Stefan Große

Experiments on the Fragility of Cooperation and Mechanisms to Overcome this Problem

Cuvillier Verlag Göttingen
Internationaler wissenschaftlicher Fachverlag

Experiments on the Fragility of Cooperation and Mechanisms to Overcome this Problem

Stefan Große

DISSERTATION

zur Erlangung des Grades
eines Doktors der Wirtschaftswissenschaften (Dr. rer. pol.)
der Universität Erfurt
Staatswissenschaftliche Fakultät

Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://dnb.d-nb.de abrufbar.
1. Aufl. - Göttingen : Cuvillier, 2011
Zugl.: Erfurt, Univ., Diss., 2011

978-3-86955-817-2

Nonnenstieg 8, 37075 Göttingen
Telefon: 0551-54724-0
Telefax: 0551-54724-21
www.cuvillier.de

1. Auflage, 2011
Gedruckt auf säurefreiem Papier

978-3-86955-817-2

To Franziska and my parents.

Acknowledgements

First and foremost I would like to thank my supervisor and coauthor Bettina Rockenbach for her advice, financial support, the many contributions, discussions, her time and the patience especially for the time when I changed my job which postponed the finalisation date of my thesis significantly and much more than I expected.

Furthermore I am much obliged to Simon Gächter for his ideas, his contributions, his advice and the nice week I had in Nottingham. He is also responsible for having ignited my interest in experimental economics by his fantastic lectures in St. Gallen.

I thank my coauthor Louis Putterman for his contributions, his insights and his always instantaneous responses.

I would like to thank Bernd Irlenbusch for allowing me together with Özgür Gürerk and Bettina Rockenbach to use their data on the endogenous public good experiment as benchmark data and furthermore for the discussion of the results of the experimental data for Chapter 4 which was my first experiment.

I am grateful to my former colleagues Irenaeus Wolff, Arne Weiss, Mareike Hoffmann, Özgür Gürerk, Thomas Dimpfl and Thomas Lauer for a magnificent time in Erfurt. They have supported me in so many regards that its is hard to write it down in only a few words. They helped me to program, test and conduct the experiments, to discuss the results, to motivate me, they supported me with the formalities and not to forget to chill out, to do sports and to enjoy the time in Erfurt. Irenaeus, Arne and Mareike proof-read the thesis.

Andrea Bäcker helped me much with the experiments as well. Her many comments helped to improve the programs and to understand the subjects. I am grateful to Torsten Huck, Mario Gruppe as well as Mandy Behrens (formerly Puschel) for their support for the experiments, especially in the last minute acquisition of subjects and the billing. I owe Kai Ahlborn for many philosophical and political discussions, my former colleagues Christian Proaño-Acosta and Pierre-Andre Gericke for the time together.

I would like to thank Enrico Schumann for pointing me at the memoir class in LaTeX, some LaTeX tricks, some hints on econometrics as well as proof-reading.

I am much obliged to Karim Sadrieh for inviting me to Magdeburg for inspiring MaXLab presentations, for his discussion, his humour, his entertaining stories and his cooking excellence.

I am thankful to Christiane Pilz who helped me numerous times in the struggle with administrative tasks, for the ability to cheer me up and the fantastic cakes she was treating us with. Not to forget for being a good proof-reader for the instructions.

I am grateful to Robert Jung for his advice on the hurdle model, Dennis Dittrich and Manfred Königstein for their econometric suggestions.

I am appreciative to my superiors at the NORD/LB, Torsten Windels and Jens Kramer, for their support, my colleague Norman Rudschuck for proof-reading. I am overgrateful to Tobias Basse for always reminding me how hard it is to write a PhD-thesis.

I like to thank the participants of my session during the GfEW conference in Cologne in 2005 most prominently Reinhard Selten as well as the participants of the workshop at the Max Planck Institute Jena in 2006 – I remember Dennis Dietrich, Werner Güth, Matthias Sutter – for their comments. Furthermore I am grateful to Tobias Heldt for his invitation to the Dalarna University Borlänge and his colleagues for their feedback.

The software I was using probably deserves acknowledgment as well: The R language for the graphics and most of the non-parametric tests as well as the hurdle models, Stata® for most of the panel regressions, Matlab®'s Symbolic Math Toolbox™ and of course LaTeX and all it contributors for the typeset of this thesis and JabRef as bibliographic database.

Last but not least I am indebted to Franziska Rippin, for supporting me all the years through all my moods, putting aside her demands for holidays and dates, always believing in me; without her I it would have been much harder if not impossible.

