No. 2009–19

VOTING ON PUNISHMENT SYSTEMS WITHIN A HETEROGENEOUS GROUP

By Charles N. Noussair, Fangfang Tan

March 2009

ISSN 0924-7815

Voting on Punishment Systems within a Heterogeneous Group

By

Charles N. Noussair and Fangfang Tan*

March 2009

Abstract

We consider a voluntary contributions game, in which players may punish others after contributions are made and observed. The productivity of contributions, as captured in the marginal-per-capita return, differs among individuals, so that there are two types: high and low productivity. Every two or eight periods, depending on the treatment, individuals vote on a punishment regime, in which certain individuals are permitted, but not required, to have punishment directed toward them. The punishment system can condition on type and contribution history. The results indicate that the most effective regime, in terms of contributions and earnings, is one that allows punishment of low contributors only, regardless of productivity. Nevertheless, only a minority of sessions converge to this system, indicating a tendency for the voting process to lead to suboptimal institutional choice.

JEL Classification: C92, D74, H41

Keywords: Voting, Punishment, Voluntary Contributions, Heterogeneity, Experiment

---

* CentER, Tilburg University, Tilburg, the Netherlands. Correspondence to F.Tan@uvt.nl. We thank Luc Bissonnette for technical support and CentER at Tilburg University for financial support. We also thank participants at the 2008 ESA European meeting, the 2008 NAKE research day, the GSS interdisciplinary workshop at CentER, as well as Wieland Muller, Owen Powell, Louis Putterman, Sigrid Suetens, Eric van Damme and Eline van der Heijden for valuable comments.

1. Introduction

When a group or a society faces a social dilemma, a potential role for an institution to promote or enforce a cooperative norm arises. If such an institutional structure is not imposed exogenously, it must arise endogenously from a social choice process involving the affected individuals. In a situation in which individuals are symmetric and their incentives to cooperate are perfectly aligned, one might argue that agreeing on a mechanism to enforce collective action might be relatively simple. The mechanism can require the individuals concerned to sacrifice an equal amount, all individuals can be punished similarly when deviating from appropriate behavior, and all individuals behaving appropriately can benefit equally.

On the other hand, suppose that players are heterogeneous. Then it is possible that the task of endogenously choosing an appropriate system to promote cooperation may be more difficult, and suboptimal institutions might emerge from the process. In this paper, we consider the effect that one type of heterogeneity among agents has on the institutions that emerge from a voting process. We employ an experimental approach. Our research strategy is the following. We take a setting, in which it is known from previous experimental results that effective institutions emerge from a simple voting process when individuals are symmetric. We then construct an experimental environment that is identical, except for the fact that there are two types of individual that differ only in the externality generated from their contributions, and introduce an analogous voting process. We find that in the heterogeneous environment poor institutions often emerge,

The environment that we consider is a version of a popular experimental paradigm to investigate social dilemmas, the voluntary contributions mechanism for public good provision. This is a game, in which players simultaneously choose a fraction of their endowment to contribute toward the provision of a public good. The level of contribution can be readily interpreted as a measure of cooperation. While total group payoff is increasing in the sum of members' contributions, and the social optimum is reached only when all individuals contribute all of their endowment, the dominant strategy for each player is to contribute zero. A recent focus has been on the role of decentralized sanctions, the ability of individuals to punish others based on their level of cooperation (Yamagishi, 1986; Ostrom et al., 1992; Fehr and Gächter, 2000, 2002; Masclet et al., 2003; Sefton et al., 2007). Such sanctions have been shown to be effective in increasing cooperation[1], but to have mixed effects on welfare (Bochet et al., 2006; Tan, 2008).

[1] Two of the limitations that apply to this result are the following. The first is that, as soon as counterpunishment is allowed, some of the beneficial effect is negated (Denant-Boemont et al., 2007; Nikiforakis, 2008). The second is that there is some tendency to punish cooperative players. This tendency has been termed anti-social or perverse punishment (Cinyabuguma et al., 2006), and the incidence of this behavior varies greatly depending on population studied (Gächter and Herrmann, 2008).

In the studies listed above, the experimenter imposed the sanctioning institution exogenously. There has been recent interest in endogenous punishment institutions that the affected individuals select themselves. Gürerk et al. (2005, 2006) permit individual players to choose, at the beginning of each period, between membership in a group with, and one without, sanctioning opportunities. They find that, while the majority of players opt for the sanction-free institution in the initial periods, the entire population eventually migrates to the group in which sanctioning is permitted. Botelho et al. (2005) construct a 21-period game in which players can vote, by majority rule, whether to allow for punishment in the last period after experiencing both systems with and without sanctioning possibilities for ten periods each. They find a tendency for groups to vote for the system that yielded them a higher payoff previously, which in their study was typically an institution that allowed no punishment. Sutter et al. (2006) let players decide whether to impose a punishment or reward regime at the beginning of a session, by unanimity, and find that individuals prefer rewards, even though payoffs are higher under punishment. Decker et al. (2008) allow individuals to vote for enforcement of the maximum, median, or minimum punishment assigned to an individual, and also report a tendency to vote for the particular institution that yielded the highest payoff previously. They find that the maximum rule is the most effective in generating high contributions. A number of studies find that contribution rates under mechanisms enacted endogenously by group members are higher than when the same institutions are imposed exogenously (Tyran and Feld, 2004; Kosfeld et al., 2008; Bó et al, 2007).[2]

Ertan et al. (2005) is the study most closely related to ours. They study a setting, in which players vote at regular intervals by majority on whether to allow punishment of group members who have made contributions that are (a) below-average, (b) above-average, and (c) exactly equal to the average for the group. If a punishment rule is passed, any group member may assign punishment to any individual meeting the criterion of the rule. The rules are not mutually exclusive: any, none, or all of punishment options (a) – (c) could be approved. They observe that most groups, while initially choosing not to allow any punishment at all, eventually vote to allow punishment of below-average contributors exclusively. A minority of groups ban any form of punishment throughout their interaction, and no groups ever vote to allow punishment of above-average contributors. Since both contributions and earnings are highest when individuals can be punished if and only if they contribute less than the group average, the authors conclude that groups successfully converge to the most

[2] Two recent studies have the feature that the punishment institution voted into place only governs players who vote in favor of it. In Kroll et al. (2007), agents first play a voluntary contributions game for ten periods, and make and vote on non-binding proposals of minimum total contributions. They report that voting is an empty commitment unless punishment is used to enforce the outcome. Kosfeld et al. (2008) have a similar finding that as long as there is no provision of a binding commitment, cooperation itself is difficult to attain.

efficient institutional structure. The focus of our study here is to consider whether this ability of a voting process to converge to the optimal institutional structure is robust to a particular change in the environment. This change is the existence of heterogeneity in the value to the group of individuals' contributions.

In all of the studies mentioned above, agents were homogenous in terms of the value that their contribution generated for the group, so that the tradeoff between the social benefit of cooperation and the private benefit of free riding was identical for each member of the group. In many situations, however, heterogeneity among group members may exist, due to differing productivity of their contributions. Consider, for example, a group of individuals that must complete a project for which all group members will receive equal credit. However, the effort of some group members, because of higher productivity in the required task, yields greater benefits for all than the same effort from other members. For example, one hour of work on the part of one individual may yield the same output as three hours of another individual's work. Because all group members, including the contributor, reap the benefits of an individual's effort, this heterogeneity in productivity is equivalent to a heterogeneous cost of effort among individuals, with those with higher productivity also having lower unit opportunity cost of contribution.[3] Thus, the gains and costs of a contribution depend on who made the contribution. The basic incentive structure of this situation can be captured within the experimental paradigm described above if the marginal per-capita return of a contribution (MPCR) differs depending on who is making the contribution. [4]

In this paper, we consider whether two key results of Ertan et al (2005) apply to a setting in which heterogeneity of group members' productivity, as expressed in the marginal-per-capita return of their

[3] Some experiments have distinguished between the private benefit to the individual making the contribution and the benefit of the contribution to other agents, calling these the internal and external returns, respectively. Palfrey and Prisbrey (1997), Brandts and Schram (2001) and Margreiter et al. (2005) vary the internal return, while holding constant the external return. In other words, contribution costs of players differ, but every group member benefits the same given a contributed token, regardless of the identity of the contributor.

[4] There are a few prior experiments in which MPCR differs among group members. Fisher et al. (1995) conduct a voluntary contributions game in which they assign half of the group members an MPCR of 0.75 and the other half an MPCR of 0.3. By comparing the group average contributions with those of homogenous groups featuring MPCR of 0.75 and 0.3, they conclude that the subjects seem to focus only on their own MPCR: players assigned 0.75 contribute more than those with 0.3. Reuben and Riedl (2007) study a setting in which one player has an MPCR of 1.5, and thus a dominant strategy to contribute, and the others have an MPCR of 0.5. They allow individuals to punish others after observing the contribution profile. They find that punishment is not as effective as in a control group where everyone is endowed with the same MPCR of 0.5. Fewer strong free-riders are punished, and they exhibit a weaker increase in contributions after being punished.

Margreiter et al (2005) study voting in a common pool resource game, with players with heterogeneous contribution costs. Players are asked to vote on proposals about the proportion of endowment each group member contributes, at the end of every period. If a certain proposal is selected by majority vote, it is automatically implemented in the next period. They find that compared to homogeneous groups, the number of distinct proposals is markedly larger in heterogeneous groups, but fewer agreements are reached by majority voting.

contributions, exists. The two results are that (1) permitting but restricting permissible punishment to below-average contributors yields the highest payoff among punishment institutions that condition on deviations from average contribution level, and (2) when engaged in repeated opportunities to vote, groups converge to this punishment institution over time. In our experiment, as in Ertan et al., individuals vote at regular intervals on whether individuals are permitted have punishment directed toward them. After a regime is selected, based on majority vote, it is in effect for that group for a fixed and known number of periods. As in the Ertan et al. study, we vary, as a treatment variable, the number of periods that the results of one vote are in effect. Studying different voting terms is a potentially important aspect of institutional design, and the effect of a punishment system could well depend on the length of time a system is locked in and not subject to change.

The parametric structure of our experimental environment follows Tan (2008). She studies a four-person voluntary contributions game with two types of agent. Two players have an MPCR of 0.9, so that each token they contribute yields 0.9 tokens to all group members, and the other two players have an MPCR of 0.3. All agents are permitted to punish any other agent in any period. Tan finds that punishment is not very effective in increasing contributions among heterogeneous agents. In groups that achieve cooperation, high MPCR players punish low MPCR players frequently for their free-riding behavior. However, when controlling for the contribution level of the recipient of punishment, high MPCR players receive more punishment than those with low MPCR.

