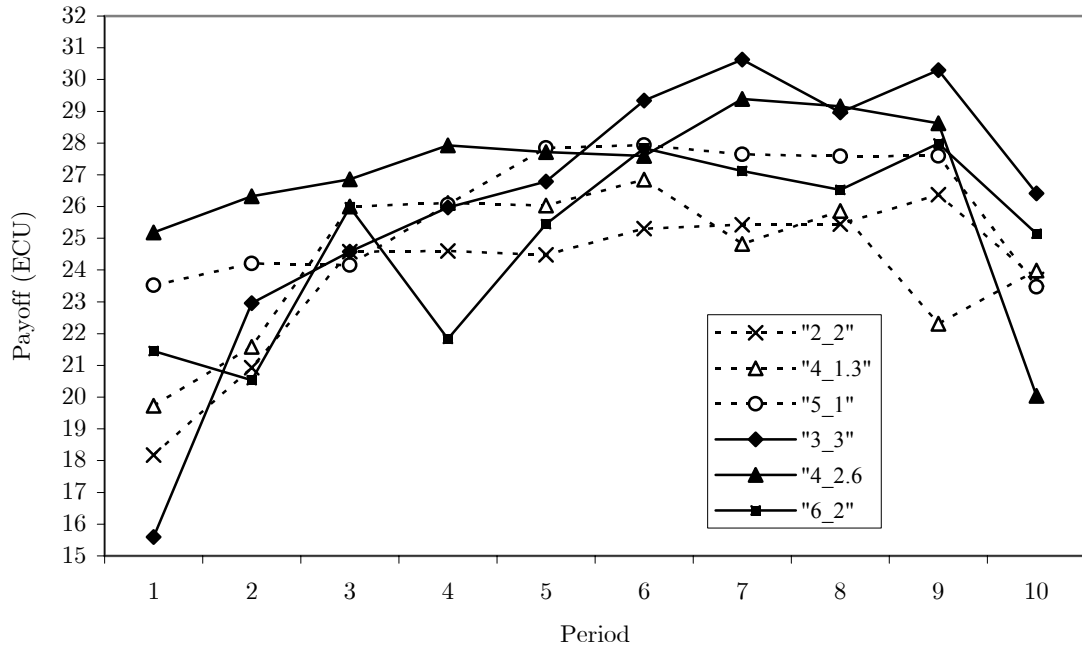


Figure 3: Evolution of Payoffs