Contents

List of Figures

List of Tables

Overview

The human race is unique in the level of cooperation among individuals: We as *homines sapientes* work together in large groups even when we are not genetically related (Trivers, 1971). This is also a puzzle for classic economic theory. It is based on the model of the *homo oeconomicus*, who only maximizes his own pecuniary interest. This model, however, has been rejected in many experimental studies. Crucially, it fails to predict behavior of aggregates of individuals in important economic settings. This can be shown most impressively by a simple experiment called the ultimatum game created by Güth et al. (1982). In this experiment one person is asked to propose an allocation of an amount of money, say € 100, between herself and another unknown person. This other person has the possibility to reject the offer. If the offer is rejected, no one gets anything; if not, both are paid according to the offer. The *homo oeconomicus* would offer the smallest possible amount, say € 1, and a *homo oeconomicus* would accept that offer because € 1 is better than € 0. In the laboratory however, offers below a certain threshold are rejected because they are perceived as not fair. This is valid even across cultures as anthropologists have shown (Henrich et al., 2005).

There are several approaches to improve the economic theory and to incorporate more aspects of social interaction. They mostly try to 'repair' the utility function for example by adding other-regarding preferences. Most prominent is perhaps the model of Fehr and Schmidt (1999), which adds linear factors for profit relations. There are other models improving the situation but the disadvantage is that it gets harder to use more complex functions for the prediction of behavior. The revision of economic theory is at its very beginning and not the focus of this thesis.

Although we observe cooperation frequently in everyday life, there are situations where it works and certainly situations where it breaks down: – take the example of an apartment-sharing community in which the kitchen gets more and more messy; or, to a larger extent, the "tragedy of the commons" (Hardin, 1968).

The question is now when does cooperation break down and what conditions are necessary to achieve a sustainable cooperation. Exactly this question is at the center of this thesis and will be experimentally investigated. Experiments have the advantage of offering a large degree of control over the environmental conditions while in real life situations it is often hard to distinguish what had been the cause of a certain behavior. Cooperation is frequently modeled by a so called public-good design (synonymously, a voluntary contribution mechanism). This design tries to depict the problem that cooperation is beneficial for every participant but that there are individual incentives to free-ride, i.e., to put no personal effort to increase the common welfare. It does so in a mathematically rather simple model: Usually there is a common project into which all subjects might invest a part of their endowment. All contributions to this project are multiplied by a factor larger than one but smaller than the number of participants and divided equally among all group members. In this case it is better for the group to put everything into the project because the overall welfare is maximized. But since the return on investment is smaller than one – which means that if one Euro is spent and if no one contributes the return is less than one Euro – there is no incentive to invest.

In *Chapter 1* we demonstrate that the cooperation modeled as a contribution to a group project breaks down over the course of the experiment if we replicate such a standard public-good design. We show that it even gets worse if we introduce dynamics – if the next period's endowment consists of the current period's profit. This is one of the first experiments that analyzes this kind of dynamics. Although the potential payoff could be much higher compared to a linear setting, cooperators are exploited much easier since they could lose most of their resources at once which leads to a much faster collapse of cooperation. One reason for the breakdown of cooperation in public-good experiments can be conditional cooperation which even does not require the presence of absolute free-riders (e.g. Fischbacher and Gächter, 2010; Fischbacher et al., 2001).

In *Chapter 2*, which is part of Grosse et al. (2011), we try to mitigate the public-good problem, which we frame as team production, by a monitoring mechanism. The monitoring mechanism can be thought of as a technology that makes the individual efforts of team members attributable. If the contribution of each individual is attributable, we argue, they will contribute as the attribution transfers the public-good into a private-good. The investment in the mechanism could be done by the team members themselves or could be "outsourced" towards an observer who is a person close to the team but not a team member. Since the monitoring investment decision is more or less a second-level public-good (the investment is costly and everyone benefits equally if the cooperation should be established) – to be specific in our case a step-level public-good – we expected that the team members will fail to sufficiently invest in the technology. In consequence they should opt for the observer, who makes the decision to monitor, even though that is somewhat costly as well as the observer will claim a share of the teams production. The construct of a team production with a residual claimant shall resemble the Alchian and Demsetz (1972) type of a capitalist firm.

The outcome is surprising: The self-monitored firm is highly preferred by the subjects. This has several causes. One problem is that the construction of the monitoring as a step-level public-good in Chapter 2 has the consequence that there exist positive contribution equilibria: for a pivotal team member the investment decision in monitoring is profitable. By pivotal we mean that a single subject might be the one subject that moves by its investment all others from not-contributing to contributing.