There is reason to believe that heterogeneity of MPCR may make a difference in which institutions emerge from the voting process. The different costs of contribution among players may inhibit the establishment of a contribution norm, and create differing beliefs among agents about the appropriate level of contribution that each type should make. This may make it more difficult to achieve consensus on which punishment system to implement and may lead to a conflict between different types of agent. Such conflicts may prove sustained and durable, with adverse long-term effects on contributions and welfare. Indeed, as described in section four, the principal results we obtain are the following. We find that, consistent with Ertan et al. (2005), the most effective institution, in terms of contributions and earnings, is one that allows punishment of below-average contributors only, regardless of productivity type. However, unlike in the Ertan et al. environment, groups often fail to enact this institution, especially when the votes are held relatively frequently. Under these conditions, groups typically establish inefficient regimes, and particularly common is a system in which no punishment is permitted. No group ever votes to enable punishment of all individuals, regardless of their type or contribution level. Players are more likely to vote to allow punishment of below-average contributors and the type other than their own, and they attempt to escape from future penalty opportunities by disallowing punishment rules targeting their own type.

For many groups, this behavior appears to create an insurmountable roadblock to the establishment of the appropriate institution.

The remainder of the paper is organized as follows. In Section 2, we describe the experiment and in Section 3, we advance several hypotheses about the performance of different punishment regimes. In Section 4, we present an analysis of the data. Finally, in Section 5, we make some concluding remarks.

2. The Experiment

2.1 General Setting

The experiment consisted of six sessions that were conducted at CentER Lab, at Tilburg University in the Netherlands. There were two treatments, the *Short-Term* and the *Long-Term* treatments. Each treatment was in effect in three of the sessions. Forty-eight subjects, among whom 42% were females, and all of whom were students at Tilburg University, participated in the study. Some of the subjects had previously participated in economic experiments, but all were inexperienced with the voluntary contributions mechanism. Each subject took part in only one session of the study. On average, a session lasted about 80 minutes (including initial instruction and payment of the subjects), and a subject earned an average of 454 tokens (approximately 18.16 euros). The experiment was programmed and conducted with the z-Tree software (Fischbacher, 2007).

Each session included eight participants that were separated into two groups of four. At the start of each session, the computer program randomly assigned the subjects into different groups according to their choices of terminal upon entering the room for the session. All individuals remained in the same group for their entire 30-period experimental session. All 30 periods of play counted toward final earnings, and there were no practice periods at the beginning of the sessions. At the beginning of each period, every player was randomly given an identification number from 1 to 4 to distinguish her actions from those of the others during that period. To prevent the formation of individual reputations, however, the numbers were randomly reallocated at the beginning of every period.

Productivity heterogeneity was generated by randomly assigning half of the group members a high MPCR of 0.9 (players of this type will be referred to as type A players) and the other half a low MPCR of 0.3 (type B players). Participants were informed of their type at the beginning of the session, and their types remained fixed for the duration of the session.[5] The instructions used in the experiment were modified on the basis of those used in Ertan et al. (2005) and Tan (2008).

[5] Neutral language was used in the experiment. Players with MPCR of 0.9 were referred to as "type A" and players with MPCR of 0.3 were "type B". Moreover, potentially biased terms such as “contribution” and “punishment” were avoided. For example, punishment was termed as "points that reduce another player's income".

2.2 Timing

The 30 periods that made up each session were divided into three segments, as illustrated in figure 1. In the first segment, comprising periods 1 – 3, subjects played the voluntary contributions game without the possibility of punishment. In the second segment, consisting of periods 4 – 6, a second stage was added to the game in which any player could punish any other player, after observing all players' contributions. In the third segment, which made up the remainder of the session (periods 7 -30), the punishment system in place depended on the outcome of a voting process. Voting took place every two periods in the Short-Term treatment, and every eight periods in the Long-Term treatment.

In each period of the first segment, the following occurred. Each subject was endowed with ten tokens, with a conversion rate of 25 tokens = 1 Euro. Subjects simultaneously and independently divided their endowment between a private account and a group account. The income of an individual equaled the number of tokens she put in her private account, plus .9 times the total contributions of type A players in her group, plus .3 times the total contribution of type B players in her group. That is, a player's income in each period equaled

(1) $$I_{ij} = 10 - C_{ij} + 0.9 \times \sum_{j=A} C_A + 0.3 \times \sum_{j=B} C_B$$

where C_{ij} is the contribution of the *i*th player of type *j*. This calculation was displayed on subject *i*'s computer screen together with the contributions and earnings of all group members at the end of each period.

In period 4 – 6, each period was made up of two stages. There was a second, punishment, stage subsequent to the contribution stage described above. In the second stage, subjects were given the opportunity to send points ranging from 0 to 10 to any group member. Every point that a particular subject sent to another reduced the sender's earnings by one token and reduced the earnings of the recipient by two tokens. Thus, subject *i*'s income in each period equaled:

(2) $$I_{ij} = 10 - C_{ij} + 0.9 \times \sum_{j=A} C_A + 0.3 \times \sum_{j=B} C_B - \sum_{k \neq i} P_{ik} - 2 \times \sum_{k \neq i} P_{ki}$$

Where $\sum_{k \neq i} P_{ik}$ was the sum of points subject *i* sent to all group members, and $\sum_{k \neq i} P_{ki}$ was the sum of points she received from all others. At the end of each period, the computer displayed the subject's own type, the tokens she and all group members contributed, the total number of points she received and assigned to others, her income for the current period and how it was calculated. Subjects were not informed about how much punishment other individuals sent or received.

In the third segment of each session, periods 7 – 30, the following took place. Every two periods in the Short-Term treatment, as well as every eight periods in the Long-Term treatment, a voting stage occurred at the beginning of a period. During the voting stage, every subject was required to answer each of the following four questions by clicking a box that corresponded to either (a) yes, (b) no, or (c) no preference.[6] The four questions were the following:

I vote to allow a person's earnings to be reduced if the person is a:

(1) Type A player assigning less than the average amount to group account.

(2) Type A player assigning more than the average amount to group account.

(3) Type B player assigning less than the average amount to group account.

(4) Type B player assigning more than the average amount to group account.

After all subjects gave their answers, the computer tabulated the votes. If the number of "Yes" votes on one of the questions exceeded the number of "No" votes, the reduction specified in the question was allowed; otherwise it was not. A "No preference" vote did not count towards the voting outcome. Since there were four questions, the number of possible outcomes, or punishment institutions, was $2^4 = 16$. Subjects were informed of the punishment system instituted, and the number of periods this institution would be in effect. In the Long-Term treatment, a vote occurred every eight periods, and the same institution remained in effect for the eight-period interval following the vote. In the Short-Term treatment, a vote took place every two periods, and the resulting system was in effect for the two periods.

[Figure 1: About Here]

In every period, regardless of whether a vote occurred in the current period, the contribution and punishment stages occurred in a similar manner as in the second segment. During the punishment stage, subjects decided how many points to send to members meeting the punishment requirement, but were required by the computer program to abide by the restrictions resulting from the last vote, whether it occurred in the current or in a prior period. The feedback presented to subjects at the end of a period in the third segment was the same as in the second segment.

2.3 The Experiment of Tan (2008)

[6] Ertan et al. (2005) also included an option to vote to allow punishment of those players whose contributions were exactly equal to the average. This option is not included in this experiment, however, because if two more questions concerning average contributors of each type are included, the potential number of punishment systems would increase to 64.

Tan (2008), in a related study, examines the effect of an exogenously imposed punishment institution on players with heterogeneous productivity. A number of features of that study are similar to the one reported here. The parametric structure of the game is the same in the two studies. Players played the voluntary contributions game under a fixed matching protocol, with two high productivity players with an MPCR of 0.9, and two low productivity players with an MPCR of 0.3. In one treatment, no punishment was possible, as in periods 1 – 3 in the study reported here. In another treatment punishment of any other player was permitted, as in periods 4 – 6 here.

However, there are important differences between the two studies. In the Tan (2008) study, the punishment system is imposed exogenously rather than enacted endogenously by participants themselves. Furthermore, in the Tan experiment, the length of a session is 15 periods, and the same punishment condition remained in effect for the entire session. While it is not the principal purpose of the study reported here, the similar parametric structure between our experiment and Tan (2008) allows us to make rough comparisons between the two studies, and we do so with regard to aspects of individual behavior in section 4.

3. Hypotheses

Our analysis is organized as a test of several hypotheses. The first two concern whether particular results obtained in Ertan et al. (2005) generalize to our environment. The first hypothesis is that the most effective system for promoting high welfare is to permit punishment of only below-average contributors, regardless of their productivity, a system we refer to hereafter as *Pun-Low*. The rationale for the hypothesis is that such a system enables the group to punish low contributors to influence their behavior, and prohibits punishment of high contributors in order to encourage them to continue their behavior. *Pun-low* was the most effective of all of the available systems in Ertan et al.'s (2005) environment.

Hypothesis 1 (Efficient Punishment Regime Hypothesis): *The most efficient punishment regime, in the sense of yielding the highest welfare, is to allow punishment of below-average contributors only, regardless of productivity (Pun-Low).*

Ertan et al. observed that *Pun-Low* was reached consistently after several iterations of the voting process. We consider whether this finding carries over to our setting with heterogeneous agents. While there is a powerful collective incentive to converge to the most efficient arrangement, there is also reason to believe that it may not do so in an environment with heterogeneous agents. The work of

Margreiter et al. (2005) indicates that voting does not guarantee that an institution with high contributions and welfare emerges when contribution costs vary among group members. More generally, heterogeneity in MPCR leads to lower contributions (Fisher et al. 1995) even in settings in which punishment is possible (Tan, 2008), and this difficulty in cooperating may carry over to the institution formation phase. Nonetheless, as a null hypothesis we propose that the voting process will behave effectively in discovering the most efficient arrangement:

Hypothesis 2 (Punishment Regime Convergence Hypothesis): *Convergence to the most efficient rule occurs over the course of the voting process.*

Note that either Hypotheses 1 or 2 may be supported while the other one is not supported. *Pun-Low* may lead to the greatest level of welfare, but may not be attained with the voting process. An institution other than *Pun-Low* may generate the highest welfare and also be the outcome of the voting process. The next hypothesis concerns the difference between treatments. A priori, the effect of lengthening the time that an institution is in effect on per-period welfare is ambiguous. On one hand, longer governance duration implies a greater commitment to the results of a given vote, and that may create greater incentives to form more effective institutions. On the other hand, the shorter governance duration in the Short-Term treatment offers groups more opportunities to search for effective institutions, and to discard ineffective ones, than does the Long-Term treatment. Since these two effects operate in different directions[7], we hypothesize that the contributions made and the welfare attained are not different between the Short-Term and the Long-Term treatments.

Hypothesis 3 (Governance Duration Hypothesis): *Contributions and welfare are not significantly different between the Short-Term and the Long-Term treatments.*

4. Results

The first hypothesis concerns the relative performance of different institutional structures in terms of contributions and welfare. Table 1 displays the average group contributions and earnings under each institution across treatments. The table shows how many times each punishment system was enacted, how many periods it was in effect, the average contribution and welfare level (measured as subject earnings) it generated, and its rank among the systems in terms of contribution and welfare levels.

[7] Ertan et al. (2005) also vary the length of time the results of a vote are in force, but do not discuss the effect of governance duration on outcomes.

