No. 67 august 2011

Lisa Bruttel
Tim Friehe

Path dependence in public-good games

Research Paper Series
Thurgau Institute of Economics and Department of Economics
at the University of Konstanz

Member of
thurgauwissenschaft
www.thurgau-wissenschaft.ch

THURGAU INSTITUTE
OF ECONOMICS
at the University of Konstanz

Path dependence in public-good games*

Lisa Bruttel† Tim Friehe‡

August 5, 2011

Abstract

This paper presents experimental evidence that contributions to a public good can be path-dependent for a limited time span. We study a repeated linear public-good game with punishment opportunities. Our data shows that subjects who had experienced a higher marginal return on public-good contributions in rounds 1-10 contributed more to the public good in rounds 11 and 12, even though they faced the same marginal return as the control group in these later rounds. In contrast, differences in contributions were not significant when comparing subjects bearing the same current costs of punishment points, but having had different costs in the past.

Keywords: public-good game, team, punishment, path dependence, experiment

JEL-Classification: C91, H26, K42

*We would like to thank Kate Bendrick, Gerald Eisenkopf, Urs Fischbacher, Katrin Schmelz, Simeon Schudy, Pascal Sulser, Verena Utikal, Irenaeus Wolff, and Natalie Zimmer for very helpful suggestions.

†University of Konstanz, Department of Economics, Box 131, 78457 Konstanz, Germany. E-mail: lisa.bruttel@uni-konstanz.de. Phone: +49-7531-88-3214. Fax: +49-7531-88-2145.

‡University of Konstanz, Department of Economics, Box 136, 78457 Konstanz, Germany. E-mail: tim.friehe@uni-konstanz.de. Phone: +49-7531-88-2534. Fax: +49-7531-88-4135.

1 Introduction

This paper addresses whether the initial and temporary use of high-powered incentives may help to lastingly promote cooperation in groups. In a variety of circumstances, group members either fail to coordinate or coordinate on a non-cooperative, often Pareto-dominated outcome. Imagine, for example, that a firm forms a team for which success is determined by the aggregate of unverifiable individual effort levels. In such a set-up, coordination failure or inefficient coordination might easily arise because team members bear the full costs of high effort but would only reap full benefits therefrom if other team members also invested high effort. The firm may therefore attempt to induce coordination on high-effort contributions (i.e., the cooperative outcome) from the start by implementing high-powered incentives over some initial time period. The firm could put positive incentives into place by leveraging performance pay in relation to team success, or it might implement negative incentives by facilitating peer monitoring and reporting, for example. If high-powered incentives prove successful in obtaining cooperation among team members, their initial use may carry over to periods in which incentives are scaled back. In such a case, incentives would help team members coordinate on the cooperative outcome while they are in effect, and would additionally produce a cooperative spillover to team interactions in periods without effective incentives. However, the limited use of high-powered incentives may also backfire, in the sense that team performance might become worse after the phase-out than it would have been without the initial use of the incentives. This could arise because team members perceive the reduction in rewards as unfair or because they lose intrinsic motivation, among other possible reasons. Finally, the individual effort provision might be unrelated to the outcomes in earlier periods, whether coordination failure or success, because past incentives are no longer relevant for current decisions. In order to address this theme of the path dependence of individual contributions, we study a repeated linear public-good game with punishment opportunities to simulate contexts such as the team scenario described above.

This paper examines whether the levels of decision parameters relevant in rounds 1-10 of the repeated linear public-good game continue to influence individual contributions in rounds 11-20, despite being no longer decisive for an individual's decision problem. We investigate the path dependence of individual contributions with respect to both positive and negative incentives in a fixed-partners design. Specifically, we consider rounds 11-20 and contrast the behavior of subjects who had a high marginal return on contributions to the public good in rounds 1-10

to the behavior of individuals who had a low marginal return on public-good contributions. Similarly, we contrast the behavior of participants who experienced low punishment costs in the first ten rounds to the behavior of subjects with high punishment costs. Principally, both low punishment costs and a high marginal return on public-good contributions are conducive to high individual public-good contributions, although the effect of punishment costs is indirect. A lower level of punishment costs makes individuals more willing to punish low contributions, all else held equal; knowing this, individuals are better off when they contribute more.[1]

The experimental results suggest that there can be path dependence of individual contributions to the public good. However, significant differences in average contributions were found only in rounds 11 and 12; i.e., there is no long-term persistence of the effect. Moreover, we obtained significant differences in average contributions only when contrasting the behavior of participants who initially had a high marginal return on public-good contributions to the behavior of individuals with a low marginal return on public-good contributions throughout all rounds. Specifically, individuals who had a high marginal return on public-good contributions in rounds 1-10 and a low one in rounds 11-20 contributed more to the public good in rounds 11 and 12 than individuals who had experienced a low marginal return in all 20 rounds. In contrast, the contributions of individuals with low punishment costs in rounds 1-10 and high costs in rounds 11-20 were not different at a conventional level of statistical significance from contributions of individuals who had high punishment costs in all 20 rounds.

Our analysis questions the long-term effectiveness of the initial and temporary usage of high-powered incentives when it comes to cooperation in groups. In addition, the present study suggests that some policy instruments are better suited to produce (at least transitory) spillovers in the level of contributions than others. The temporary use of high levels of the marginal return on the public good produced significant differences in rounds 11 and 12, whereas the use of low punishment costs failed to do so.

The structure of the article is as follows: In Section 2, we relate our contribution to the literature; we present the experimental design and procedures in Section 3. The experimental results are discussed in Section 4, and Section 5 discusses our findings and concludes the paper.

[1]The hypothesis underlying the mechanism via punishment costs has empirical support. Carpenter (2007) has established that subjects in voluntary contribution mechanisms indeed demand more punishment when the price decreases, and Nikiforakis and Normann (2008) have shown that contributions increase monotonically in punishment effectiveness. See also the very recent survey by Chaudhuri (2011).

2 Connections to the literature

Our paper explores the potential path dependence of individual contributions in a public-good game with punishment opportunities.[2] A number of papers have studied related issues, with Bruttel and Friehe (2010), Bohnet and Huck (2004), Brandts and Cooper (2006), and Hamman et al. (2007) particularly relevant to the current paper.

Bruttel and Friehe (2010) investigate path dependence in a tax-evasion setting that is repeated for 20 rounds but not characterized by strategic interdependence, finding that the penalty for tax evasion valid in rounds 1-10 persistently influences income declarations in rounds 11-20. While the present study uses a similar format to that of Bruttel and Friehe (2010), strategic interdependence and uncertainty of participants are vital elements of our setup. Bohnet and Huck (2004) study a binary-choice trust game involving two subjects; they find that individuals who took part in a fixed-partner setup in rounds 1-10 trust more and are more trustworthy in rounds 11-20 when interacting with random strangers on a one-shot basis than subjects who were in the stranger condition in all rounds. The present study challenges their results by adding to both the complexity of the strategy space and the degree of strategic uncertainty. However, like Bohnet and Huck (2004), we are interested in whether the initial provision of circumstances conducive to cooperation leaves imprints on decisions at later points in time. In contrast to the two studies cited above (and to the experiment in the present paper), which consider the effect of the initial provision of incentives on cooperation, Brandts and Cooper (2006) and Hamman et al. (2007) are interested in persistent effects when teams that are failing to coordinate are temporarily confronted with stronger incentives. Specifically, Brandts and Cooper (2006) study 3×10 rounds of a minimum-effort game with fixed groups of four subjects and in which most groups suffer from coordination failure after the first ten rounds, exploring the consequences of various changes to the bonus rate in rounds 11-30. Inter alia, the authors contrast the behavior of subjects who face a low bonus rate in rounds 1-10, a high bonus rate in rounds 11-20, and a low bonus rate in rounds 21-30 to pooled behavior in rounds 1-10; they find that reducing the bonus rate did not bring the level of minimum effort back to the level measured in rounds 1-10. This is interpreted as indicating the path dependence of effort provision; however, their design lacks a control group with low bonus rates

[2]The literature concerning individual contributions in public-goods games is extensive. A recent survey is provided by Chaudhuri (2011).

in all rounds. This gap was filled by Hamman et al. (2007) who study a similar minimum-effort game with five different effort levels, fixed groups, and 30 rounds. The authors depart from the design investigated by Brandts and Cooper (2006) by (i) studying positive and negative incentives, and (ii) restricting bonus payments (penalties) to a fixed magnitude instead of setting them to be proportional to the minimal effort level, and applying them only in outcomes with sufficiently high (low) effort levels. They present in particular two related findings: (i) the long-term persistent benefits of incentives are smaller than those presented by Brandts and Cooper (2006), and (ii) positive and negative incentives are effective to a comparable degree. Like Hamman et al. (2007), we consider positive and negative incentives. However, in our design, negative repercussions of small individual contributions do not arise automatically from the compensation scheme but are endogenous, because the subjects themselves determine the utilization of punishment opportunities. In contrast to Hamman et al. (2007), we will show that there are marked differences between what we consider to be positive and negative incentives in our setup. In contrast to the last three studies cited above (i.e., Bohnet and Huck 2004, Brandts and Cooper 2006, and Hamman et al. 2007), we examine a public-good game with punishment opportunities in which participants choose from a rich set of alternatives. We believe that the increase in the number of alternatives makes the decision problem both more realistically complex and less of a binary choice, so that our study provides a more conservative test for potential path dependence.

The above studies examine the possibility of effecting desirable long-term effects by temporarily changing the rules to a game. There is also evidence for the detrimental long-term consequences of explicit incentives after their abolishment. Gneezy and Rustichini (2000) show that the introduction of fines for tardy parents at a child-care center increases tardiness even after fines are no longer used, and Gächter et al. (2008) establish that the use of explicit incentives in an employment relationship can crowd out voluntary cooperation even after the reliance on explicit incentives is discontinued. Finally, instead of focusing on how behavior changes over time in a given game (i.e., exploring whether there are spillovers across periods of time), it is also possible to consider spillovers across games. For example, Cason et al. (2009) show that cooperation in a minimum-effort game can be achieved through a behavioral spill-over from a median-effort game when the former is played after the latter. Similar results have been found by Ahn et al. (2001) and Knez and Camerer (2000), among others.