In order to test the positive contribution equilibria we alternate the design of the experiment for *Chapter 3*. First, we change the monitoring decision in itself from a step-level public-good that implies a coordination problem to a linear public-good and furthermore we apply a quadratic cost function which is necessary to prevent the threshold where the optimal contribution would be the maximal instead of the minimal contribution. We can show that this leads to the success of the residual claimant. A part of that chapter has been published in Grosse et al. (2011).

Another mechanism that has been proven to be successful in the prevention of free-riding behavior is to allow for sanctions, most often in a decentralized way (for example Fehr and Gächter, 2000a). When cooperators are allowed to individually sanction non-cooperators they are able to raise the overall contribution level. This "works" quite well in order to ensure a high cooperation level but gives rise to an efficiency issue. The sanctions are costly. It is even possible that the overall welfare in a public-good experiment with sanctioning possibility is lower than compared with a standard public-good game even with free-riding behavior present (Herrmann et al., 2008). It has been proven as successful to give the subjects the choice whether they want to join a group that has a sanctioning possibility or a group without that (Gürerk et al., 2006, 2010a) in order to enhance the efficiency in a sanctioning environment.

One phenomenon that still is an issue is that of "anti-social" punishment even though the efficiency with the endogenous institution choice might be an improvement. Some subjects punish especially maximum contributors. Those sanctions are socially not desirable since the welfare maximum contributors behave optimal for all (i.e., are "pro-social").

In *Chapter 4* we create a design that by the introduction of a simple rule shall measure the combined detrimental effect of the "anti-social" punishment in an endogenous-institution-choice design. With combined detrimental effects we mean that beside the direct costs of punishments there could be revenge effects and furthermore it could be that uncontrolled sanctioning leads to high contributors decreasing their contributions after being punished. We only allowed sanctions towards subjects that contributed less than the sanctioning subject. By this rule we intended increase the efficiency and thus the attractiveness of the punishment institution by preventing not only "anti-social" but most (direct) revenge activities. The punishment institution was avoided by most subjects in the beginning of the experiment of Gürerk et al. (2010a). We compare our data with the punishment constraint with the data of Gürerk et al. (2010a). The introduction of the punishment limit has a great impact by raising the efficiency and making the sanctioning institution more attractive, not in the first period but in the course of the experiment. There are far fewer sanctions necessary to achieve a high contribution level even compared with Gürerk et al. (2010a) if one subtracts the "anti-social" sanctions.

1

Dynamic Public Goods[1]

Abstract

We conduct a dynamic public-good experiment: The subjects receive an endowment only in the first period of the experiment. After that the subjects' endowment consists of the previous periods' payoff. Although the potential benefit through the dynamics is huge the subjects fail massively to contribute. This can be attributed mostly to cooperative subjects running out of resources.

1.1 Introduction

A commonly used tool in experimental economics research on cooperation is the public-good game, also known as voluntary contribution mechanism (VCM).[2] In many variations of VCM experiments in a multi-period setting, the subjects receive a fresh endowment every period. The subjects thus are allowed to repeat their decision with fresh resources; a decision in one period does not have the consequence of potentially losing resources in the next period – given the anonymity of the participants within the experiment and no kind of a reputation building mechanism. But in some settings it might be interesting whether it makes a difference if a decision in one period has consequences for all other periods. As a real life example, think of a small society in an agricultural context: If such a small community is cooperating successfully in one year and by cooperating is having a plentiful harvest, then this improves its survival probability and furthermore possibly enables it to invest in technology which in turn could improve the harvest in the next year. Until now surprisingly little research has been conducted on comparable dynamics especially in the

[1]This part is based on a joint project with Simon Gächter, Torsten Huck, Bettina Rockenbach.

[2]See Ledyard (1995) for an older, Gächter and Herrmann (2009) as well as Chaudhuri (2011) for more recent reviews.

public-good context. We will give an overview on related literature in the next section.

Unlike a standard public-good experiment with a constant endowment over all periods we created a dynamic variant wherein the subjects receive an initial endowment only in the beginning of the experiment. Each subjects' endowment from there on will equal the profit of the previous period. Thus there could be a remarkable growth and a huge benefit if all subjects contribute all the time. However, there is a free-riding incentive and the danger of losing almost everything.