Nine out of 16 possible combinations of punishment rules are enacted at least once in our dataset. The four most common combinations are: (1) to disallow punishment of any agent (which we will refer to as *No-Pun*), (2) to allow punishment of below-average contributors regardless of productivity (*Pun-Low*), (3) to allow punishment of Type B players making below-average contributions (*Pun-B-Low*) and (4) to allow punishment of Type A players making below-average contributions (*Pun-A-Low*). These four structures account for almost 90% of the total voting outcomes. No group ever votes to permit punishment of all agents. Result 1 summarizes the main findings concerning the relative performance of the institutions with regard to contributions and efficiency.

Insert Table 1 about here

RESULT 1: The efficient punishment regime hypothesis (Hypothesis 1) is supported. The most effective regime, in terms of both contributions and earnings, is *Pun-Low*, which allows punishment of players with below-average contributions only, regardless of productivity. This is the case in both the Short-Term and Long-Term treatments.

SUPPORT: According to table 1, the four most successful institutions all allow punishment of at least some below-average contributors. *Pun-Low* is the most effective institution in terms of contributions in both treatments, and in terms of welfare in the Short-Term treatment. In the Long-term treatment, *Pun-Low* is the second-ranked system of welfare after *Pun-A-Low*. Overall, in *Pun-Low*, the mean contribution level is almost three quarters of the total endowment, which is 73% more than the next best system, *Pun-A-Low*. A Mann-Whitney rank-sum test, using average contributions in each session for the periods that the system is in effect as the unit of observation, indicates that contributions in *Pun-Low* are significantly greater than in *Pun-A-Low* ($z = -2.364$, $p < 0.05$) and than in *Pun-B-Low* ($z = -2.030$, $p < 0.05$). A similar result holds for welfare. Although welfare is not significantly greater in *Pun-Low* compared to *Pun-A-Low* ($z = -0.447$), it is significantly greater than under *Pun-B-Low* ($z = -2.030$, $p < 0.05$). □

There are a number of other interesting patterns evident in the table. *No-Pun* is considerably less effective in generating contributions and earnings than the systems that allow punishment of below-average contributors. There are also some differences in the incidence and relative performance of the institutions between treatments. Institutions permitting punishment of only above-average but not below-average contributors appear only in the Short-Term treatment. The inefficient *No-Pun* institution is in effect in more than twice as many periods in the Short-Term treatment than in the

Long-Term treatment. The *Pun-A-Low* institution is more effective in the Long-Term treatment than in the Short-Term treatment both in terms of contribution and earnings, while the opposite holds for *Pun-B-Low*.

Insert Table 2 about here

Table 2 reports the results of a regression estimating the effect of the different institutions on contribution and welfare levels. The data in the first three periods of the sessions, in which no punishment regime is in effect, are the baseline of the regressions. Unrestricted punishment, in effect in periods 4 – 6 of each session, and in which players can reduce the earnings of any other player, does not lead to higher contribution levels, but does lower earnings, in both treatments. This is indicated by the estimates for β_1. The significantly positive β_2 across all equations confirms the robust effect of allowing for punishment of below-average contributors: this increases group average contribution levels and earnings relative to the baseline. The significantly negative coefficient β_5 in indicates that if players vote out to disallow any form of punishment during the voting stage, group average contributions and earnings decrease relative to a situation in which the same system is imposed exogenously.

The second hypothesis concerned whether the most effective institutional structure emerges from the voting process. Our findings are summarized in result 2.

Insert Figure 2 about here

RESULT 2: Punishment Regime Convergence Hypothesis (Hypothesis 2) is not supported. Institutional rules fail to converge to the efficient *Pun-Low* system in either treatment.

SUPPORT: Figure 2 shows the incidence of each institution in each of the sequence of votes in the two treatments. The horizontal axis of the figures represents the timing of the vote, with voting time "1" indicating the first vote in a session, which occurs at the beginning of period 7. The second "voting time" occurs in period 9 in the Short-Term and in period 15 in the Long-Term treatment. The vertical axis represents the number of groups, out of a total of six groups, that choose each system. None of the six groups votes for *Pun-Low* during the last voting stage in the Short-Term treatment, while only three of the six groups do so in the Long-Term treatment. □

As we can see from the data in the figures, the relatively efficient *Pun-Low* institution is chosen with greater frequency in the Long-Term treatment. However, the positive effect on welfare of the relatively frequent choice of Pun-Low in the Long-Term treatment is not sufficient to offset the even greater increase in contributions and welfare that occurs when subjects in the Short-Term treatment select *Pun-Low*. As stated in the introduction section, we vary the duration of the time interval that the results of an individual vote are in effect. Result 3 summarizes our findings.

RESULT 3: The Governance Duration Hypothesis (Hypothesis 3) cannot be rejected. That is, we cannot reject the null hypothesis that the Short-Term and the Long-Term treatments are the same in terms of both contributions and welfare.

SUPPORT: Mann-Whitney rank sum tests of differences in contributions and welfare between the Short-Term and Long-Term treatment in periods 7 to 30 suggest that neither distributional difference is significant between the two treatments ($z = 0.320$ for contribution and 0.801 for earnings). □

Figures 3a and 3b show the time series of earnings and punishment points assigned for each group, in the Long-Term and Short-Term treatments respectively. The vertical axis indicates the per-capita earnings in tokens (the maximum possible is 24, and the level corresponding to zero contribution and zero punishment is 10), and the number of punishment points allocated per capita. The horizontal axis is the period number. Both figures show that, while *Pun-Low* performs better than the other systems on average in terms of earnings, it remains inconsistent and only reaches welfare levels close to the potential maximum in some instances. It is also clear that punishment is effective in raising contributions, at least in the short run; in almost every period after which any punishment points are assigned, there is an increase in group earnings. The *No-Pun* institution consistently leads to zero or close to zero contributions, as reflected in average earnings near ten tokens. In the Long-Term treatment, three groups achieve close to the maximum possible level of earnings, and they do by enacting *Pun-Low* or *Pun-A-Low*. In the Short-Term treatment, institutional changes are quite frequent with at least four changes from one vote to the next occurring in each group. Only two groups achieve close to maximal earnings by the end of their session. One does so by enacting *Pun-Low*, and the other with *Pun-B-Low*.

Figures 3a and 3b About Here

Figures 4a and 4b show the voting behavior of individuals based on their type and contribution level in the period immediately preceding the vote. Each panel in the figures corresponds to the votng behavior of one four types/contribution profiles in one of the treatments. Each bar indicates the percentage voting in favor, voting against, and abstaining from each of the four punishment rules. The figures are constructed by classifying each player into one of the four categories: type A below-average contributors (abbreviated to AL), type A above-average contributors (AH), type B below-average contributors (BL) and type B above-average contributors (BH) based on her actual contribution one period before the voting stage. Then the number of "yes", "no" and "no preference" votes are summed.

Insert Figure 4a and 4b about here

The figures illustrate the sharp conflicts between above-and below-average contributors, as well as between type A and type B players. When above-average contributors vote in favor of punishment of below-average contributors, they are much more likely to vote in favor of punishment of the other type. Likewise, when they vote against allowing punishment of above-average contributors, they are more likely to vote in favor of banning this punishment for their own type. Below-(above-) average contributors are more willing to vote to allow punishment of above-(below-) average contributors than of players who contribute similarly to themselves.[8] These patterns suggest that players try to shut down punishment channels that may point to them in the future.

Insert Table 3 about here

Consider the following probit regression.

$$V_{ik}^{t} = \beta_0 + \beta_1 yourself + \beta_2 opp\_MPCR + \beta_3 low\_con + \beta_4 type_i + \beta_5 typeA + \beta_6 opp\_con + \beta_6 pun\_rec + \beta_7 pun\_sent + \varepsilon_{ik}^{t} \quad (3)$$

The dependent variable equals 1 if subject *i* votes to permit a specific punishment rule *k* in period *t*, and 0 otherwise. The first explanatory variable, *yourself*, is a dummy variable that equals 1 if the rule allows punishment of a player with the same type and similar contribution behavior as player *i*. *Opp_MPCR* equals 1 if the voted item *k* refers to the other type of player, and *Opp_con* is analogous

[8] There is one exception. In the Long-Term treatment, AH players rather than AL players are more willing to allow for punishment of BH players: 43.5% of AH players vote to allow for punishment of BH players while only 30.8% of AL players vote to allow for punishment of BH players.

for the opposite contribution level relative to the average. *Low_con* equals 1 if the voted item refers to someone making a below-average contribution. *Typei* is also a dummy equal to 1 if the player voting is a type A player, whereas *TypeA* equals to 1 if the voted rule targets type A players. *Pun_rec* and *Pun_sent* are continuous variables representing the total number of punishment points received from other players and sent to other players, respectively, in the period immediately preceding each vote. Result 4 summarizes the findings.

RESULT 4: (Voting Behavior) In both treatments, the willingness of players to vote on punishment of a certain player is greater (i) if the punishment rule refers to the opposite MPCR, (ii) if the rule refers to below-average contributors, and (iii) if the punisher has a high MPCR.

SUPPORT: The estimates in table 3 show highly significant positive coefficients of β_2 - β_4. This indicates that players are more willing to vote in favor of a punishment rule if it targets the opposite productivity type (β_2), below-average contributors (β_3), and if the player voting has a high productivity level (β_4). □

How does *Pun-Low* promote high contributions? The decision to free ride can be viewed as a result of a cost-benefit calculation. The possibility of punishment of below-average contributors may lower the return from making low contributions, and those for whom it is relatively low-cost to increase their contributions do so in response to the availability of punishment of low contributors.

Consider first who receives punishment. Previous research indicates that the amount of punishment points assigned is influenced by the difference in contribution between the punishing and the punished agent, as well as the difference between the negative deviations of the recipient's contribution from the group average level (Fehr and Gächter, 2000; Masclet et al., 2003; and Falk et al., 2005). Consider the following estimated equation, whose estimates are given in table 4.

$$
\begin{aligned}
P_{ik}^{t} &= \beta_0 + \beta_1 \left(\max\left\{0, c_i^t - c_k^t\right\}\right) + \beta_2 \left(\max\left\{0, c_k^t - c_i^t\right\}\right) \\
&+ \beta_3 \left(\max\left\{0, \overline{c}^t - c_k^t\right\}\right) + \beta_4 \left(\max\left\{0, c_k^t - \overline{c}^t\right\}\right) + \beta_5 type_i + \beta_6 type_k + \varepsilon_{ik}^t
\end{aligned} \tag{4}
$$

where $type_i = 1$ if the punisher *i* has an MPCR of 0.9; $type_k = 1$ if the punished *k* has an MPCR of 0.9, and $\bar{c}_t$ is the average contribution in the group in period *t*. Because of the large number of zero values for the dependent variable, we estimate this specification by Tobit models with standard errors robust to within group correlation.