3 Experimental design and procedures

3.1 Design

The experiment was based on a repeated linear public-good game with punishment opportunities, played by fixed groups of four subjects.[3] There were two stages to each round and two sequences of ten rounds each. In *Stage 1*, participants decided (simultaneously and without the possibility of communicating) how to split their endowments of 20 points between a private account and a public account. The contribution to the public good by subject i is denoted c_i. The contributions to the public good by all four group members were collected in the public account, and the total was multiplied by $\alpha > 1$ before reallocating the returns equally to group members. This implies that participant i's payoffs from the first stage are given by

$$\pi_i = 20 - c_i + \alpha/4 \sum_{j=1}^{4} c_j \tag{1}$$

At the end of Stage 1, participants were informed about π_i and the sum of contributions to the public good. At the beginning of *Stage 2*, group members obtained information about the other group members' first-stage income and their contributions to the public good. Individual i could then assign punishment points to be applied to subject j, denoted p_{ij}, at a cost of $\beta > 0$. One punishment point lowered subject j's income by one, i.e., it had a punishment effectiveness equal to one. This implies that payoffs at the end of Stage 2 and thus from any given round can be stated as

$$\Pi_i = \pi_i - \sum_{j \neq i} p_{ji} - \beta \sum_{j \neq i} p_{ij} \tag{2}$$

The impact of punishment points imposed by others on subject i, i.e. the effect of $\sum_{j \neq i} p_{ji}$, was restricted to the degree that $\pi_i - \sum_{j \neq i} p_{ji} \geq 0$. Nonetheless, it was possible that $\Pi_i < 0$.[4] The identification of a given participant to the other members of the group on the screen (group members 2-4) was subject to a random draw in every round in order to avoid spillovers across rounds.[5] At the end of Stage 2, participant i obtained a summary detailing the contribution to

[3] The design of individual rounds is similar to Fehr and Gächter (2000). A translation of the instructions is presented in an appendix to our paper.

[4] Participants were given an initial endowment to compensate for potential losses in initial rounds.

[5] The instructions in the appendix include a screenshot showing how this was implemented.

the public good c_i, the sum of individual contributions, first-stage income π_i, both the number and the costs of punishment points allocated by this participant to other group members, the number of punishment points received, and Π_i. The payoff structure was the same for all participants, which in turn was common knowledge.

Before the start of the experiment, all participants were informed that a parameter of the decision-problem might change after the first ten rounds, but they did not know which parameter would be altered, or in which direction. In order to ensure that participants were indeed aware of the change, they were asked to answer a question about the current levels of the two decision parameters at the beginning of round 11. In our study, the marginal return on public-good contributions α and (marginal) punishment costs β are the treatment variables of interest, and we distinguish between a low and a high level for each. Both α and β stayed constant for rounds 1-10 and 11-20 in all treatments. Thus, we may denote α_{1-10} (β_{1-10}) as the marginal return (punishment costs) applicable in the first ten rounds. The levels applied in rounds 11-20 were the same for all three treatments and represent the low marginal return on public-good contributions and the high level of punishment costs, i.e., $\alpha_{11-20} = 1.2$ and $\beta_{11-20} = 0.5$. The experimental design consisted of 3 treatments (see Table 1).

Treatment	α_{1-10}	β_{1-10}	α_{11-20}	β_{11-20}	Subjects	Observations
Base	1.2	0.5	1.2	0.5	56	14
HighReturn	2	0.5	1.2	0.5	56	14
LowCost	1.2	0.25	1.2	0.5	52	13

Table 1: Treatment parameters, number of subjects, and independent observations

In our setup, game-theoretic predictions based on the concept of subgame-perfectness in view of monetary aspects predict zero punishment and zero contributions in all treatments. Accordingly, there is no straightforward way to calibrate the two decision parameters in a way that they would be equivalent regarding equilibrium behavior in theory. Regarding the precise levels of our treatment variables, the two levels of $\alpha/4$ were chosen to be the neighboring levels of the commonly used level 0.4 (see, e.g., Fehr and Gächter 2000, Kube and Traxler forthcoming, Nikiforakis and Normann 2008). As a result, $\alpha = 2$ is relatively high and $\alpha = 1.2$ relatively low when compared to the standard level of this parameter. It must be noted that this fact (i.e., that $0.3 < 0.4$) will tend to tilt behavior in our experiment against cooperation in rounds 11-20, all else held equal. The levels of punishment costs were chosen in view of the findings by

Nikiforakis and Normann (2008). In their experiment, which was based on a repeated linear public-good game with punishment opportunities, a punishment effectiveness of 2 (comparable to punishment costs of 0.5) yielded constant average contributions, whereas a punishment effectiveness of 4 (comparable to punishment costs of 0.25) yielded strictly increasing average contributions, averaging at 17.93 in rounds 6-10.

Based on the findings of other analyses of individual contributions in public-good games with punishment opportunities (see, e.g., Chaudhuri 2011), we expect that average contributions in HighReturn and LowCosts will be higher than in Base in rounds 1-10. However, we do not have a clear-cut prediction about path dependence, that is, the differences between the average contributions to the public good in rounds 11-20 in our three treatments. Although the study by Brandts and Cooper (2006) suggests that path dependence may occur, this need not extend to the public-good setting we scrutinize. Indeed, the relatively small variation to the Brandts/Cooper-setup introduced by Hamman et al. (2007) has already cast doubt on the generality of path dependence. Moreover, if the removal of the incentive to punish by the introduction of high punishment costs is perceived as the removal of a sanctioning device, the introduction of low punishment costs in rounds 1-10 may actually backfire in rounds 11-20, as suggested by Gneezy and Rustichini (2000). In other words, even if we were to find path dependence, it is not necessarily clear whether subjects would "underreact" or "overreact" relative to the decision-making behavior in our control group.

3.2 Procedures

The experiment was computerized using z-Tree (Fischbacher 2007). A total of 164 students from various disciplines took part in at most one of the three treatments. They were recruited via ORSEE (Greiner 2004). The experiment took place in the *Lakelab*, the laboratory for experimental economics at the University of Konstanz, in November and December 2010. Sessions lasted less than 90 minutes. The experimental currency was points, with 30 points converted into 1 euro after the experiment. On average, participants earned 17.85 euros in the experiment. The protocol before the start of the experiment was as follows: Subjects first received written instructions for participating in the experiment, and then had to answer control questions displayed on their computer screens. The experiment started only after all subjects had

answered the control questions correctly. At the end of the session, the participants were asked to complete a brief questionnaire. We elicited beliefs about the expected average contribution for the upcoming ten rounds at the beginning of rounds 1 and round 11 without incentivizing them. According to Gächter and Renner (2010), eliciting beliefs in a non-incentivized manner has the advantage that it does not systematically distort contribution levels.

4 Results

In this section, we discuss the experimental data. We turn first to our main interest; the pattern of contributions over time. Subsequently, we briefly discuss the assigned punishment points and the participants' beliefs about the expected average contribution for the upcoming ten rounds.

4.1 Contributions

Figure 1 shows average individual contributions to the public good in our three treatments across all 20 rounds. In line with other studies (see, e.g., Chaudhuri 2011), the average contribution in the baseline treatment was about 50 percent of the endowment. Average contributions in rounds 1-10 were very stable over time. In particular, there was no decline in round 9 or 10. In contrast to other studies, there was no restart effect in BASE, meaning that group members' allocations to the public good in round 11 were not higher than in round 10 (see, e.g., Hamman et al. 2007, Sefton et al. 2007).[6]

From Figure 1, both the high marginal return on public-good contributions and the low level of the punishment costs increased contribution levels in rounds 1-10. The difference is statistically significant only when we compare contributions in HIGHRETURN and in BASE ($p = 0.0204$).[7] With regard to the difference in average contributions in rounds 1-10 between LOWCOST and BASE, we obtain $p = 0.2758$.

[6] It seems that this effect is not contingent on the regulation of information. In Sefton et al. (2007), subjects were informed of the whole sequence of decision periods at the beginning of the experiment. In Hamman et al. (2007), participants received instructions for later rounds only after the experiment had already started.

[7] Unless noted otherwise, we report the significance levels of two-sided Wilcoxon rank sum test in parentheses.

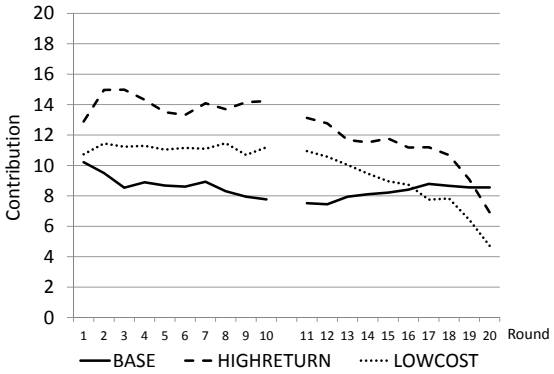


Figure 1: Average public-good contributions

Result 1 *(a) Individuals with high marginal returns to public-good contributions in rounds 1-10 ($\alpha_{1-10} = 2$) and low marginal returns in rounds 11-20 ($\alpha_{11-20} = 1.2$) contributed more to the public good in rounds 1-10 than subjects with low marginal returns in rounds 1-20 ($\alpha_{1-10} = \alpha_{11-20} = 1.2$). (b) Individuals with low punishment costs in rounds 1-10 ($\beta_{1-10} = 0.25$) and high punishment costs in rounds 11-20 ($\beta_{11-20} = 0.5$) did not contribute significantly more to the public good in rounds 1-10 than subjects with high punishment costs in rounds 1-20 ($\beta_{1-10} = \beta_{11-20} = 0.5$).*

Next, we turn to the path dependence of contributions to the public good. Table 2 reports the results of round-wise Wilcoxon rank sum tests, examining whether contributions in the HighReturn (LowCost) condition were equal to those in Base. The test results show that there was indeed transitory path dependence in HighReturn. Contributions in this treatment were higher than those in Base in rounds 11 and 12. However, experiencing high marginal return on contributions to the public good in rounds 1-10 did not yield persistent differences. From round 13 onwards, contributions in Base and HighReturn were statistically indistinguishable. Comparing contributions in LowCost and in Base, we find no significant differences from round 11 onwards. This does not come as a surprise, given that the difference between contribution levels in rounds 1-10 was not statistically significant, either. In summary, our experimental data affirms the possibility of path dependence, but casts a doubt on the long-term persistence of such effects.

	p-value	
Round	BASE vs. HIGHRETURN	BASE vs. LOWCOST
11	0.0204	0.1970
12	0.0512	0.2340
13	0.1556	0.4354
14	0.2420	0.7872
15	0.2076	0.9044
16	0.3370	0.9840
17	0.3576	0.7188
18	0.3472	0.9442
19	0.8336	0.5754
20	0.8180	0.1970

Table 2: Comparison of contributions: p-values from two-sided Wilcoxon rank sum test.