To decompose the impact of the dynamics we first compare the initial treatment with a treatment where subjects are randomly assigned to an endowment history of the dynamic public-good treatment and add two other treatments – one standard public-good experiment and a public-good experiment with equal but per-period increasing endowment.

We find indeed an effect of the dynamics. Contributions deteriorate in the course of the dynamic experiment much faster compared with the other treatments. We attribute this detrimental development mostly to cooperators running out of resources while we observe only minor differences in the behavior of the subjects between the dynamic and the other three treatments.

1.2 Related Literature

Dynamics

The term "dynamic" in the public-good context can be interpreted in different ways. Fershtman and Nitzan (1991) for example theoretically investigated the effect of an exogenously introduced decay in a continuous public-good problem with infinite duration where contributions are aggregated. The decay makes it necessary to provide a certain contribution level for it to survive and the contributions were added up over the time. Thus it resembles more some kind of a threshold public-good but with a dynamic component. They found that under certain circumstances an equilibrium might emerge but if the subjects act conditional the contribution will deteriorate. Noussair and Soo (2008) introduced some dynamics by letting the marginal per capita return change in a step-wise fashion if a certain threshold of the sum of contributions (50%) was met. In this setting only a minority of groups with deteriorating contributions emerged but there was a great heterogeneity among the group contribution paths.

In our interpretation dynamic means that the subjects have to live with the resources (i.e., endowments) created by their payoffs during the experiment. Most closely related in this regard is Gürerk et al. (2010b). Like in our dynamic experiment the subjects in this study received an initial endowment that was equal in the first period. The profit of one period was the endowment of the next period. The difference to our experiment is that the subjects' total payoff of Gürerk et al. (2010b) only consisted in the profit of the last period while in our experiment it is the sum of all period-profits. The authors compared a dynamic public-good treatment without punishment option with a dynamic treatment with a punishment option included. They observed a decreasing trend of contributions relative to the endowment without the punishment option contrasting with the sanctioning treatment where the contributions remained at the same level. The detrimental effects of the punishment – i.e., efficiency losses – are slightly more than compensated by the benefits of the induced higher contributions leading to almost equal final wealth levels.

Another strand of the experimental literature that could very roughly be attributed to our definition of dynamics is that which combines a real effort with an experiment wherein the profit of the real effort is connected to the later endowment. One paper in this regard would be Cherry et al. (2005). In their experiment the subjects' endowments for a one-shot public-good game were earned by taking a GMAT questionnaire beforehand. They found no impact of the source of the endowment heterogeneity – whether it was endogenously earned or exogenously assigned. This is somewhat opposite to another paper of the same authors – Cherry et al. (2002) – who found a strong impact in a dictator game context: If the endowment was self-earned the subjects behaved much more selfish.

Heterogeneity

Our experiment leads in one treatment to within-group differences in the endowments, in one treatment caused by the actions of the subjects (endogenous heterogeneity) and in another treatment because the subjects will follow a randomly assigned endowment history of the dynamic treatment (exogenous heterogeneity).

There is plenty of literature on the heterogeneity of endowments[3] with somewhat mixed results. One can distinguish between those experiments that assign heterogeneous endowments with or without the influence of the subjects. We will refer to the latter as exogenous heterogeneity. In fact most experiments on heterogeneity can be regarded as falling into that category. A few others introduce heterogeneity by different actions before the public-good actually was conducted – we call this endogenous heterogeneity. For the latter the aforementioned paper of Cherry et al.

[3]There can be of course also heterogeneity in the ability allowing some subjects to be more 'productive' than others.

(2005) would be an example. A special kind would be household money experiments. In those the people bring in their own money at the beginning of the experiment (Clark, 2002; Harrison, 2007, for example)).

Cherry et al. (2005) found lower contributions with exogenous as well as endogenous heterogeneity in the sense of earned heterogeneous endowments compared with homogeneous endowments in their one shot public-good setting. Contrary to that, Chan et al. (1999) concluded that (exogenous) heterogeneity increased the cooperation level in a non-communication environment. Their public-good design was slightly unusual and the heterogeneity was differently distributed in comparison to Cherry et al. (2005). Levati et al. (2007) observed mixed results in a public-good in presence of a leader who was inducing higher contributions with (endogenous) heterogeneity than compared with homogeneity in the endowments. Anderson et al. (2008) reported a negative impact of (endogenous) inequality of endowments.