Empirical evidence also shows that low contributors on average respond to punishment by raising their contributions in the subsequent period (Fehr and Gächter, 2000; Masclet et al, 2003). The change in the contribution of player *i* between period *t* and *t*+1 can be modeled as:

$$(5)\quad \begin{aligned} c_i^{t+1} - c_i^t &= \beta_0 + \beta_1 \left(\sum_k P_{ki} \right) \\ &+ \beta_2 \left(\overline{c}^t - c_i^t \right) + \beta_3 type_i + \beta_4 \left(type_i \times \sum_k P_{ki} \right) + \varepsilon_i^t \end{aligned}$$

where $type_i = 1$ if *i* is a high-productivity player. β_1 measures the effect of the total number of points subject *i* receives on her change in contribution from one period to the next, and β_2 is the effect of the difference between individual *i*'s contribution and her group average contribution level in period *t*. β_3 measures any difference in overall contribution change between the two types, and β_4 registers a differential response to punishment on the part of high and low productivity types. The estimates of models (4) and (5), for the data from the exogenously-imposed unrestricted punishment system studied in Tan (2008), are also included in tables 4 and 5 under the column labeled Unrestricted Punishment. In table 5, only the observations in which an individual's contribution in period *t* is lower than his group's average for the period are included. Result 5 summarizes the main findings from the estimation of (4) and (5).

Insert Tables 4 and 5 about here

RESULT 5: (Punishment Behavior and Responses) Under *Pun-Low*, the level of monetary sanction is increasing in the negative difference between the contributions of the recipient and the punisher in both treatments. Players increase their contributions more in the subsequent period, the farther their contribution is below the group average. The two types of player respond similarly to the receipt of punishment.

SUPPORT: The estimates in table 4 show that in both *Pun-Low* and the unrestricted punishment regime, there is a positive relation between the punishment points player *i* sends to player *k* and the extent to which player *k*'s contribution below that of player *i*'s. Unlike under unrestricted punishment, there is no relationship between the type of either the sanctioner or the sanctioned party in terms of punishment behavior. Table 5 indicates that in the *Pun-Low* regime, the contribution level increases significantly, the more a player's contribution is below group average (β_2). The insignificance of β_1 coefficient suggests that it is not the actual sanction that, but rather the possibility of punishment, triggers increases in contribution when punishment of below-average contributors in enabled. The

significant β_4 coefficient in the Unrestricted Punishment data indicates that type A players are more likely to increase their contribution in response to punishment than type B players. However, this difference between types is not observed under *Pun-Low*. □

5. Conclusion

We have studied the voting behavior of groups that face a social dilemma. At regular intervals they vote to select a punishment institution, a set of conditions under which individuals may punish others. The issue is whether the most efficient institution, in terms of yielding maximal gains to the group, emerges from the voting process. We pose this question for an environment in which players are heterogeneous in terms of the benefit that their contributions yield to the group.

We observe that institutions that allow punishment of low contributors while immunizing high contributors perform well in generating high average contributions and welfare levels. This extends a previous result obtained by Ertan et al (2005) in a setting with symmetric players. Little punishment is actually applied; the threat of punishment is sufficient to generate high levels of cooperation.

However, we find that generally, groups do not adopt the most profitable institution even after having repeated opportunities to vote for its enactment. The heterogeneity of players generates conflicts as players attempt to prevent punishment that can be directed at themselves, while attempting to enable it for other people. The result is that groups often find themselves with no ability to punish some or all free riders, and thus without a mechanism for enforcing high contributions.

References

Bó, P., A. Foster and L. Putterman (2007), "Institutions and behavior: Experimental evidence on the effects of democracy", Working papers 2007-9, Department of Economics, Brown University.

Bochet, O., T. Page and L. Putterman (2006), "Communication and punishment in voluntary contribution experiments", *Journal of Economic Behavior and Organization*, 60: 11–26.

Botelho, A., G. Harrison, L. Pinto and E. Rutstrom (2005), "Social norms and social choice", Working paper, Department of Economics, University of Central Florida.

Cinyabuguma, M., T. Page and L. Putterman (2006), "Can second-order punishment deter perverse punishment?", *Experimental Economics*, 9:265–279.

Brandts, J. and A. Schram (2001): Cooperation and noise in public goods experiments: "applying the contribution function approach", *Journal of Public Economics*, Vol.79 (2), 399-427.

Decker, T., A.Stiehler and M. Strobel (2003): "A comparison of punishment rules in repeated public good games", *Journal of Conflict Resolution*, Vol. 47 (6), 751-772.

Denant-Boemont, L., D. Masclet and C.Noussair (2007), "Punishment, counter-punishment and sanction enforcement in a social dilemma experiment", *Economic Theory*, Vol.33 (1): 145-167.

Ertan, A., T. Page and L. Putterman (2005), "Can endogenously chosen institutions mitigate the free-rider problem and reduce perverse punishment?", Working papers, Department of Economics, Brown University.

Falk, A., E. Fehr and U. Fischbacher (2005), "Driving forces behind informal sanctions", *Econometrica*, Vol.73 (6): 2017-2030.

Fehr, E., and S. Gächter (2000), "Cooperation and punishment in public goods experiments", *American Economic Review*, Vol. 90(4): 980-994.

Fehr, E., and S. Gächter (2002), "Altruistic punishment in humans", *Nature*, 415:137-140.

Fischbacher, U. (2007), "Z-Tree - Zurich toolbox for readymade economic experiments", *Experimental Economics*, 10(2): 171-178.

Fisher, J, R. Isaac, J. Schatzberg and J.Walker (1995), "Heterogeneous demand for public goods: Behavior in the voluntary contributions mechanism", *Public Choice*, 85: 249-266.

Gächter, S., and B. Herrmann (2008), "Reciprocity, culture, and human cooperation: Previous insights and a new cross-cultural experiment", Discussion Papers 2008-14, The Centre for Decision Research and Experimental Economics, School of Economics, University of Nottingham.

Gürerk, O., B. Irlenbusch and B Rockenbach (2005), "On the evolvement of institutions in social dilemmas", Working paper, University of Erfurt.

Gürerk, O., B. Irlenbusch and B Rockenbach (2006), "The competitive advantage of sanctioning institutions", *Science*, Vol. 312: 108-110.

Kosfeld, M., A. Okada and A. Riedl (2006), "Institution formation in public goods games", CESifo Working paper No. 1794.

Kroll, S., T. Cherry and J. Shogren (2007), "Voting, punishment, and public goods", *Economic Inquiry*, Vol. 45 (3): 557-570.

Margreiter, M., M. Sutter and D.Dittrich (2005), "Individual and collective choice and voting in common pool resource problem with heterogeneous actors", *Environmental and Resource Economics*,

Vol.32 (2): 241-271.

Masclet, D., C. Noussair, S. Tucker and M. Villeval (2003), "Monetary and non-monetary punishment in the voluntary contributions mechanism", *American Economic Review*, 93(1): 366-380.

Nikiforakis (2008), "Punishment and counter-punishment in public good games: Can we really govern ourselves?" *Journal of Public Economics*, Vol.92 (1-2):91-112.

Ostrom, E.; J. Walker, and R. Gardner (1992), "Covenant with and without a sword: Self - governance is possible", *American Political Science Review*, 86(2): 404–417.

Palfrey, T., and J. Prisbrey (1997), "Anomalous behavior in public goods experiment: how much and why", the American Economic Review, Vol. 87(5): 829-846.

Reuben, E., and A. Riedl (2007), "Public Goods Provision and Sanctioning in Privileged Groups", IZA Discussion Papers 2916, Institute for the Study of Labor (IZA).

Sefton, M., R. Shupp and J. Walker (2007), "The effect of rewards and sanctions in provision of public goods", *Economic Inquiry*, Vol. 45 (4):671-690.

Sutter, M., S. Haigner and M. Kocher (2006), "Choosing the stick or the carrot? – Endogenous institutional choice in social dilemma situations", CEPR Discussion paper No. 5497.

Tan, F. (2008), "Punishment in a linear public good game with productivity heterogeneity", *De Economist*, 156 (3): 269-293.

Tyran, JR. and L. Feld (2004), "Achieving compliance when legal sanctions are non-deterrent", *Scandinavian Journal of Economics*, Vol.108 (1): 135-156.

Yamagishi, T. (1986), "The provision of a sanctioning system as a public good", *Journal of Personality and Social Psychology*, Vol. 51(1): 110-116.

Table 1: Frequency and Average Outcomes of Different Punishment Systems

	Long- Term Treatment						Short-Term Treatment					
	Number of Times Enacted	Number of Periods in Effect	Contribution rank	Average Contributions	Welfare rank	Average Welfare	Number of Times Enacted	Number of Periods in Effect	Contribution rank	Average Contributions	Welfare rank	Average Welfare
Pun-Low	8	64	1	6.94	2	18.15	12	24	1	8.20	1	20.67
Pun-A-Low	4	32	2	4.13	1	20.22	9	18	3	3.97	4	13.27
Pun-B-Low	1	8	4	0.91	5	10.43	18	36	2	4.57	2	15.21
No-Pun	3	24	5	0.22	4	10.29	25	50	7	0.41	7	10.50
PunAL&PunBH	1	8	3	2.34	3	12.98	2	4	5	2.88	3	14.40
PunAH&PunBL	1	8	6	0.28	6	9.66	1	2	4	3.63	6	11.03
Pun-B-High	--	--	--	--	--	--	1	2	9	0.00	9	10.0
Pun-A-High	--	--	--	--	--	--	3	6	6	2.00	5	12.3
Pun-B	--	--	--	--	--	--	1	2	8	0.13	8	10.03
Total	18	144					72	144				

Table 2. Average Group Contributions and Earnings as a Function of Punishment System in Effect

Dependent variable: Group average contributions in period t, $\overline{C}_i$ and group average earnings $\overline{I}_i$ in period t.

	Average Contributions		Average Earnings	
	Long-Term Treatment	Short-Term Treatment	Long-Term Treatment	Short-Term Treatment
β_1 Unrestricted Punishment	0.306	0.657	-4.979***	-9.434***
	(0.578)	(0.492)	(1.127)	(2.486)
β_2 Pun Low Contributors	2.733***	2.989***	1.978**	5.351**
	(0.497)	(0.502)	(0.969)	(2.317)
β_3 Punish A Low Contributors	-1.780***	-0.535	-1.669	-1.448
	(0.626)	(0.507)	(1.220)	(2.486)
β_4 Punish B Low Contributors	-0.875	1.030**	-2.295	0.496
	(0.826)	(0.451)	(1.615)	(2.134)
β_5 No Punishment	-1.934**	-1.971***	-2.921*	-4.217**
	(0.843)	(0.417)	(1.642)	(2.025)
β_0 Constant	4.062***	3.459***	16.704***	14.718***
	(0.411)	(0.339)	(0.801)	(1.171)
Adjusted R squared	0.353	0.399	0.273	0.459
Observations	164	166	164	166

Notes: *10% significance; **5% significance, ***1% significance. Contribution data corresponding to infrequently enacted institutions such as *PunAL&PunBH*, *PunAH&PunBL*, *PunAH, PunBH and PunB* are excluded because of the insufficient number of observations. The model specification is a fixed effect model with the variable "group" as the individual effect. The standard errors are robust within group correlation. A Chow test rejects the null hypothesis that the coefficients of Long-Term and Short-Term treatments are equal. Therefore, we conduct a separate estimation for each treatment.