Result 2 *In* HIGHRETURN*, there was transitory path dependence of the level of contributions to the public good. Individuals with high marginal returns to public-good contributions in rounds 1-10 ($\alpha_{1-10} = 2$) and low marginal returns in rounds 11-20 ($\alpha_{11-20} = 1.2$) contributed more to the public good in rounds 11 and 12 than subjects with low marginal returns in rounds 1-20 ($\alpha_{1-10} = \alpha_{11-20} = 1.2$).*

In Figure 1, average contributions in rounds 11-20 followed a downward trend in the LOWCOST and HIGHRETURN conditions. In contrast, average contributions were more or less constant in BASE. However, patterns of average contributions might not tell the whole story. The average in BASE hides a wide variance in group contribution levels.[8] In contrast to what held true in the other two treatments, contribution levels of three groups in BASE actually increased over time. Figure 2 shows the results from clustering groups according to the change in the level of average contributions in rounds 16-20 when compared to the level in rounds 11-15.[9] In the LOWCOST and HIGHRETURN conditions, the contributions of the vast majority of groups declined in the fourth quarter of the experiment. In the BASE treatment, the con-

[8]The pattern of contributions over time is very similar to that from a comparable treatment in Nikiforakis and Normann (2008). For a punishment effectiveness of 2, they also report average contributions which are roughly constant but result from very diverse group contribution patterns.

[9]We denote groups that experienced a change of not more than one unit as constant.

tributions of most groups either remained relatively stable or even showed an increase. The differences illustrated in Figure 2 are statistically highly significant in a two-sided 2x3 Fisher exact test (p-values: LowCost vs. Base = 0.0021, HighReturn vs. Base = 0.0094).

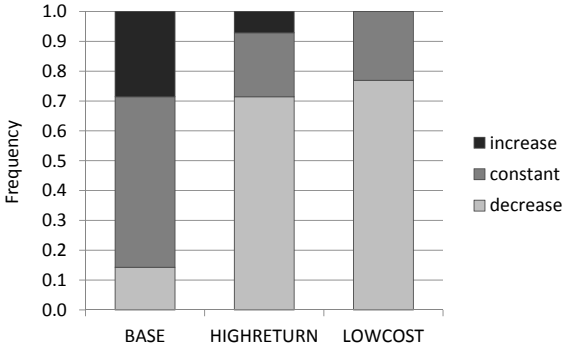


Figure 2: Direction of change of the average group contribution in rounds 16-20 compared to rounds 11-15.

Result 3 *In* HighReturn *and* LowCost*, there was a downward trend in average contributions to the public good after the reduction in the marginal returns to contributions to the public good and after the increase in the marginal costs of punishment, respectively (i.e., in rounds 11-20).*

As a result of the findings above, the data represented in Figures 1 and 2 does not help to rule out the possibility that the initial and temporary use of high-powered incentives in our design may backfire in the long-run by yielding a lower contribution level than that which would have been obtained without the use of such incentives. Neither Brandts and Cooper (2006) nor Hamman et al. (2007) have reported this potential downside.

Next, we would like to address the reaction to the change in the level of the decision parameter in round 11. Figure 3 shows the distribution of changes in average group contributions from round 10 (the last round with different parameters across treatments) to round 11 (the first one with the same parameters across treatments). In both Base and LowCost, most changes were concentrated in the interval $(-1, 1]$ and the average change across groups was

−0.25. In HighReturn, changes displayed a higher variance and their average was −1.11. The difference in the distributions of changes between HighReturn and Base is statistically significant in a two-sided Fisher exact test with $p = 0.0670$. In other words, some participants in HighReturn reacted relatively strongly to the parameter change in round 11. This strong initial response was followed by a rapid decay in contribution levels in later rounds.

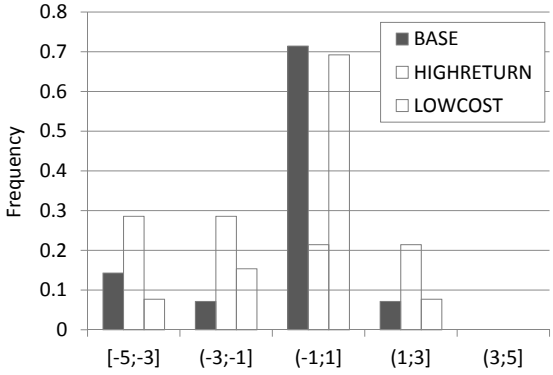


Figure 3: Frequency distribution of the size of the change in average contributions from round 10 to round 11

4.2 Punishment

In addition to allowing a comparison of the effects of a high initial marginal return on public-good contributions and low initial punishment costs on the average contribution to the public account, our data supplies data on the actual use of the punishment possibility. Figure 4 shows the average punishment points assigned over time.

For rounds 1-10, we do not obtain a statistically significant difference in the use of punishment across treatments. On average, subjects assigned 1.46 punishment points to the other three group members in Base, 1.57 points in HighReturn, and 1.56 points in LowCost. This is surprising in view of the fact that the incentives to induce other group members to contribute were stronger in LowCost than in the other two treatments (due to the low marginal cost of punishment), but this finding is in line with the results of Nikiforakis and Normann

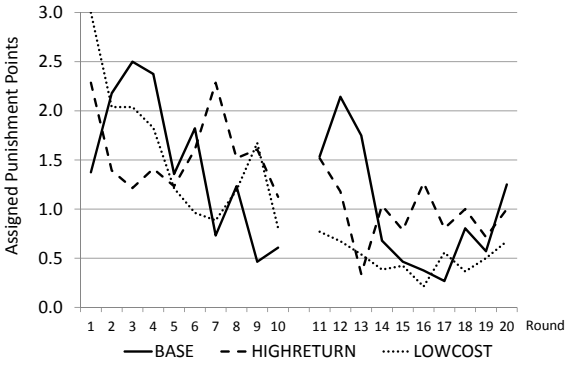


Figure 4: Average assigned punishment points

(2008).[10] Even more surprising is that subjects in LowCost reduced their assignment of punishment points in the second half of the experiment by more than participants in Base did, although incentives for punishment were the same in all treatments in rounds 11-20.[11] This difference in punishment is not statistically significant. However, it may help to explain why average contributions in LowCost decreased: the above finding actually suggests that subjects in LowCost "overreacted" to the change in the price of punishment points, as they assigned fewer punishment points in rounds 11-20 than subjects in Base did, although the two average contributions were similar and the punishment costs were the same.[12] Punishment no longer serves as a threat to potential free riders when they learn that its execution becomes increasingly unlikely. Consequently, contributions in LowCost declined.

[10]In their sample, average point deductions due to the assignment of punishment points amounted to 2.13 with a punishment effectiveness of 2, and to 1.70 with a punishment effectiveness of 4.

[11]In accordance with results obtained in other studies (see, e.g., Sefton et al. 2007), there is a decrease in punishment over time.

[12]This type of overreaction to new information has also been reported , for example, by De Bondt and Thaler (1985) for stock markets.

4.3 Beliefs

We elicited participants' beliefs about the average contribution over the next ten rounds at the beginning of rounds 1 and 11. The beliefs of participants in the different treatment conditions are compared to the actual average contributions in Figure 5.

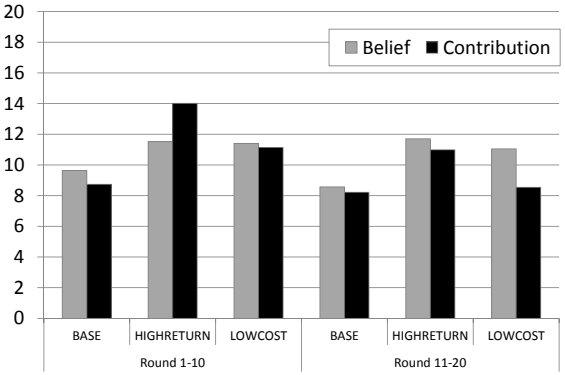


Figure 5: Beliefs v. actual contributions

The comparison between beliefs and actual average contributions illustrates that (i) participants in HighReturn underestimated contributions in rounds 1-10 ($p = 0.0418$), and (ii) participants in LowCost overestimated contributions in rounds 11-20 ($p = 0.0034$). Thus, one may conclude that subjects who were confronted with a high initial marginal return were actually surprised about how well the coordination went in rounds 1-10. This may have helped to create the path dependence in contributions that we have referred to above. In contrast, subjects who experienced low initial punishment costs were quickly disappointed by the way cooperative behavior unfolded in the second ten periods. This could help to explain the very rapid decay in contribution levels in treatment LowCost, in which average contributions were higher than in Base in round 11, but fell rapidly in subsequent rounds.

5 Discussion and conclusion

This paper investigates the path dependence of individual contributions to a public good by analyzing experimental data. The repeated linear public-good game we use replicates key features of the types of coordination challenges faced by many organizations. The results show that the initial and temporary use of high-powered incentives can produce temporary path dependence. The present study differentiates between positive and negative incentives and finds statistically significant differences in behavior only for positive incentives. It must be noted that the pattern of average contributions in the second half of the experiment could be interpreted as an indication that the limited use of incentives can actually create long-term drawbacks for an organization.

The translation of our findings to practical applications, such as the provision of incentives for newly-formed teams, suggests that organizations should carefully consider whether to resort to temporary incentives at all. However, as has been pointed out by Hamman et al. (2007), the empirical results that have been produced to date suggest that the consequences of temporary incentives rely to a great deal on the precise implementation and circumstances, opening an avenue for further research. For example, comparison of the findings of the present paper to the results of Bruttel and Friehe (2010) indicates that the extent of strategic uncertainty may play a critical role for path dependence of behavior.

Appendix: Instructions[13]

Introduction

Thank you for participating in the experiment.

From this point on, please remain seated and stop communicating with other participants. These instructions are identical for all participants. Please read the instructions carefully. If you have any questions or should any uncertainties arise, please ask one of the supervisors for help. We will come to your place to answer your questions in private.

During the experiment you will run through a sequence of decisions and events 20 times. In each of the 20 rounds, there will be two decision stages, which will be explained later on. After 10 rounds, a modification of the decision parameters might possibly be announced on the computer screen. This (possibly modified) set of decision parameters will then remain unchanged for the final 10 rounds.