In the context of step-level public-goods we also discovered mixed results in the literature. While van Dijk and Wilke (1995) observed that participants with different endowments contributed the same percentage of the endowments – which was further support to van Dijk and Grodzka (1992), Aquino et al. (1992) wrote that in their experiments the cooperation was lower with higher resource inequality – independent of the framing. Bagnoli and McKee (1991) and Rapoport and Suleiman (1993) reported evidence for inequality reducing the contributions to the group account while on the contrary Marwell and Ames (1979), Marwell and Ames (1980) did not discover a difference.

Sutter and Weck-Hannemann (2003) introduced some asymmetry by two different minimum contribution levels. Those levels were either externally imposed or the subjects could vote if the asymmetric rule is implemented. The externally set rule produced higher contributions.

Another article invoking a dynamic component is Sadrieh and Verbon (2006). Although they did not use the classical linear public-good design having a non-linear production function and only three choices (cooperate, sabotage or Nash) their design has the character of a public-good. An important additional difference is that the payoffs were determined by the subjects' share of the total production. They allowed the accumulation of endowments over 5 periods and furthermore tested different degrees of heterogeneity in endowments and different distributions of endowments (i.e., number of lowly-endowed subjects). The authors observed that the degree of heterogeneity had no effect on the contribution level and on the growth in the dynamic game while a static variant experienced a positive effect of a greater degree of heterogeneity.

Reuben and Riedl (2009) found no difference in their treatments with heterogeneous endowments without punishment although with punishment they seem to agree on certain social norms that are then enforced: Subjects with higher endowment should thus contribute more.

Summing up, results on the impact of heterogeneity are mixed. It seems to be very dependent on the context whether contributions are higher, equal or lower. We will contribute to this literature by adding insights on the dynamics of endogenous heterogeneity and compare them with exogenous heterogeneity.

1.3 Experimental Design

At the center of our experiments is the dynamic linear public-good treatment over 20 periods which we dub DYN. The subjects received in the first period a start endowment of $e_{i,t=1} = 20$ experimental units. They could contribute these to a public-good. The contributions were multiplied by the factor $a = 1.4$ and equally distributed among the $n = 4$ group members. The profit of the current round was then the endowment of the subjects in the next period:

$$\pi_{i,t} = \underbrace{\pi_{i,t-1}}_{endowment} - c_{i,t} + \frac{a}{n} \sum_{j=1}^{n} c_{j,t} \,. \qquad (1.1)$$

This is similar to Gürerk et al. (2010b). Our experiment differs to theirs in the regard that the wealth level that was relevant for the subjects' payment at the end of their experiment was $\Omega_i = \pi_{i,T}$ while ours was $\Omega_i = \sum_{t=1}^{T} \pi_{i,t}$. A second difference was that their a was larger (1.6).

Since we wanted to investigate the effect of the dynamics in the main treatment we separated the different factors that could have an effect on the contribution levels. Thus beside the DYN treatment we created the HIST treatment in which the heterogeneous endowment is exogenously introduced in as much as one subject was randomly assigned to a unique endowment history path of a subject of the DYN treatment. The CONST treatment is the control treatment and resembles the classical linear public-good experiment with the difference that the endowment was the average endowment of the DYN treatment over all periods. The INC treatment then deals with the issue that we observed per period increasing endowments in DYN. In consequence we created this treatment as control for the increasing endowment effect and assigned the per period average of DYN as the per period equal endowment for all subjects in this treatment. Table 1.1 summarizes the treatments.

The experiments were conducted at the University of Erfurt using z-tree (Fischbacher, 2007) for the experiment and ORSEE (Greiner, 2004) for the recruitment. The total number of participants was 160.

Table 1.1: Overview of the treatments.

	DYN	HIST	CONST	INC
Endowment e in $t=1$	$e^{\text{DYN}}_{i,t=1}=20$	$e^{\text{HIST}}_{i,t=1}=20$	$e^{\text{CONST}}_{i,t=1}=\bar{e}^{\text{DYN}}$	$e^{\text{INC}}_{i,t=1}=20$
Endowment e in $t>1$	$e^{\text{DYN}}_{i,t}=\pi^{\text{DYN}}_{i,t-1}$	$e^{\text{HIST}}_{i,t}=e^{\text{DYN}}_{i,t}$	$e^{\text{CONST}}_{i,t}=\bar{e}^{\text{DYN}}$	$e^{\text{INC}}_{i,t}=\bar{e}^{\text{DYN}}_{t}$
Profit $\pi_{i,t}=\Big\{$	$e^{\text{treatment}}_{i,t}-c_{i,t}+\frac{a}{n}\sum_{j=1}^{n}c_{j,t}$			
Independent observations	12	12	8	8

1.4 Theoretical Considerations and Predictions

Standard Theory

Traditional game theory with rational money-maximizers as players predicts zero contribution for all players as the subgame perfect Nash-equilibrium for all treatments. This follows from the fact that the marginal per capita return ($\mu = mpcr$) of a unit of contribution in period T, $\mu = \frac{a}{n}$, is smaller than one with our parameters (0.35); by backward induction, this holds for any other period. This results in the equilibrium with no contribution at all and holds for any of our treatments.