Table 3. Voting Patterns

Dependent variable: Voting of player *i* in favor of permitting punishment of player *k* at time *t*, $\overline{V}_{ik}^{\,t}$

	Long-Term Treatment	Short-Term Treatment
β_1 Yourself	-0.720*	0.182
	(0.397)	(0.174)
β_2 Opposite MPCR	0.917***	1.259***
	(0.279)	(0.126)
β_3 Below Average Contributor	1.467***	0.968***
	(0.221)	(0.093)
β_4 Type i	1.017**	0.697***
	(0.406)	(0.217)
β_5 Type A Player	0.208	-0.151*
	(0.203)	(0.091)
β_6 Opposite Contribution	0.252	0.526***
	(0.261)	(0.119)
β_7 Punishment Received at Period (t-1)	-0.007	-0.045
	(0.073)	(0.033)
β_8 Punishment Sent at Period (t-1)	0.051	0.010
	(0.072)	(0.017)
β_0 Constant	-2.060***	-1.330***
	(0.489)	(0.219)
Log-Likelihood	-123.72	-553.26
Observations	268	1098

Notes: *10% significance; **5% significance, ***1% significance. Only "yes" votes and "no" votes arte included in the estimation; abstentions are excluded. A random effect probit model with observations clustered within group correlation is reported. The results of a random effect logit model and a fixed effect logit model are highly similar.

Table 4. Determinants of Sanctioning Behavior under *Pun-Low*

Dependent Variable: Punishment points player *i* sends to player *k* at time t: P_{ik}^{t}

	Unrestricted punishment	Pun-Low	
	(source Tan(2008))	Long-Term	Short-Term
β_0 Constant	-5.326***	-3.226***	-2.968***
	(1.975)	(0.630)	(0.803)
β1 Negative Deviation from i's Own Contribution $(\max\{0, c_i - c_k\})$	0.546**	0.752***	1.769***
	(0.259)	(0.130)	(0.671)
β2 Positive Deviation from i's Own Contribution $(\max\{0, c_k - c_i\})$	0.078	1.144***	--
	(0.223)	(0.423)	--
β3 Negative Deviation from Average $(\max\{0, \bar{c} - c_k\})$	0.799**	-0.201	-1.346
	(0.352)	(0.199)	(0.862)
β_4 Positive Deviation from Average $(\max\{0, c_k - \bar{c}\})$	-0.162	--	--
	(0.242)	--	--
β_5 Type i	-0.497	-0.325	0.431
	(1.096)	(0.529)	(0.673)
β_6 Type k	0.787*	-0.268	0.084
	(0.475)	(0.533)	(0.669)
Log-Likelihood	-744.01	-277.957	-75.205
Observations	1080	278	99

Notes: *10% significance; **5% significance, ***1% significance. Only the observations of individuals who could potentially be punished are included. Since the earnings of above-average contributors are not allowed to be reduced, the β_4 coefficient is not included in the *Pun-Low* estimation.

Table 5. Subsequent changes in Contributions of Below-average Contributors as a Function of Punishment Received and Type

Dependent variable: changes of contribution $C_i^{t+1} - C_i^t$

	Unrestricted Punishment (source, Tan (2008))	*Pun-Low*, Short- and Long-term Treatments Pooled
β_1 Punishment Received at Period t	0.289***	0.185
	(0.067)	(0.162)
β_2 Deviation from average	0.169***	0.512**
	(0.056)	(0.205)
β_3 Type i	0.058	0.314
	(0.401)	(0.805)
β_4 Punishment Received * Type i	0.340***	-0.304
	(0.127)	(0.202)
β_0 Constant	0.891***	0.886
	(0.258)	0.603
Adjusted R squared	0.250	0.324
Observations	161	66

Notes: *10% significance; **5% significance, ***1% significance. The model specification procedure is as follows. Firstly, for *Pun-Low* institution, a Chow-breaking point test cannot reject the null hypothesis that the contribution responses of the Long-Term treatment and Short-Term treatment are the statistically equivalent (F (3, 58) = 0.93, p=0.432). Therefore we only report one result by combining two treatments together. We then compare a pooled OLS with robust standard errors; a fixed effect model and a random effect model. The pooled OLS proves to be the best specification through a Language-Multiplier test comparing with the random effect model and an F-test with a fixed effect model. For the unrestricted punishment institution (source Tan (2008)), we run the same regression on players who contribute less than average at period t.

Figure 1: Timing of Activity in Each Treatment

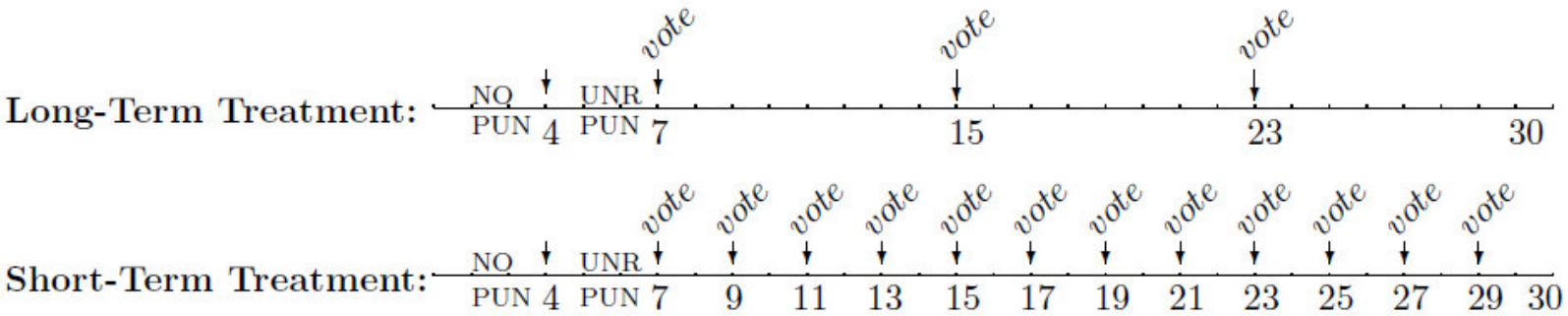
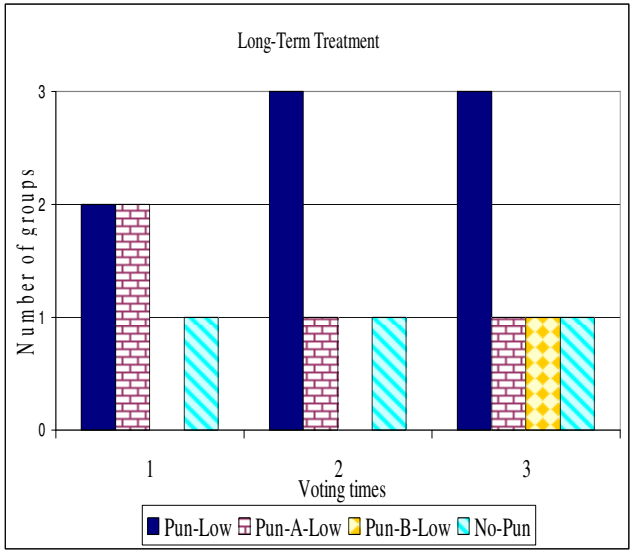