For all 20 rounds, you will be grouped with three randomly selected participants. No group member will receive information about the identity of the other group members. Your payoff in this experiment depends on your own decisions and the decisions of these three participants.

Your earnings during the experiment will be denoted in points. Your total payoff is determined by the sum of your income levels over the 20 rounds. Directly after the experiment you will be paid 1 euro in cash for each 30 points you have earned. An allocation of 90 points will be applied to your point account as an initial endowment.

Before the actual experiment begins, please answer some control questions about the experiment on the computer screen in order to ensure that all participants understand the instructions correctly. Your responses to the control questions do not impact your payoff in the experiment.

Detailed description of the two decision stages:

During the experiment each round consists of two decision stages, which will now be described:

[13] These are the translations of the German instructions for the BASE condition. Instructions in the HIGHRETURN and LOWCOST treatments were identical except for the levels of the return on public-good contributions and the costs of punishment, respectively.

First stage:

At the beginning of each round, each participant will receive 20 points. We will refer to these points as your endowment. Your task is to make a decision about how to use your endowment. You must decide how many out of the 20 points you want to invest in a project and how many you want to keep for yourself. The consequences of this decision will be explained later on.

At the beginning of each round, the following screen will appear in the first stage:

You decide on your contribution to the project by entering a number between 0 and 20 into the array and clicking the "OK" button. Thereafter, you will not be able to change your decision for the period in question. Your contribution automatically determines how many points you keep for yourself, namely (20 minus your contribution) points. Thus, if you were to invest 2 points in the project, you would keep 18 points for yourself.

Your earnings from the first stage depend on your decisions and the decisions of the three other group members. You keep all points that you do not invest in the project. The points you contribute to the project will be multiplied by 1.2, i.e. increased by 20%, and the increased amount will be split uniformly among all four group members. For example, if you contribute 5 points to the project, they will be increased to 6 points and split among all 4 group members. Each group member (you included) therefore earns 1.5 points due to your contribution of 5 points. If another group member contributes 5 points to the project, then every group member will again earn 1.5 points. Contributions to the project will always be increased by 20% and the resulting amount will be split equally among all four group members. Every group member consequently receives the same income from the project. This income is calculated in the following way:

$$\textbf{Your income from the project} = 1.2 * \frac{\text{Sum of all contributions to the project}}{4}$$

Thus, if you invest in the project, you thereby increase the incomes of the other participants. Conversely, if other group members contribute to the project, they increase your income. The points you keep for yourself do not increase other participants' incomes.

Your earnings from the first stage of a period are given by the sum of the points you have kept for yourself and the points you receive from the project:

Your earnings during the first stage =
points kept for yourself + earnings from the project

Two examples should clarify this calculation for your earnings:

- If all group members keep their endowment of 20 points for themselves, everyone earns 20 points from the points kept and nothing from the project. Thus, everyone gets 20 points.
- If all group members invest their entire endowment of 20 points in the project, the sum of all contributions will add up to 80 points. This amount will then be multiplied by 1.2, i.e., increased by 20%, and the resulting 96 points will be split among four group members, i.e., everyone will earn 24 points from the project and nothing from the points kept. Thus, everyone gets 24 points.

After you and the other three group members have made your decisions, you will receive information about the sum of contributions (including your contribution) all group members have made to the project on the following earnings screen. Furthermore, you will be informed of the points you earned in the first stage.

Summary of the first stage:

Second stage:

In the second stage, you will receive information about how many points the other group members invested in the project. In addition, you can reduce other group members' income by assigning deduction points in the second stage. The other group members can also reduce your income. If you assign 0 deduction points, you will not change the income levels of the other members . This will become clear immediately from the following screen for the second stage:

Screen at the second stage

Here you can see how much each group member has invested in the project. The earnings of a group member are stated below the member's contribution. Your contribution to the project is always in the first column, written in blue. The contributions of the other group members are listed in the three remaining columns. Please note that the labels group member 1, 2, and 3 will not refer to the same participant in every round. A different group member will be labeled as the second group member in rounds 1, 2 and subsequent rounds.

You must decide for each group member whether to assign deduction points and, if so, how many you want to assign to this participant. You can assign between 0 and 20 deduction points to each group member. Please enter your desired point level into the designated field.

You reduce the earnings of a group member by 1 point per deduction point you assign. However, the assignment of deduction points also reduces your earnings. Each deduction point you assign reduces your earnings by 1/2 point. If, for example, you assign 2 deduction points to another group member, his or her earnings will be reduced by 2 points and you will have costs of 1 point. If you additionally assign 18 deduction points to another group member, you will reduce that member's income by 18 points, while you will have costs of 9 points. If you assign 0 points to another group member, neither that member's earnings nor yours will be reduced.

If you are assigned deduction points by other group members, your earnings will be reduced accordingly. If, for example, you earned 15 points in the first stage and the other group members were to assign 12 deduction points to you in the second stage, your earnings would be reduced by 12 points. However, your earnings cannot be reduced by more than what you have. This means that, for instance, if your first stage income was 8 points but you were assigned 12 deduction points, your earnings would be reduced by 8 points. The group member who assigned 12 deduction points to you would still have to pay the entire costs of the 12 deduction points, amounting to 6 points.

To make the final decision concerning the assignment of deduction points, you must click on the "OK" button again. As long as you do not click on the "OK" button, you can revise your decision.

Your entire earnings for the two decision stages are calculated as follows:

Earnings of the round = earnings at the end of the second stage

= earnings from the first stage
- sum of deduction points assigned to you
- costs of deduction points you assigned.

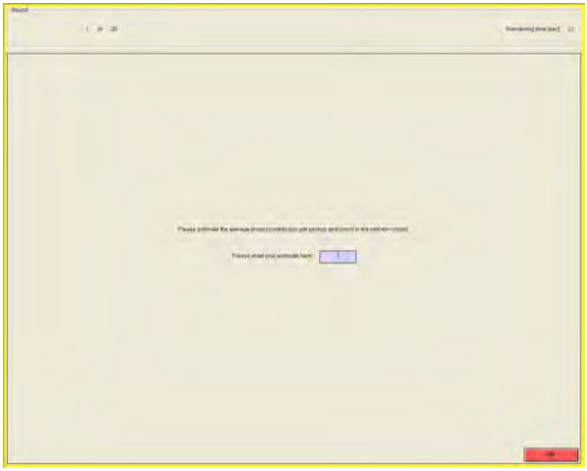
Your earnings at the end of the second stage if the deduction points you were assigned are equal to or higher than your first-stage income
= 0 - costs of the deduction points you assigned

Please note that your earnings at the end of the second stage can be negative if the costs of the deduction points you assigned exceed your earnings from the first stage minus the sum of the deduction points assigned to you. After all participants have made their decision, your earnings for the round will be shown on the following screen:

Earnings screen at the end of the second stage:

Your expectations:

Before the beginnings of the 1st and the 11th rounds, you will be called upon to make an estimation of the average contribution of group members. You must make a guess about how much the other three group members will on average invest in the project per round in the next 10 rounds. Since the other participants also receive an endowment of 20 points per round, your estimated value accordingly must range from 0 to 20. This estimation has no impact on your payoff. For this purpose the following screen will be displayed:



After the experiment, we will ask you to fill out a short questionnaire. Your earnings will then be paid to you in cash. The exchange rate is 30 points for 1 euro.

References

Ahn, T.K., Ostrom, E., Schmidt, D., Shupp, R., and J. Walker, 2001. Cooperation in PD games: Fear, greed, and history of play. Public Choice 106, 137-155.

Bohnet, I., and S. Huck, 2004. Repetition and reputation: Implications for trust and trustworthiness when institutions change. American Economic Review Papers & Proceedings 94, 362-366.

Brandts, J., and D.J. Cooper, 2006. A change would do you good ... An experimental study on how to overcome coordination failure in organizations. American Economic Review 96, 669-693.

Bruttel, L., and T. Friehe, 2010. On the path dependence of tax compliance. TWI Research Paper No. 59.

Carpenter, J.P., 2007. The demand for punishment. Journal of Economic Behavior and Organization 62, 522-542.

Cason, T.N., Savikhin, A., and R.M. Sheremeta, 2009. Cooperation spillovers in coordination games. Unpublished working paper.

Chaudhuri, A., 2011. Sustaining cooperation in laboratory public goods experiments: A selective survey of the literature. Experimental Economics 14, 47-83.

De Bondt, W.F.M., and R. Thaler, 1985. Does the stock market overreact? Journal of Finance 15, 793-805.

Fehr, E., and S. Gächter, 2000. Cooperation and punishment. American Economic Review 90, 980-994.

Fischbacher, U., 2007. z-Tree: Zurich Toolbox for Ready-made Economic Experiments. Experimental Economics 10, 171-178.

Gächter, S., Kessler, E., and M. Königstein, 2008. Performance incentives and the dynamics of voluntary cooperation. Unpublished working paper.

Gächter, S., and E. Renner, 2010. The effects of (incentivized) belief elicitation in public good experiments. Experimental Economics 13, 364-377.

Gneezy, U., and A. Rustichini, 2000. A fine is a price. Journal of Legal Studies 29, 1-17.

Greiner, B., 2004. An online recruitment system for economic experiments. In: Kremer, K. and V. Macho (Eds.): Forschung und wissenschaftliches Rechnen 2003. GWDG Bericht 63, Göttingen: Gesellschaft für Wiss. Datenverarbeitung, 79-93.

Hamman, J., Rick, S., and R.A. Weber, 2007. Solving coordination failure with "all-or-none" group-level incentives. Experimental Economics 10, 285-303.

Knez, M., and C. Camerer, 2000. Increasing cooperation in prisoner's dilemmas by establishing a precedent of efficiency in coordination games. Organizational Behavior and Human Decision Processes 82, 194-216.

Kube, S., and C. Traxler, forthcoming. The interaction of legal and social norm enforcement. Journal of Public Economic Theory.

Nikiforakis, N., and H.T. Normann, 2008. A comparative statics analysis of punishment in public-good experiments. Experimental Economics 11, 358-369.

Sefton, M., Shupp, R., and J.M. Walker, 2007. The effect of rewards and sanctions in provision of public goods. Economic Inquiry 45, 671-690.