Figure 2: Punishment Systems Enacted, Both Treatments, By Timing of Vote



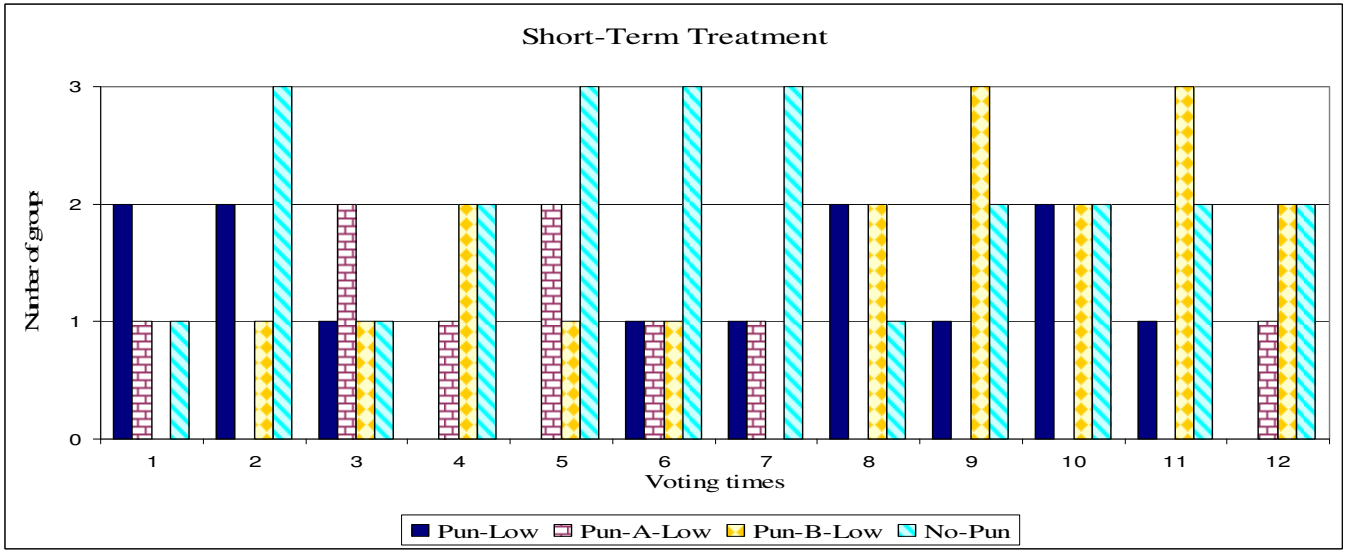
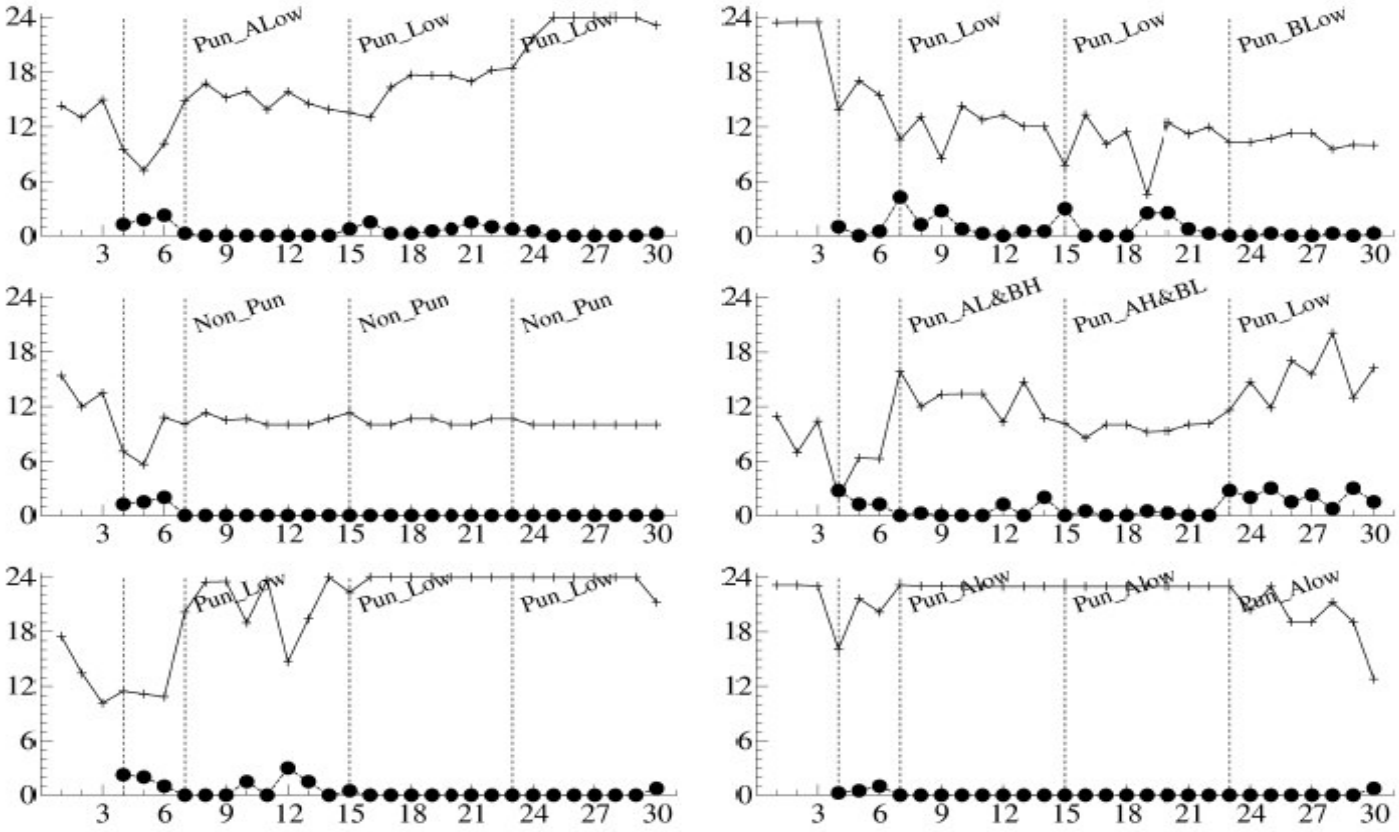
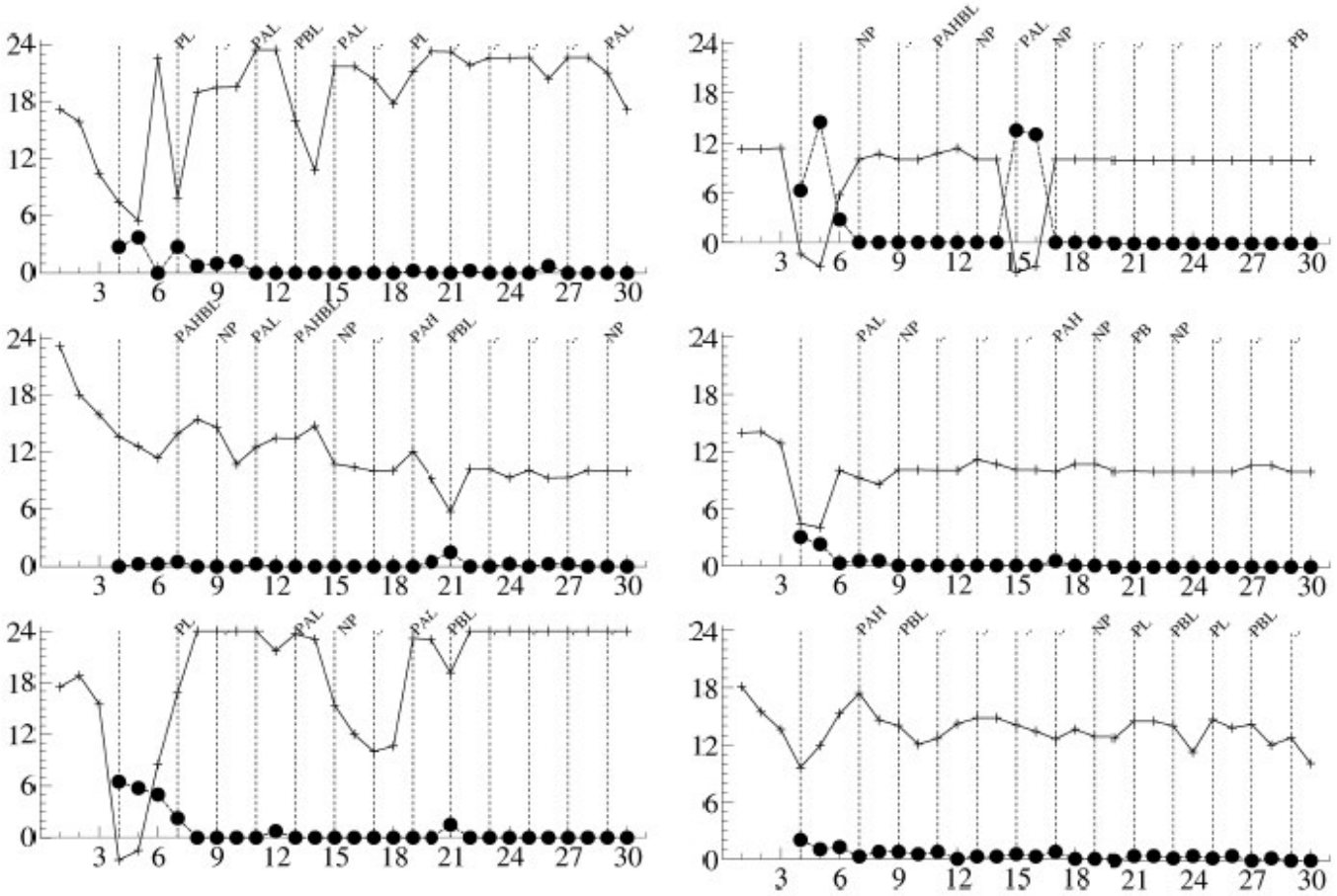


Figure 3a: Earnings and Punishment Levels in the Long-Term Treatment



Notes: Each panel corresponds to one group in the treatment. The horizontal axis designates the number of periods, with the segments indicating the periods in which a specific institution is in effect. The names of the voted institutions are written in the upper part of each segment. The lines with crosses represent the group average earnings, and the lines with dots represent average sanction points.

Figure 3b: Earnings and Punishment Levels in the Short-Term Treatment



Notes (Cont'd): *PL* is short for "allowing punishment of players with below average contributions"; *PAL* is short for "allowing punishment of type A players with below average contributions"; *PBL* is short for "allowing punishment of type B players with below average contributions". *NP* is short for "not allowing any form of punishment". *PB* is short for "allowing punishment of type B players. *PunAHBL* is short for "allowing punishment of type A with above average contributions or type B players with below average contributions". *PAH* is short for "allowing punishment of type A players with above average contributions.

Figure 4a: Voting Patterns in the Long-Term Treatment, Percentage of Players Voting to Punish Each Type and Contribution Level

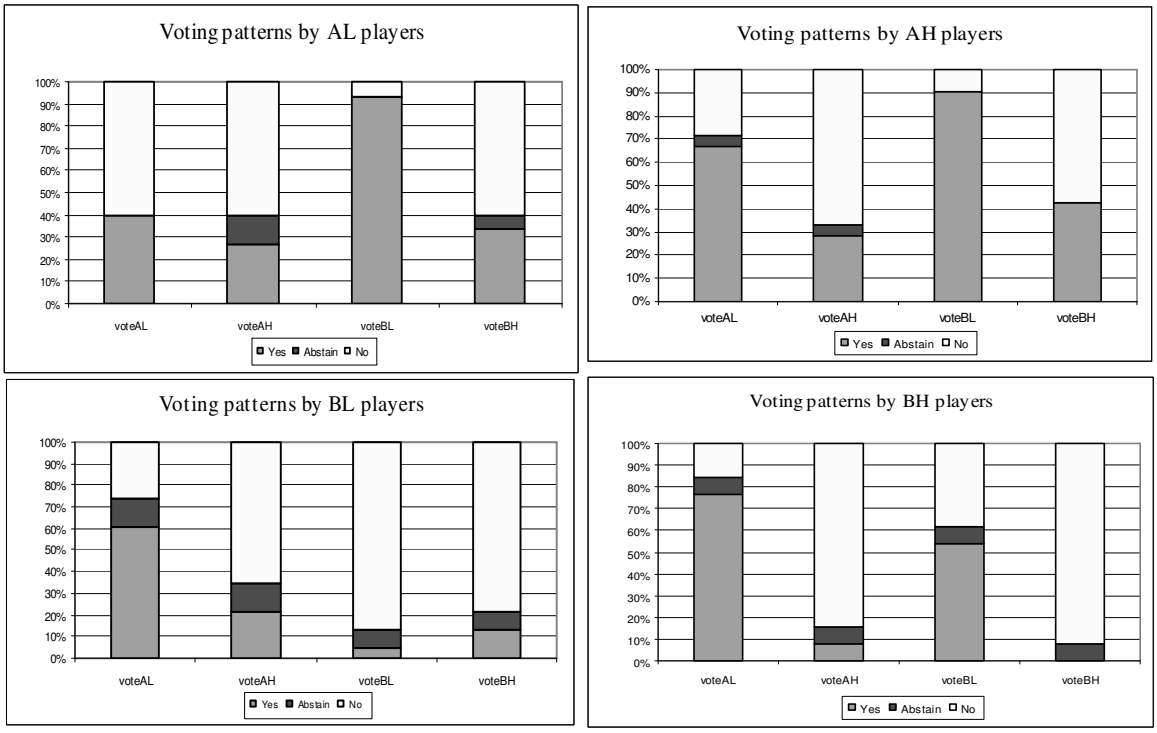
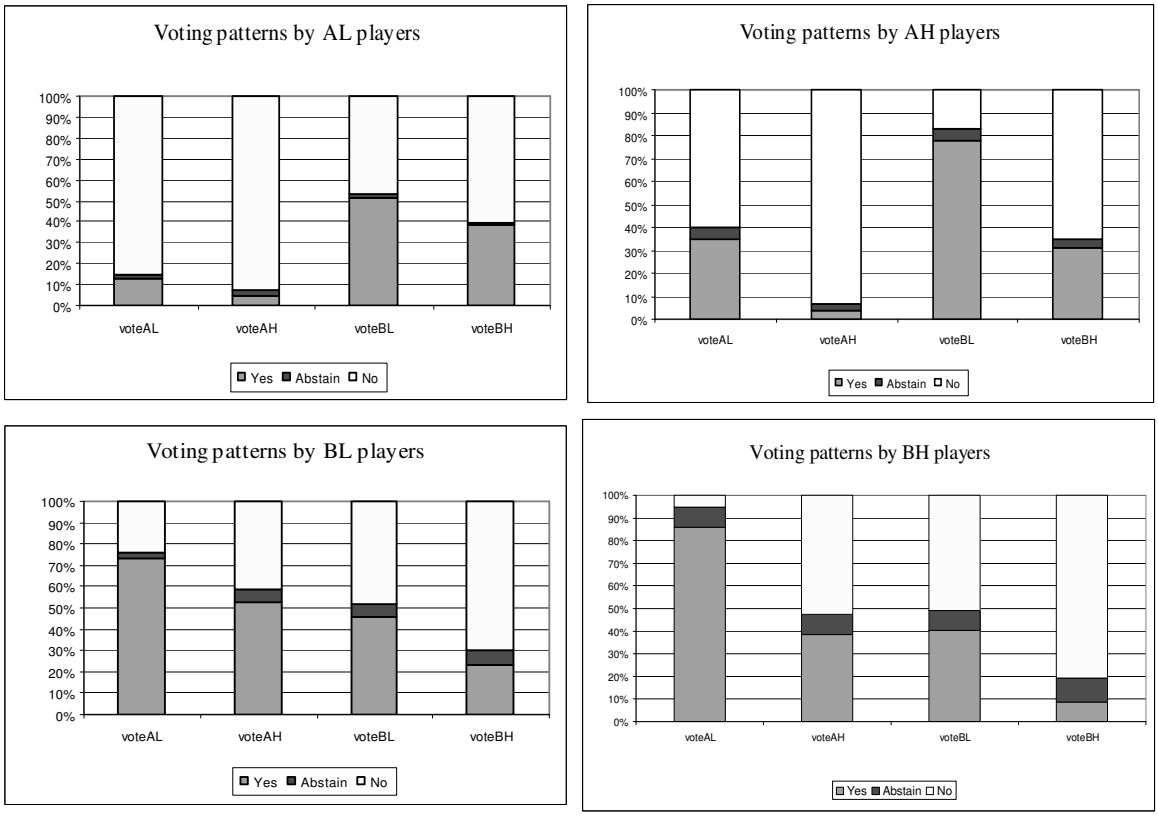


Figure 4b: Voting Patterns in the Short-Term Treatment, Percentage of Players Voting to Punish Each Type and Contribution Level



Appendix: Experiment Instructions

Presented below are the instructions for the Long-Term Treatment. The instructions for the Short-Term Treatment are identical except for the number of rounds an institution is in effect. For instance, in the sentence "After the voting, the decision is in effect for eight rounds. Then you will be asked to vote again for every eight rounds", the number of rounds is changed from "eight" to "two".

The Long-Term Treatment Instruction

EXPERIMENT INSTRUCTIONS (PART I)

You are now taking part in an economic experiment. If you read the following instructions carefully, you can, depending on your decisions and the decisions of others, earn a considerable amount of money. It is therefore very important that you read these instructions with care.

The instructions we have distributed to you are solely for your private information. **It is prohibited to communicate with the other participants during the experiment.** Should you have any questions please ask us. If you violate this rule, we shall have to exclude you from the experiment and from all payments.

During the experiment your entire earnings will be calculated in TOKENS. At the end of the experiment the total number of tokens you have earned will be converted to Euros at the following rate:

25 TOKENS= 1 Euro

Before the experiment starts the computer will assign you with a type. This type can be either "A" or "B". The meaning of type A and type B will be explained in the "Detailed Instructions" below. Your type remains unchanged during the entire experiment.

The experiment is divided into rounds. In each round the participants are divided into groups of four. You will therefore be in a group with 3 other participants. *Note that each group consists of 2 participants with type "A" and 2 participants with type "B".* You will stay in the same group for 30 rounds, but each participant will receive a different identity name, ID 1, 2, 3 or 4 within the group in each round. For example, a participant with ID 1 in this round may not be the same as a participant with ID 1 in another round.

Detailed Instructions:

At the beginning of each round each participant receives 10 tokens. In the following we call this his or her endowment. Your task is to decide how to use your endowment. You have to decide how many of the 10 tokens you want to put into a project and how many of them to keep for yourself. Your choice should be an integer, i.e.numbers such as 0, 1, 2, …10.

Your income consists of two parts:

1) the tokens which you have kept for yourself;
2) the income from the project. This equals 90 percent of the total input of group members with type "A" to the project plus 30 percent of the total input of group members with type "B" to the project (including your own input).

Your income in tokens in each round is therefore:
(10-your input to the project)+
0.9* (total input to the project of members with type "A") +
0.3 * (total input to the project of members with type "B")

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.

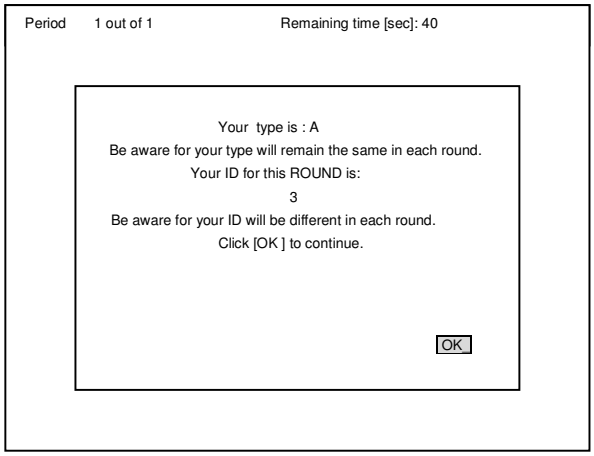
For example, suppose the total sum that all group members put into the project is 30 tokens. Among these 30 tokens, 18 tokens are put by participants with type "A"; and 12 tokens are put by participants with type "B". In this case each member of the group receives an income from the project of 0.9 * 18 + 0.3 * 12 = 19.8 tokens. If the total sum put into the project is 9 tokens, among which 3 tokens are put by participants with type "A"; and 6 tokens are put by participants with type "B", then each member of the group receives an income of 0.9 * 3 + 0. 3 * 6 =4.5 tokens from the project.

For each token that you keep for yourself you earn an income of 1 token. For every token you put into the project instead, the total input rises by one token. If you are type "A", your income from the project would rise by 0.9*1=0.9 tokens. However the income of the other group members would also increase by 0.9 tokens each, so that the total income of the group from the project would rise by 3.6 tokens. If you are type "B", your income from the project would rise by 0.3*1=0.3 tokens. However the income of the other group members would also increase by 0.3 tokens each, so that the total income of the group from the project would rise by 1.2 tokens. Your input to the project therefore also raises the income of the other group members. On the other hand you earn an income for each token put by the other members to the project. For each token put in by a participant with type "A" you earn 0.9*1=0.9 tokens; for each token put in by a participant with type "B" you earn 0.3 * 1=0.3 tokens.

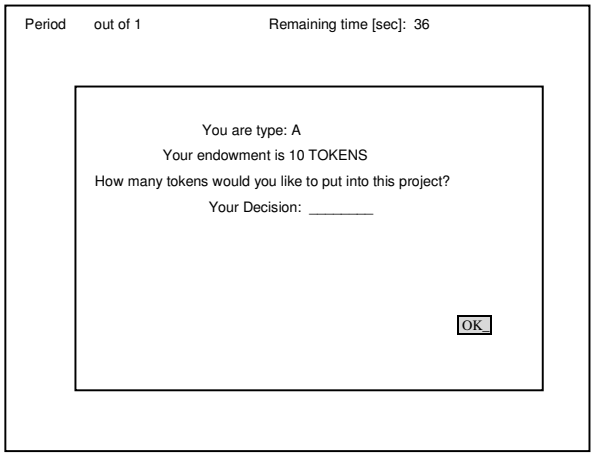
We will now explain how the computer screens look like.

SCREEN 1

This is the screen which shows your type and your ID for this round. The ID will range from 1 to 4. After checking this information, click on OK to proceed.



SCREEN 2

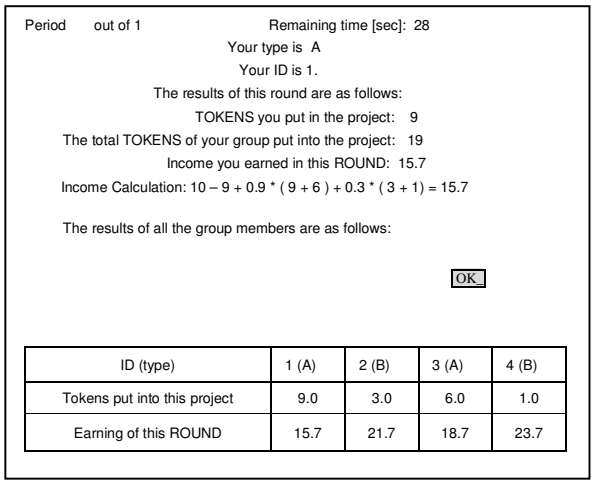


Here you decide on how many tokens you will use for the project in this round. Use the keyboard to type in one of the numbers 0,1 … 10 and confirm your choice by pressing OK.

Warning: Before pressing OK, make sure your choice is correct. You cannot change your decision after you have pressed OK.

After having pressed OK, you will be asked to wait until all experiment participants have done the same. The experiment continues only after all experiment participants have pressed OK. We therefore kindly ask you not to delay your decision too much. After pressing OK, a waiting screen will appear. After all experiment participants have pressed OK, Screen 3 will appear.

SCREEN 3



ID (type)	1 (A)	2 (B)	3 (A)	4 (B)
Tokens put into this project	9.0	3.0	6.0	1.0
Earning of this ROUND	15.7	21.7	18.7	23.7

In the upper part of your screen you find a table with information on your type and your ID, the number of tokens chosen by all participants in your group, the income you earned and its calculation. In the lower part, you find a table with information on tokens put into the project and earnings for all group subjects.

Click on OK if you are done with checking the information.

The experiment will begin with three rounds of play. Each round you begin with a new 10 tokens to allocate, and each round's earnings are independent of the others. After these three rounds, there will be further instructions.

Please raise your hand if you have any questions at this moment.

The experiment now starts with a quiz to make sure that everybody understands how you earn your points. After finishing the quiz, please raise your hand for answer checking. After all participants answered all the questions correctly, the experiment will begin.