No. 67 august 2011

Lisa Bruttel
Tim Friehe

Path dependence in public-good games

Research Paper Series
Thurgau Institute of Economics and Department of Economics
at the University of Konstanz

Member of
thurgauwissenschaft
www.thurgau-wissenschaft.ch

THURGAU INSTITUTE
OF ECONOMICS
at the University of Konstanz

Path dependence in public-good games*

Lisa Bruttel† Tim Friehe‡

August 5, 2011

Abstract

This paper presents experimental evidence that contributions to a public good can be path-dependent for a limited time span. We study a repeated linear public-good game with punishment opportunities. Our data shows that subjects who had experienced a higher marginal return on public-good contributions in rounds 1-10 contributed more to the public good in rounds 11 and 12, even though they faced the same marginal return as the control group in these later rounds. In contrast, differences in contributions were not significant when comparing subjects bearing the same current costs of punishment points, but having had different costs in the past.

Keywords: public-good game, team, punishment, path dependence, experiment

JEL-Classification: C91, H26, K42

*We would like to thank Kate Bendrick, Gerald Eisenkopf, Urs Fischbacher, Katrin Schmelz, Simeon Schudy, Pascal Sulser, Verena Utikal, Irenaeus Wolff, and Natalie Zimmer for very helpful suggestions.

†University of Konstanz, Department of Economics, Box 131, 78457 Konstanz, Germany. E-mail: lisa.bruttel@uni-konstanz.de. Phone: +49-7531-88-3214. Fax: +49-7531-88-2145.

‡University of Konstanz, Department of Economics, Box 136, 78457 Konstanz, Germany. E-mail: tim.friehe@uni-konstanz.de. Phone: +49-7531-88-2534. Fax: +49-7531-88-4135.

1 Introduction

This paper addresses whether the initial and temporary use of high-powered incentives may help to lastingly promote cooperation in groups. In a variety of circumstances, group members either fail to coordinate or coordinate on a non-cooperative, often Pareto-dominated outcome. Imagine, for example, that a firm forms a team for which success is determined by the aggregate of unverifiable individual effort levels. In such a set-up, coordination failure or inefficient coordination might easily arise because team members bear the full costs of high effort but would only reap full benefits therefrom if other team members also invested high effort. The firm may therefore attempt to induce coordination on high-effort contributions (i.e., the cooperative outcome) from the start by implementing high-powered incentives over some initial time period. The firm could put positive incentives into place by leveraging performance pay in relation to team success, or it might implement negative incentives by facilitating peer monitoring and reporting, for example. If high-powered incentives prove successful in obtaining cooperation among team members, their initial use may carry over to periods in which incentives are scaled back. In such a case, incentives would help team members coordinate on the cooperative outcome while they are in effect, and would additionally produce a cooperative spillover to team interactions in periods without effective incentives. However, the limited use of high-powered incentives may also backfire, in the sense that team performance might become worse after the phase-out than it would have been without the initial use of the incentives. This could arise because team members perceive the reduction in rewards as unfair or because they lose intrinsic motivation, among other possible reasons. Finally, the individual effort provision might be unrelated to the outcomes in earlier periods, whether coordination failure or success, because past incentives are no longer relevant for current decisions. In order to address this theme of the path dependence of individual contributions, we study a repeated linear public-good game with punishment opportunities to simulate contexts such as the team scenario described above.

This paper examines whether the levels of decision parameters relevant in rounds 1-10 of the repeated linear public-good game continue to influence individual contributions in rounds 11-20, despite being no longer decisive for an individual's decision problem. We investigate the path dependence of individual contributions with respect to both positive and negative incentives in a fixed-partners design. Specifically, we consider rounds 11-20 and contrast the behavior of subjects who had a high marginal return on contributions to the public good in rounds 1-10

to the behavior of individuals who had a low marginal return on public-good contributions. Similarly, we contrast the behavior of participants who experienced low punishment costs in the first ten rounds to the behavior of subjects with high punishment costs. Principally, both low punishment costs and a high marginal return on public-good contributions are conducive to high individual public-good contributions, although the effect of punishment costs is indirect. A lower level of punishment costs makes individuals more willing to punish low contributions, all else held equal; knowing this, individuals are better off when they contribute more.[1]

The experimental results suggest that there can be path dependence of individual contributions to the public good. However, significant differences in average contributions were found only in rounds 11 and 12; i.e., there is no long-term persistence of the effect. Moreover, we obtained significant differences in average contributions only when contrasting the behavior of participants who initially had a high marginal return on public-good contributions to the behavior of individuals with a low marginal return on public-good contributions throughout all rounds. Specifically, individuals who had a high marginal return on public-good contributions in rounds 1-10 and a low one in rounds 11-20 contributed more to the public good in rounds 11 and 12 than individuals who had experienced a low marginal return in all 20 rounds. In contrast, the contributions of individuals with low punishment costs in rounds 1-10 and high costs in rounds 11-20 were not different at a conventional level of statistical significance from contributions of individuals who had high punishment costs in all 20 rounds.

Our analysis questions the long-term effectiveness of the initial and temporary usage of high-powered incentives when it comes to cooperation in groups. In addition, the present study suggests that some policy instruments are better suited to produce (at least transitory) spillovers in the level of contributions than others. The temporary use of high levels of the marginal return on the public good produced significant differences in rounds 11 and 12, whereas the use of low punishment costs failed to do so.

The structure of the article is as follows: In Section 2, we relate our contribution to the literature; we present the experimental design and procedures in Section 3. The experimental results are discussed in Section 4, and Section 5 discusses our findings and concludes the paper.

[1]The hypothesis underlying the mechanism via punishment costs has empirical support. Carpenter (2007) has established that subjects in voluntary contribution mechanisms indeed demand more punishment when the price decreases, and Nikiforakis and Normann (2008) have shown that contributions increase monotonically in punishment effectiveness. See also the very recent survey by Chaudhuri (2011).

2 Connections to the literature

Our paper explores the potential path dependence of individual contributions in a public-good game with punishment opportunities.[2] A number of papers have studied related issues, with Bruttel and Friehe (2010), Bohnet and Huck (2004), Brandts and Cooper (2006), and Hamman et al. (2007) particularly relevant to the current paper.

Bruttel and Friehe (2010) investigate path dependence in a tax-evasion setting that is repeated for 20 rounds but not characterized by strategic interdependence, finding that the penalty for tax evasion valid in rounds 1-10 persistently influences income declarations in rounds 11-20. While the present study uses a similar format to that of Bruttel and Friehe (2010), strategic interdependence and uncertainty of participants are vital elements of our setup. Bohnet and Huck (2004) study a binary-choice trust game involving two subjects; they find that individuals who took part in a fixed-partner setup in rounds 1-10 trust more and are more trustworthy in rounds 11-20 when interacting with random strangers on a one-shot basis than subjects who were in the stranger condition in all rounds. The present study challenges their results by adding to both the complexity of the strategy space and the degree of strategic uncertainty. However, like Bohnet and Huck (2004), we are interested in whether the initial provision of circumstances conducive to cooperation leaves imprints on decisions at later points in time. In contrast to the two studies cited above (and to the experiment in the present paper), which consider the effect of the initial provision of incentives on cooperation, Brandts and Cooper (2006) and Hamman et al. (2007) are interested in persistent effects when teams that are failing to coordinate are temporarily confronted with stronger incentives. Specifically, Brandts and Cooper (2006) study 3×10 rounds of a minimum-effort game with fixed groups of four subjects and in which most groups suffer from coordination failure after the first ten rounds, exploring the consequences of various changes to the bonus rate in rounds 11-30. Inter alia, the authors contrast the behavior of subjects who face a low bonus rate in rounds 1-10, a high bonus rate in rounds 11-20, and a low bonus rate in rounds 21-30 to pooled behavior in rounds 1-10; they find that reducing the bonus rate did not bring the level of minimum effort back to the level measured in rounds 1-10. This is interpreted as indicating the path dependence of effort provision; however, their design lacks a control group with low bonus rates

[2]The literature concerning individual contributions in public-goods games is extensive. A recent survey is provided by Chaudhuri (2011).

in all rounds. This gap was filled by Hamman et al. (2007) who study a similar minimum-effort game with five different effort levels, fixed groups, and 30 rounds. The authors depart from the design investigated by Brandts and Cooper (2006) by (i) studying positive and negative incentives, and (ii) restricting bonus payments (penalties) to a fixed magnitude instead of setting them to be proportional to the minimal effort level, and applying them only in outcomes with sufficiently high (low) effort levels. They present in particular two related findings: (i) the long-term persistent benefits of incentives are smaller than those presented by Brandts and Cooper (2006), and (ii) positive and negative incentives are effective to a comparable degree. Like Hamman et al. (2007), we consider positive and negative incentives. However, in our design, negative repercussions of small individual contributions do not arise automatically from the compensation scheme but are endogenous, because the subjects themselves determine the utilization of punishment opportunities. In contrast to Hamman et al. (2007), we will show that there are marked differences between what we consider to be positive and negative incentives in our setup. In contrast to the last three studies cited above (i.e., Bohnet and Huck 2004, Brandts and Cooper 2006, and Hamman et al. 2007), we examine a public-good game with punishment opportunities in which participants choose from a rich set of alternatives. We believe that the increase in the number of alternatives makes the decision problem both more realistically complex and less of a binary choice, so that our study provides a more conservative test for potential path dependence.

The above studies examine the possibility of effecting desirable long-term effects by temporarily changing the rules to a game. There is also evidence for the detrimental long-term consequences of explicit incentives after their abolishment. Gneezy and Rustichini (2000) show that the introduction of fines for tardy parents at a child-care center increases tardiness even after fines are no longer used, and Gächter et al. (2008) establish that the use of explicit incentives in an employment relationship can crowd out voluntary cooperation even after the reliance on explicit incentives is discontinued. Finally, instead of focusing on how behavior changes over time in a given game (i.e., exploring whether there are spillovers across periods of time), it is also possible to consider spillovers across games. For example, Cason et al. (2009) show that cooperation in a minimum-effort game can be achieved through a behavioral spill-over from a median-effort game when the former is played after the latter. Similar results have been found by Ahn et al. (2001) and Knez and Camerer (2000), among others.