Quiz

To check your understanding of the experiment, please answer the following questions:

About the experiment setting (Yes/ No):

a) If you are assigned with type "A", does your type change in different rounds? Yes/No
b) Are there 2 participants with type "A" and 2 participants with type "B" in a group? Yes/No
c) Are you in the same group in different rounds? Yes/No
d) Is a person with ID1 in Round 2 definitely the same with a person with ID1 in Round 3? Yes/No

2. You are assigned with type "A". Suppose each group member has an endowment of 10 tokens. Nobody (including yourself) put in any tokens to the project. How high is:

a) Your income for the period? _____
b) The income of the other group members for the period? _____

3. You are assigned with type "B". Suppose each group member has an endowment of 10 tokens. You put in 10 tokens to the project. Besides you, a participant with type "A" puts in 3 tokens into the project; another participant with type "A" puts in 6 tokens into the project; and the third participant with type "B" puts in 2 tokens into the project . What is:

a) Your income for the period? _____
b) The income of the group member which is type A and put 3 tokens into the project for the period? _____

4. You are assigned with type "A". Suppose each group member has an endowment of 10 tokens. Besides you, a participant with type "A" puts in 4 tokens into the project; another participant with type "B" puts in 5 tokens into the project; and the third participant with type "B" puts in 3 tokens into the project .

a) What is your income if you put in 0 tokens to the project? _____
b) What is your income if you put in 5 tokens to the project? _____

EXPERIMENT INSTRUCTIONS (PART II)

After this break for instructions, you and the same three members of your group will be interacting for another three rounds. As with the three rounds just completed, each of these rounds begins with a decision on assigning ten tokens to a group account or to a personal account. This time, however, each round also includes a second stage of decision-making.

At the beginning of the second stage, a screen will show you how much each of your group members puts into the project. In this stage you have the opportunity to register your disapproval of each other group member's decision by assigning points to the other three participants in your group.

You must decide how many points to send to each of the other three group members. If you do not wish to change the income of a specific group member then you must enter 0. Every point you

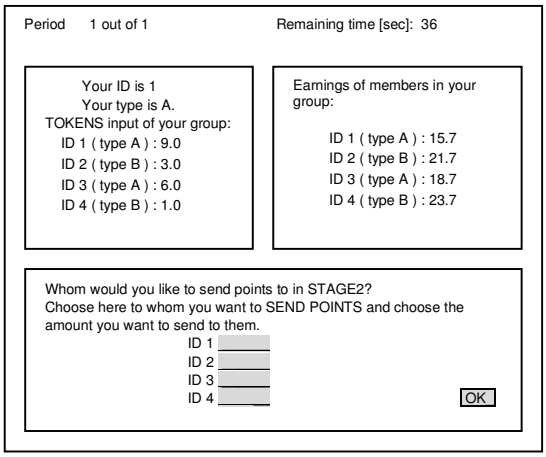
send will reduce your earnings by 1 token AND reduce the earnings of the participant receiving it by 2 tokens.

Whether and by how much a person's income from the first stage is reduced depends on the total of the points he/ she received from all of the other members of his/her group. If somebody received a total of 3 points (from all other group members in this round), his or her income would be reduced by 6 tokens. If somebody received a total of 4 points, his or her income would be reduced by 8 tokens. The other group members can also assign points to you if they wish to.

Your total income from this round (two stages together) is therefore calculated as follows:
= (income from the 1st stage –
points assigned to other participants) –
2 * total points received by three other participants

We will now explain how the computer screens look like. Note that Screen 1 to Screen 3 are exactly the same as the first three rounds.

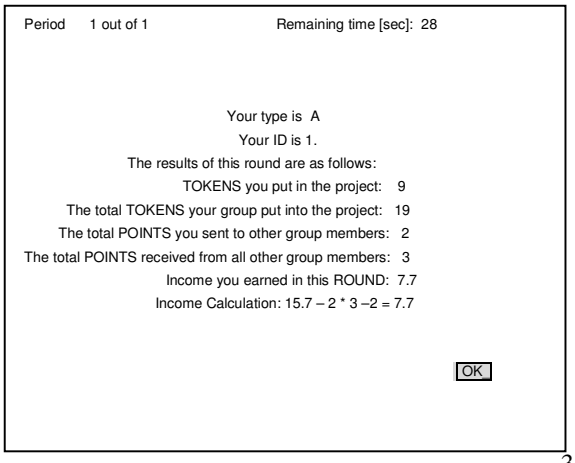
SCREEN 4



In the upper part of this screen you find a table with information on the type of each participant, the number of tokens chosen for the project by each subject in stage 1 of this round and the number of tokens earned in Stage 1.

In the lower part of this screen, you are asked to make a decision on how many points you would like to assign to reduce earnings of each of the three other participants. Your choice must be integer, i.e. numbers like 0,1,2,…10. Select OK, when you are ready to continue. A waiting screen will appear. The experiment continues only after all participants have pressed OK, and therefore we kindly ask you not to delay your decision too much.

SCREEN 5



In this screen you will be provided with information about this round. You will be shown the tokens you and all participants put into the project, the total number of points you received and assigned to others, the income of this round and its calculation.

Click on OK if you are done with checking

the information.

The experiment will continue with another three rounds of play. After these three rounds, there will be further instructions.

Please raise your hand if you have any question at the moment.

The experiment now starts with a quiz to make sure that everybody understands how you earn your points. After finishing the quiz, please raise your hand for answer checking. After all participants answered all the questions correctly, the experiment will continue.

1. Suppose in the second stage of a period, you distribute the following amounts of monetary points to the other three group members: 9, 5, and 0.What is the total cost of the tokens you distribute? _____

2. What are your costs if you send a total of 0 tokens? _____

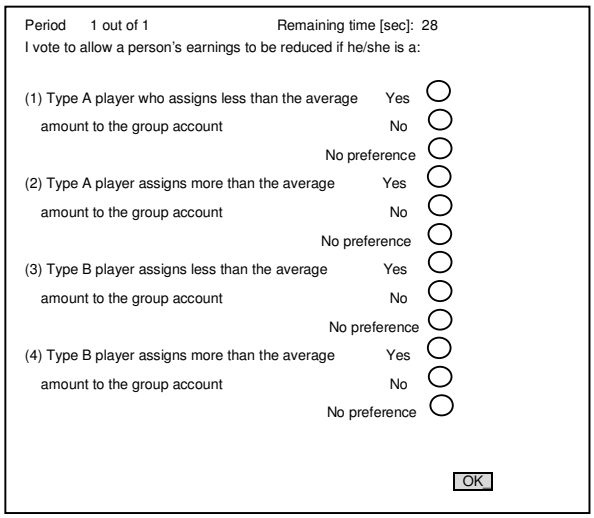
3. By how many tokens will your income from the first stage be reduced, when you receive a total of 0 monetary points from the other group members? _____

4. By how many tokens will your income from the first stage be reduced, when you receive a total of 5 tokens from the other group members? _____

EXPERIMENT INSTRUCTIONS (PART III)

In the remaining parts of the experiment, you will play for three sets of eight rounds each in the same group of four subjects. Before this part begins, each group will decide, by voting, whether to permit subjects to reduce one another's earnings after learning of their assignments to the group account. It will be possible to allow reductions of a type A and/or type B subject who assigns more than the average to the group account, and/or of type A or type B subjects who assign less than the average to the group account. Once the decision has been made by your group, it will be in force for the next eight rounds of the experiment.

We will now explain how the computer screens look like.



SCREEN 6

In this screen you are asked to answer "Yes", "No", or "No preference" to four questions by clicking the box to the right of each of the three choices. For each question, if the number of "Yes" vote in your group exceeds the

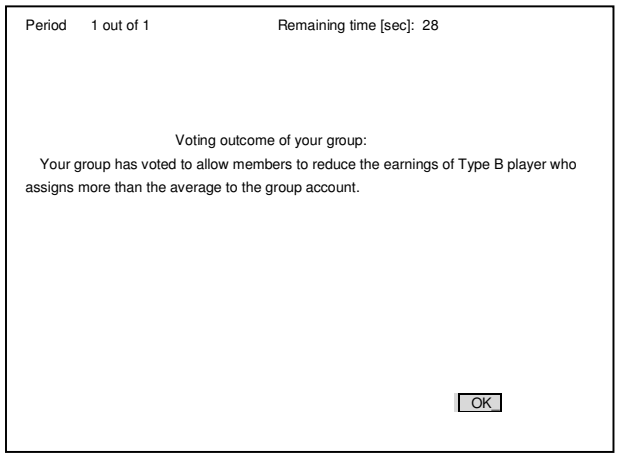
number of “No” vote, the reductions in question will be allowed; otherwise they will not. A “No preference” vote does not count towards the voting outcome.

Click on OK if you are done with answering the questions.

Warning: Before pressing OK, make sure your choice is correct. You cannot change your decision after you have pressed OK.

After having pressed OK, you will be asked to wait until all experiment participants have done the same. The experiment continues only after all experiment participants pressed OK. We therefore kindly ask you not to delay your decision too much. After pressing OK, a waiting screen will appear. After all experiment participants have pressed OK, Screen 7 will appear.

SCREEN 7



In this screen you will be informed of the outcome under which your group will operate for the next eight rounds. The possible messages are listed in the appendix below. Note that only one of these messages will show up on the screen.

Click on OK if you are done with checking the information.

After the voting, the decision is in effect for eight rounds. Then you will be asked to vote again for every eight rounds. During the reduction stage of each round, if the earnings of certain group members are voted “allow to be reduced” because of the rules decided by your group, you can decide whether to send points to the group members meeting the description. On the other hand, the reduction boxes for any individuals whom your group has decided cannot have their earnings reduced will automatically appear with zeros inside, which cannot be changed.

It is important that you fully understand the voting process before we continue. Please raise your hand if you have any questions at this moment. If not, the experiment will continue.

Possible messages regarding to voting outcomes:

(1) Your group has voted not to allow group members to reduce one another's earnings

(2) Your group has voted to allow members to reduce the earnings of any other group member.

(3) Your group has voted to allow members to reduce the earnings of Type B player who assigns less than the average to the group account.

(4) Your group has voted to allow members to reduce the earnings of Type B player who assigns more than the average to the group account.

(5) Your group has voted to allow members to reduce the earnings of Type A player who assigns less than the average to the group account.

(6) Your group has voted to allow members to reduce the earnings of Type A player who assigns more than the average to the group account.

(7) Your group has voted to allow group members to reduce the earnings of players assigning less than average to the group account regardless of their types.

(8) Your group has voted to allow group members to reduce the earnings of players assigning more than average to the group account regardless of their types.

(9) Your group has voted to allow group members to reduce the earnings of Type B players.

(10)Your group has voted to allow group members to reduce the earnings of Type B player assigning less than group average AND Type A player assigning more than average to the group account.

(11)Your group has voted to allow group members to reduce the earnings of Type B player assigning more than group average AND Type A player assigning less than average to the group account.

(12) Your group has voted to allow group members to reduce the earnings of Type A players.

(13) Your group has voted to allow group members to reduce the earnings of Type B players AND Type A players assigning less than average to the group account.

(14) Your group has voted to allow group members to reduce the earnings of Type B players AND Type A players assigning more than average to the group account.

(15) Your group has voted to allow group members to reduce the earnings of Type A players AND Type B players assigning more than average to the group account.

(16) Your group has voted to allow group members to reduce the earnings of Type A players AND Type B players assigning less than average to the group account.