3 Experimental design and procedures

3.1 Design

The experiment was based on a repeated linear public-good game with punishment opportunities, played by fixed groups of four subjects.[3] There were two stages to each round and two sequences of ten rounds each. In *Stage 1*, participants decided (simultaneously and without the possibility of communicating) how to split their endowments of 20 points between a private account and a public account. The contribution to the public good by subject i is denoted c_i. The contributions to the public good by all four group members were collected in the public account, and the total was multiplied by $\alpha > 1$ before reallocating the returns equally to group members. This implies that participant i's payoffs from the first stage are given by

$$\pi_i = 20 - c_i + \alpha/4 \sum_{j=1}^{4} c_j \tag{1}$$

At the end of Stage 1, participants were informed about π_i and the sum of contributions to the public good. At the beginning of *Stage 2*, group members obtained information about the other group members' first-stage income and their contributions to the public good. Individual i could then assign punishment points to be applied to subject j, denoted p_{ij}, at a cost of $\beta > 0$. One punishment point lowered subject j's income by one, i.e., it had a punishment effectiveness equal to one. This implies that payoffs at the end of Stage 2 and thus from any given round can be stated as

$$\Pi_i = \pi_i - \sum_{j \neq i} p_{ji} - \beta \sum_{j \neq i} p_{ij} \tag{2}$$

The impact of punishment points imposed by others on subject i, i.e. the effect of $\sum_{j \neq i} p_{ji}$, was restricted to the degree that $\pi_i - \sum_{j \neq i} p_{ji} \geq 0$. Nonetheless, it was possible that $\Pi_i < 0$.[4] The identification of a given participant to the other members of the group on the screen (group members 2-4) was subject to a random draw in every round in order to avoid spillovers across rounds.[5] At the end of Stage 2, participant i obtained a summary detailing the contribution to

[3] The design of individual rounds is similar to Fehr and Gächter (2000). A translation of the instructions is presented in an appendix to our paper.

[4] Participants were given an initial endowment to compensate for potential losses in initial rounds.

[5] The instructions in the appendix include a screenshot showing how this was implemented.

the public good c_i, the sum of individual contributions, first-stage income π_i, both the number and the costs of punishment points allocated by this participant to other group members, the number of punishment points received, and Π_i. The payoff structure was the same for all participants, which in turn was common knowledge.

Before the start of the experiment, all participants were informed that a parameter of the decision-problem might change after the first ten rounds, but they did not know which parameter would be altered, or in which direction. In order to ensure that participants were indeed aware of the change, they were asked to answer a question about the current levels of the two decision parameters at the beginning of round 11. In our study, the marginal return on public-good contributions α and (marginal) punishment costs β are the treatment variables of interest, and we distinguish between a low and a high level for each. Both α and β stayed constant for rounds 1-10 and 11-20 in all treatments. Thus, we may denote α_{1-10} (β_{1-10}) as the marginal return (punishment costs) applicable in the first ten rounds. The levels applied in rounds 11-20 were the same for all three treatments and represent the low marginal return on public-good contributions and the high level of punishment costs, i.e., $\alpha_{11-20} = 1.2$ and $\beta_{11-20} = 0.5$. The experimental design consisted of 3 treatments (see Table 1).

Treatment	α_{1-10}	β_{1-10}	α_{11-20}	β_{11-20}	Subjects	Observations
Base	1.2	0.5	1.2	0.5	56	14
HighReturn	2	0.5	1.2	0.5	56	14
LowCost	1.2	0.25	1.2	0.5	52	13

Table 1: Treatment parameters, number of subjects, and independent observations

In our setup, game-theoretic predictions based on the concept of subgame-perfectness in view of monetary aspects predict zero punishment and zero contributions in all treatments. Accordingly, there is no straightforward way to calibrate the two decision parameters in a way that they would be equivalent regarding equilibrium behavior in theory. Regarding the precise levels of our treatment variables, the two levels of $\alpha/4$ were chosen to be the neighboring levels of the commonly used level 0.4 (see, e.g., Fehr and Gächter 2000, Kube and Traxler forthcoming, Nikiforakis and Normann 2008). As a result, $\alpha = 2$ is relatively high and $\alpha = 1.2$ relatively low when compared to the standard level of this parameter. It must be noted that this fact (i.e., that $0.3 < 0.4$) will tend to tilt behavior in our experiment against cooperation in rounds 11-20, all else held equal. The levels of punishment costs were chosen in view of the findings by

Nikiforakis and Normann (2008). In their experiment, which was based on a repeated linear public-good game with punishment opportunities, a punishment effectiveness of 2 (comparable to punishment costs of 0.5) yielded constant average contributions, whereas a punishment effectiveness of 4 (comparable to punishment costs of 0.25) yielded strictly increasing average contributions, averaging at 17.93 in rounds 6-10.

Based on the findings of other analyses of individual contributions in public-good games with punishment opportunities (see, e.g., Chaudhuri 2011), we expect that average contributions in HighReturn and LowCosts will be higher than in Base in rounds 1-10. However, we do not have a clear-cut prediction about path dependence, that is, the differences between the average contributions to the public good in rounds 11-20 in our three treatments. Although the study by Brandts and Cooper (2006) suggests that path dependence may occur, this need not extend to the public-good setting we scrutinize. Indeed, the relatively small variation to the Brandts/Cooper-setup introduced by Hamman et al. (2007) has already cast doubt on the generality of path dependence. Moreover, if the removal of the incentive to punish by the introduction of high punishment costs is perceived as the removal of a sanctioning device, the introduction of low punishment costs in rounds 1-10 may actually backfire in rounds 11-20, as suggested by Gneezy and Rustichini (2000). In other words, even if we were to find path dependence, it is not necessarily clear whether subjects would "underreact" or "overreact" relative to the decision-making behavior in our control group.

3.2 Procedures

The experiment was computerized using z-Tree (Fischbacher 2007). A total of 164 students from various disciplines took part in at most one of the three treatments. They were recruited via ORSEE (Greiner 2004). The experiment took place in the *Lakelab*, the laboratory for experimental economics at the University of Konstanz, in November and December 2010. Sessions lasted less than 90 minutes. The experimental currency was points, with 30 points converted into 1 euro after the experiment. On average, participants earned 17.85 euros in the experiment. The protocol before the start of the experiment was as follows: Subjects first received written instructions for participating in the experiment, and then had to answer control questions displayed on their computer screens. The experiment started only after all subjects had

answered the control questions correctly. At the end of the session, the participants were asked to complete a brief questionnaire. We elicited beliefs about the expected average contribution for the upcoming ten rounds at the beginning of rounds 1 and round 11 without incentivizing them. According to Gächter and Renner (2010), eliciting beliefs in a non-incentivized manner has the advantage that it does not systematically distort contribution levels.

4 Results

In this section, we discuss the experimental data. We turn first to our main interest; the pattern of contributions over time. Subsequently, we briefly discuss the assigned punishment points and the participants' beliefs about the expected average contribution for the upcoming ten rounds.

4.1 Contributions

Figure 1 shows average individual contributions to the public good in our three treatments across all 20 rounds. In line with other studies (see, e.g., Chaudhuri 2011), the average contribution in the baseline treatment was about 50 percent of the endowment. Average contributions in rounds 1-10 were very stable over time. In particular, there was no decline in round 9 or 10. In contrast to other studies, there was no restart effect in BASE, meaning that group members' allocations to the public good in round 11 were not higher than in round 10 (see, e.g., Hamman et al. 2007, Sefton et al. 2007).[6]

From Figure 1, both the high marginal return on public-good contributions and the low level of the punishment costs increased contribution levels in rounds 1-10. The difference is statistically significant only when we compare contributions in HIGHRETURN and in BASE ($p = 0.0204$).[7] With regard to the difference in average contributions in rounds 1-10 between LOWCOST and BASE, we obtain $p = 0.2758$.

[6] It seems that this effect is not contingent on the regulation of information. In Sefton et al. (2007), subjects were informed of the whole sequence of decision periods at the beginning of the experiment. In Hamman et al. (2007), participants received instructions for later rounds only after the experiment had already started.

[7] Unless noted otherwise, we report the significance levels of two-sided Wilcoxon rank sum test in parentheses.

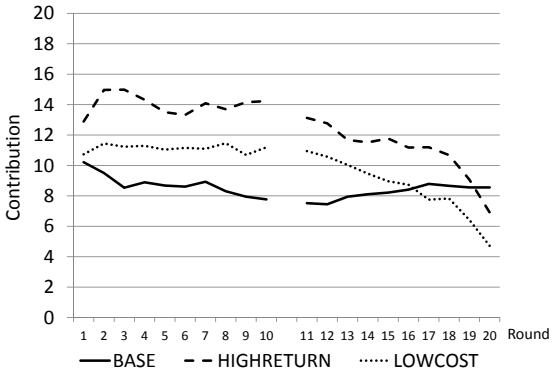


Figure 1: Average public-good contributions

Result 1 *(a) Individuals with high marginal returns to public-good contributions in rounds 1-10 ($\alpha_{1-10} = 2$) and low marginal returns in rounds 11-20 ($\alpha_{11-20} = 1.2$) contributed more to the public good in rounds 1-10 than subjects with low marginal returns in rounds 1-20 ($\alpha_{1-10} = \alpha_{11-20} = 1.2$). (b) Individuals with low punishment costs in rounds 1-10 ($\beta_{1-10} = 0.25$) and high punishment costs in rounds 11-20 ($\beta_{11-20} = 0.5$) did not contribute significantly more to the public good in rounds 1-10 than subjects with high punishment costs in rounds 1-20 ($\beta_{1-10} = \beta_{11-20} = 0.5$).*

Next, we turn to the path dependence of contributions to the public good. Table 2 reports the results of round-wise Wilcoxon rank sum tests, examining whether contributions in the HighReturn (LowCost) condition were equal to those in Base. The test results show that there was indeed transitory path dependence in HighReturn. Contributions in this treatment were higher than those in Base in rounds 11 and 12. However, experiencing high marginal return on contributions to the public good in rounds 1-10 did not yield persistent differences. From round 13 onwards, contributions in Base and HighReturn were statistically indistinguishable. Comparing contributions in LowCost and in Base, we find no significant differences from round 11 onwards. This does not come as a surprise, given that the difference between contribution levels in rounds 1-10 was not statistically significant, either. In summary, our experimental data affirms the possibility of path dependence, but casts a doubt on the long-term persistence of such effects.

	p-value	
Round	BASE vs. HIGHRETURN	BASE vs. LOWCOST
11	0.0204	0.1970
12	0.0512	0.2340
13	0.1556	0.4354
14	0.2420	0.7872
15	0.2076	0.9044
16	0.3370	0.9840
17	0.3576	0.7188
18	0.3472	0.9442
19	0.8336	0.5754
20	0.8180	0.1970

Table 2: Comparison of contributions: p-values from two-sided Wilcoxon rank sum test.

Result 2 *In* HIGHRETURN*, there was transitory path dependence of the level of contributions to the public good. Individuals with high marginal returns to public-good contributions in rounds 1-10 ($\alpha_{1-10} = 2$) and low marginal returns in rounds 11-20 ($\alpha_{11-20} = 1.2$) contributed more to the public good in rounds 11 and 12 than subjects with low marginal returns in rounds 1-20 ($\alpha_{1-10} = \alpha_{11-20} = 1.2$).*

In Figure 1, average contributions in rounds 11-20 followed a downward trend in the LOWCOST and HIGHRETURN conditions. In contrast, average contributions were more or less constant in BASE. However, patterns of average contributions might not tell the whole story. The average in BASE hides a wide variance in group contribution levels.[8] In contrast to what held true in the other two treatments, contribution levels of three groups in BASE actually increased over time. Figure 2 shows the results from clustering groups according to the change in the level of average contributions in rounds 16-20 when compared to the level in rounds 11-15.[9] In the LOWCOST and HIGHRETURN conditions, the contributions of the vast majority of groups declined in the fourth quarter of the experiment. In the BASE treatment, the con-

[8]The pattern of contributions over time is very similar to that from a comparable treatment in Nikiforakis and Normann (2008). For a punishment effectiveness of 2, they also report average contributions which are roughly constant but result from very diverse group contribution patterns.

[9]We denote groups that experienced a change of not more than one unit as constant.

tributions of most groups either remained relatively stable or even showed an increase. The differences illustrated in Figure 2 are statistically highly significant in a two-sided 2x3 Fisher exact test (p-values: LowCost vs. Base = 0.0021, HighReturn vs. Base = 0.0094).

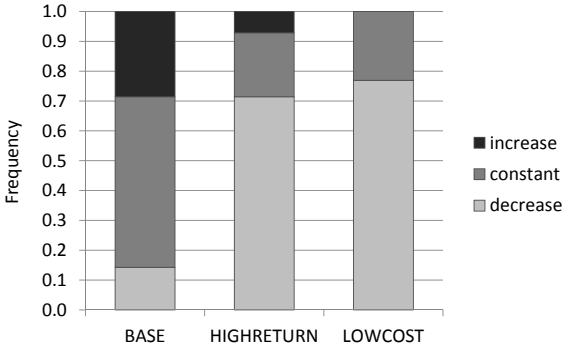


Figure 2: Direction of change of the average group contribution in rounds 16-20 compared to rounds 11-15.

Result 3 *In* HighReturn *and* LowCost*, there was a downward trend in average contributions to the public good after the reduction in the marginal returns to contributions to the public good and after the increase in the marginal costs of punishment, respectively (i.e., in rounds 11-20).*

As a result of the findings above, the data represented in Figures 1 and 2 does not help to rule out the possibility that the initial and temporary use of high-powered incentives in our design may backfire in the long-run by yielding a lower contribution level than that which would have been obtained without the use of such incentives. Neither Brandts and Cooper (2006) nor Hamman et al. (2007) have reported this potential downside.

Next, we would like to address the reaction to the change in the level of the decision parameter in round 11. Figure 3 shows the distribution of changes in average group contributions from round 10 (the last round with different parameters across treatments) to round 11 (the first one with the same parameters across treatments). In both Base and LowCost, most changes were concentrated in the interval $(-1, 1]$ and the average change across groups was

−0.25. In HighReturn, changes displayed a higher variance and their average was −1.11. The difference in the distributions of changes between HighReturn and Base is statistically significant in a two-sided Fisher exact test with $p = 0.0670$. In other words, some participants in HighReturn reacted relatively strongly to the parameter change in round 11. This strong initial response was followed by a rapid decay in contribution levels in later rounds.

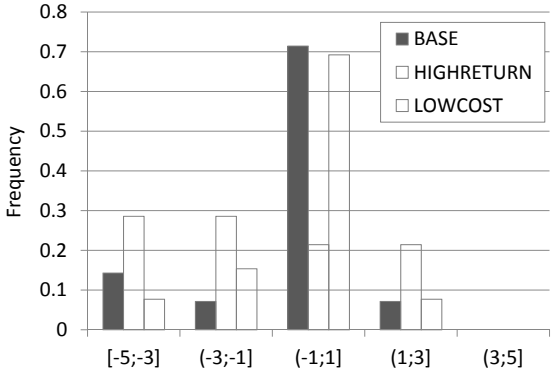


Figure 3: Frequency distribution of the size of the change in average contributions from round 10 to round 11

4.2 Punishment

In addition to allowing a comparison of the effects of a high initial marginal return on public-good contributions and low initial punishment costs on the average contribution to the public account, our data supplies data on the actual use of the punishment possibility. Figure 4 shows the average punishment points assigned over time.

For rounds 1-10, we do not obtain a statistically significant difference in the use of punishment across treatments. On average, subjects assigned 1.46 punishment points to the other three group members in Base, 1.57 points in HighReturn, and 1.56 points in LowCost. This is surprising in view of the fact that the incentives to induce other group members to contribute were stronger in LowCost than in the other two treatments (due to the low marginal cost of punishment), but this finding is in line with the results of Nikiforakis and Normann

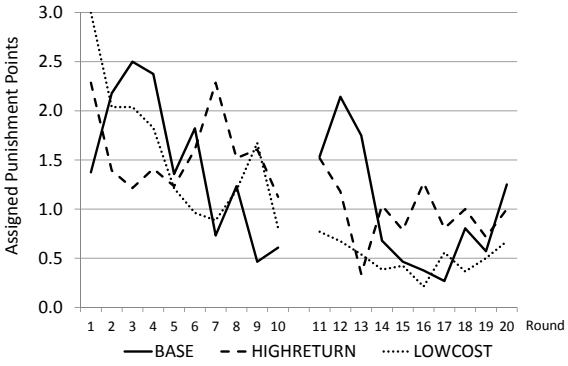


Figure 4: Average assigned punishment points

(2008).[10] Even more surprising is that subjects in LowCost reduced their assignment of punishment points in the second half of the experiment by more than participants in Base did, although incentives for punishment were the same in all treatments in rounds 11-20.[11] This difference in punishment is not statistically significant. However, it may help to explain why average contributions in LowCost decreased: the above finding actually suggests that subjects in LowCost "overreacted" to the change in the price of punishment points, as they assigned fewer punishment points in rounds 11-20 than subjects in Base did, although the two average contributions were similar and the punishment costs were the same.[12] Punishment no longer serves as a threat to potential free riders when they learn that its execution becomes increasingly unlikely. Consequently, contributions in LowCost declined.

[10]In their sample, average point deductions due to the assignment of punishment points amounted to 2.13 with a punishment effectiveness of 2, and to 1.70 with a punishment effectiveness of 4.

[11]In accordance with results obtained in other studies (see, e.g., Sefton et al. 2007), there is a decrease in punishment over time.

[12]This type of overreaction to new information has also been reported , for example, by De Bondt and Thaler (1985) for stock markets.

4.3 Beliefs

We elicited participants' beliefs about the average contribution over the next ten rounds at the beginning of rounds 1 and 11. The beliefs of participants in the different treatment conditions are compared to the actual average contributions in Figure 5.

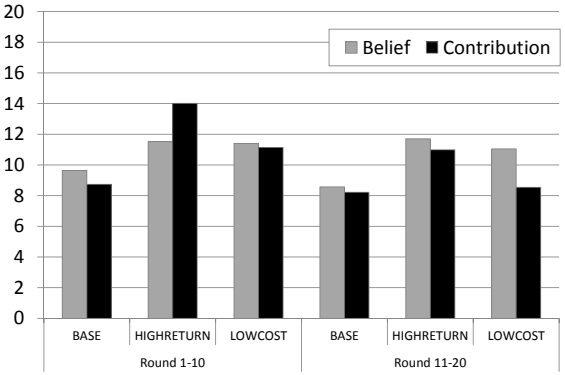


Figure 5: Beliefs v. actual contributions

The comparison between beliefs and actual average contributions illustrates that (i) participants in HighReturn underestimated contributions in rounds 1-10 ($p = 0.0418$), and (ii) participants in LowCost overestimated contributions in rounds 11-20 ($p = 0.0034$). Thus, one may conclude that subjects who were confronted with a high initial marginal return were actually surprised about how well the coordination went in rounds 1-10. This may have helped to create the path dependence in contributions that we have referred to above. In contrast, subjects who experienced low initial punishment costs were quickly disappointed by the way cooperative behavior unfolded in the second ten periods. This could help to explain the very rapid decay in contribution levels in treatment LowCost, in which average contributions were higher than in Base in round 11, but fell rapidly in subsequent rounds.

5 Discussion and conclusion

This paper investigates the path dependence of individual contributions to a public good by analyzing experimental data. The repeated linear public-good game we use replicates key features of the types of coordination challenges faced by many organizations. The results show that the initial and temporary use of high-powered incentives can produce temporary path dependence. The present study differentiates between positive and negative incentives and finds statistically significant differences in behavior only for positive incentives. It must be noted that the pattern of average contributions in the second half of the experiment could be interpreted as an indication that the limited use of incentives can actually create long-term drawbacks for an organization.

The translation of our findings to practical applications, such as the provision of incentives for newly-formed teams, suggests that organizations should carefully consider whether to resort to temporary incentives at all. However, as has been pointed out by Hamman et al. (2007), the empirical results that have been produced to date suggest that the consequences of temporary incentives rely to a great deal on the precise implementation and circumstances, opening an avenue for further research. For example, comparison of the findings of the present paper to the results of Bruttel and Friehe (2010) indicates that the extent of strategic uncertainty may play a critical role for path dependence of behavior.

Appendix: Instructions[13]

Introduction

Thank you for participating in the experiment.

From this point on, please remain seated and stop communicating with other participants. These instructions are identical for all participants. Please read the instructions carefully. If you have any questions or should any uncertainties arise, please ask one of the supervisors for help. We will come to your place to answer your questions in private.

During the experiment you will run through a sequence of decisions and events 20 times. In each of the 20 rounds, there will be two decision stages, which will be explained later on. After 10 rounds, a modification of the decision parameters might possibly be announced on the computer screen. This (possibly modified) set of decision parameters will then remain unchanged for the final 10 rounds.

For all 20 rounds, you will be grouped with three randomly selected participants. No group member will receive information about the identity of the other group members. Your payoff in this experiment depends on your own decisions and the decisions of these three participants.

Your earnings during the experiment will be denoted in points. Your total payoff is determined by the sum of your income levels over the 20 rounds. Directly after the experiment you will be paid 1 euro in cash for each 30 points you have earned. An allocation of 90 points will be applied to your point account as an initial endowment.

Before the actual experiment begins, please answer some control questions about the experiment on the computer screen in order to ensure that all participants understand the instructions correctly. Your responses to the control questions do not impact your payoff in the experiment.

Detailed description of the two decision stages:

During the experiment each round consists of two decision stages, which will now be described:

[13] These are the translations of the German instructions for the BASE condition. Instructions in the HIGHRETURN and LOWCOST treatments were identical except for the levels of the return on public-good contributions and the costs of punishment, respectively.

First stage:

At the beginning of each round, each participant will receive 20 points. We will refer to these points as your endowment. Your task is to make a decision about how to use your endowment. You must decide how many out of the 20 points you want to invest in a project and how many you want to keep for yourself. The consequences of this decision will be explained later on.

At the beginning of each round, the following screen will appear in the first stage:

You decide on your contribution to the project by entering a number between 0 and 20 into the array and clicking the "OK" button. Thereafter, you will not be able to change your decision for the period in question. Your contribution automatically determines how many points you keep for yourself, namely (20 minus your contribution) points. Thus, if you were to invest 2 points in the project, you would keep 18 points for yourself.

Your earnings from the first stage depend on your decisions and the decisions of the three other group members. You keep all points that you do not invest in the project. The points you contribute to the project will be multiplied by 1.2, i.e. increased by 20%, and the increased amount will be split uniformly among all four group members. For example, if you contribute 5 points to the project, they will be increased to 6 points and split among all 4 group members. Each group member (you included) therefore earns 1.5 points due to your contribution of 5 points. If another group member contributes 5 points to the project, then every group member will again earn 1.5 points. Contributions to the project will always be increased by 20% and the resulting amount will be split equally among all four group members. Every group member consequently receives the same income from the project. This income is calculated in the following way:

$$\textbf{Your income from the project} = 1.2 * \frac{\text{Sum of all contributions to the project}}{4}$$

Thus, if you invest in the project, you thereby increase the incomes of the other participants. Conversely, if other group members contribute to the project, they increase your income. The points you keep for yourself do not increase other participants' incomes.

Your earnings from the first stage of a period are given by the sum of the points you have kept for yourself and the points you receive from the project:

Your earnings during the first stage =
points kept for yourself + earnings from the project

Two examples should clarify this calculation for your earnings:

- If all group members keep their endowment of 20 points for themselves, everyone earns 20 points from the points kept and nothing from the project. Thus, everyone gets 20 points.
- If all group members invest their entire endowment of 20 points in the project, the sum of all contributions will add up to 80 points. This amount will then be multiplied by 1.2, i.e., increased by 20%, and the resulting 96 points will be split among four group members, i.e., everyone will earn 24 points from the project and nothing from the points kept. Thus, everyone gets 24 points.

After you and the other three group members have made your decisions, you will receive information about the sum of contributions (including your contribution) all group members have made to the project on the following earnings screen. Furthermore, you will be informed of the points you earned in the first stage.

Summary of the first stage:

Second stage:

In the second stage, you will receive information about how many points the other group members invested in the project. In addition, you can reduce other group members' income by assigning deduction points in the second stage. The other group members can also reduce your income. If you assign 0 deduction points, you will not change the income levels of the other members . This will become clear immediately from the following screen for the second stage:

Screen at the second stage

Here you can see how much each group member has invested in the project. The earnings of a group member are stated below the member's contribution. Your contribution to the project is always in the first column, written in blue. The contributions of the other group members are listed in the three remaining columns. Please note that the labels group member 1, 2, and 3 will not refer to the same participant in every round. A different group member will be labeled as the second group member in rounds 1, 2 and subsequent rounds.

You must decide for each group member whether to assign deduction points and, if so, how many you want to assign to this participant. You can assign between 0 and 20 deduction points to each group member. Please enter your desired point level into the designated field.

You reduce the earnings of a group member by 1 point per deduction point you assign. However, the assignment of deduction points also reduces your earnings. Each deduction point you assign reduces your earnings by 1/2 point. If, for example, you assign 2 deduction points to another group member, his or her earnings will be reduced by 2 points and you will have costs of 1 point. If you additionally assign 18 deduction points to another group member, you will reduce that member's income by 18 points, while you will have costs of 9 points. If you assign 0 points to another group member, neither that member's earnings nor yours will be reduced.

If you are assigned deduction points by other group members, your earnings will be reduced accordingly. If, for example, you earned 15 points in the first stage and the other group members were to assign 12 deduction points to you in the second stage, your earnings would be reduced by 12 points. However, your earnings cannot be reduced by more than what you have. This means that, for instance, if your first stage income was 8 points but you were assigned 12 deduction points, your earnings would be reduced by 8 points. The group member who assigned 12 deduction points to you would still have to pay the entire costs of the 12 deduction points, amounting to 6 points.

To make the final decision concerning the assignment of deduction points, you must click on the "OK" button again. As long as you do not click on the "OK" button, you can revise your decision.

Your entire earnings for the two decision stages are calculated as follows:

Earnings of the round = earnings at the end of the second stage

= earnings from the first stage
- sum of deduction points assigned to you
- costs of deduction points you assigned.

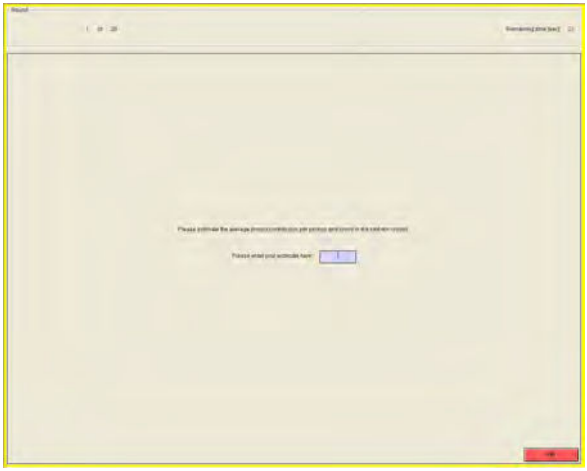
Your earnings at the end of the second stage if the deduction points you were assigned are equal to or higher than your first-stage income
= 0 - costs of the deduction points you assigned

Please note that your earnings at the end of the second stage can be negative if the costs of the deduction points you assigned exceed your earnings from the first stage minus the sum of the deduction points assigned to you. After all participants have made their decision, your earnings for the round will be shown on the following screen:

Earnings screen at the end of the second stage:

Your expectations:

Before the beginnings of the 1st and the 11th rounds, you will be called upon to make an estimation of the average contribution of group members. You must make a guess about how much the other three group members will on average invest in the project per round in the next 10 rounds. Since the other participants also receive an endowment of 20 points per round, your estimated value accordingly must range from 0 to 20. This estimation has no impact on your payoff. For this purpose the following screen will be displayed:



After the experiment, we will ask you to fill out a short questionnaire. Your earnings will then be paid to you in cash. The exchange rate is 30 points for 1 euro.

References

Ahn, T.K., Ostrom, E., Schmidt, D., Shupp, R., and J. Walker, 2001. Cooperation in PD games: Fear, greed, and history of play. Public Choice 106, 137-155.

Bohnet, I., and S. Huck, 2004. Repetition and reputation: Implications for trust and trustworthiness when institutions change. American Economic Review Papers & Proceedings 94, 362-366.

Brandts, J., and D.J. Cooper, 2006. A change would do you good ... An experimental study on how to overcome coordination failure in organizations. American Economic Review 96, 669-693.

Bruttel, L., and T. Friehe, 2010. On the path dependence of tax compliance. TWI Research Paper No. 59.

Carpenter, J.P., 2007. The demand for punishment. Journal of Economic Behavior and Organization 62, 522-542.

Cason, T.N., Savikhin, A., and R.M. Sheremeta, 2009. Cooperation spillovers in coordination games. Unpublished working paper.

Chaudhuri, A., 2011. Sustaining cooperation in laboratory public goods experiments: A selective survey of the literature. Experimental Economics 14, 47-83.

De Bondt, W.F.M., and R. Thaler, 1985. Does the stock market overreact? Journal of Finance 15, 793-805.

Fehr, E., and S. Gächter, 2000. Cooperation and punishment. American Economic Review 90, 980-994.

Fischbacher, U., 2007. z-Tree: Zurich Toolbox for Ready-made Economic Experiments. Experimental Economics 10, 171-178.

Gächter, S., Kessler, E., and M. Königstein, 2008. Performance incentives and the dynamics of voluntary cooperation. Unpublished working paper.

Gächter, S., and E. Renner, 2010. The effects of (incentivized) belief elicitation in public good experiments. Experimental Economics 13, 364-377.

Gneezy, U., and A. Rustichini, 2000. A fine is a price. Journal of Legal Studies 29, 1-17.

Greiner, B., 2004. An online recruitment system for economic experiments. In: Kremer, K. and V. Macho (Eds.): Forschung und wissenschaftliches Rechnen 2003. GWDG Bericht 63, Göttingen: Gesellschaft für Wiss. Datenverarbeitung, 79-93.

Hamman, J., Rick, S., and R.A. Weber, 2007. Solving coordination failure with "all-or-none" group-level incentives. Experimental Economics 10, 285-303.

Knez, M., and C. Camerer, 2000. Increasing cooperation in prisoner's dilemmas by establishing a precedent of efficiency in coordination games. Organizational Behavior and Human Decision Processes 82, 194-216.

Kube, S., and C. Traxler, forthcoming. The interaction of legal and social norm enforcement. Journal of Public Economic Theory.

Nikiforakis, N., and H.T. Normann, 2008. A comparative statics analysis of punishment in public-good experiments. Experimental Economics 11, 358-369.

Sefton, M., Shupp, R., and J.M. Walker, 2007. The effect of rewards and sanctions in provision of public goods. Economic Inquiry 45, 671-690.

THURGAU INSTITUTE
OF ECONOMICS
at the University of Konstanz

Hauptstr. 90
CH-8280 Kreuzlingen 2

Telefon: +41 (0)71 677 05 10
Telefax: +41 (0)71 677 05 11

info@twi-kreuzlingen.ch
www.twi-kreuzlingen.ch