Inequity Aversion

Using the model framework of Fehr and Schmidt (1999) with their inequity aversion concept:

$$U(x_i) = x_i - \frac{\alpha}{n-1}\sum_{j\neq i}^{n}(x_j - x_i)^{+} - \frac{\beta}{n-1}\sum_{j\neq i}^{n}(x_i - x_j)^{+} \tag{1.2}$$

there are cooperative equilibria possible if certain conditions are met. Especially the analysis of CONST but also that of INC can be broken down to Proposition 4 of Fehr and Schmidt (1999, p. 839). There exist equilibria with conditional cooperative subjects if those do not suffer too much from disadvantageous inequality. But if there are not enough subjects with $\mu + \beta_i < 1$ then the cooperative players will also not contribute in absence of a punishment possibility.

The HIST and DYN treatment are a little different. With DYN a contribution early in the experiment affects the endowment of all subjects in all later periods. For an extensive game theoretic analysis of the dynamics of this approach see Gürerk et al. (2010b). Although the calculation of the wealth is slightly different the analysis for the inequity is quite similar.

The dynamic property of DYN has the implication that it can be optimal for the money-maximizer (*mm*) to contribute because an early contribution yields "compound interests" if others contribute in subsequent periods. Suppose the money-maximizer (*mm*) contributes nothing in every period and now increases his contribution in period t by $\Delta c_{mm,t} = 1$. In t this creates a loss of $1 - \mu$ which is why in a non-dynamic context a *mm* would contribute nothing. But given the dynamics and if there exist k conditional cooperators (*cc*) that fully contribute their whole endowment then the contribution of the *mm* will increase the endowment of the *cc*'s in $t + 1$ by μ. If we assume they contribute that additional endowment, the project payout in $t + 1$ will increase by $(k\mu)\mu = k\mu^2$. In $t + 2$ this additional payout will be again contributed and yields a project payout of $(k\mu)^2\mu$ and so on.

To generalize: the benefit of increasing $c_{mm,t}$ (only) in t by $\Delta c_{mm,t}$ until a period t' leads to a change in the period t_a, $t_a \in \{0, 1, \ldots, t'\}$ profits of $\Delta\pi_{mm,t+t_a} = \mu \cdot (\mu \cdot k)^{t_a}$ which leads to a change in the *mm*'s wealth of

$$\begin{aligned}\Delta\Omega_{mm,t'} = &- \Delta c_{mm,t} + \Delta c_{mm,t}\mu + \Delta c_{mm,t}(k\mu)\mu + \Delta c_{mm,t}(k\mu)^2\mu + \\ &\cdots + \Delta c_{mm,t}(k\mu)^{t'}\mu.\end{aligned}$$

in order to let the wealth increase the condition

$$\begin{aligned}0 \leq &- \Delta c_{mm,t} + \Delta c_{mm,t}\mu + \Delta c_{mm,t}(k\mu)\mu + \Delta c_{mm,t}(k\mu)^2\mu + \\ &\cdots + \Delta c_{mm,t}(k\mu)^{t'}\mu\end{aligned}$$

should be met. Dividing by $\Delta c_{mm,t}$ this further simplifies to

$$1 \leq \mu + (k\mu)\mu + (k\mu)^2\mu + \ldots + (k\mu)^{t'}\mu$$

and because this is a geometric progression we can reformulate this as

$$1 \leq \mu\frac{(k\mu)^{t_a} - 1}{(k\mu) - 1}\,.$$

This condition makes it easy to calculate the number of t' periods that are needed to make the contribution profitable:

$$t' = \log_{k\mu}\left(k - \frac{1}{\mu} + 1\right)\,.$$

With our parameters we get with $k = 3$ cooperators $t' = 2.73 \approx 3$ as the number of necessary periods and for $k = 2$ a $t' \approx 6$. So even with "only" two cooperators we might observe money maximizers that positively contribute until $T - t'$.

Of course it is necessary that the *cc* stick to full contribution. They will do so if their $\mu + \beta > 1$ as specified by Fehr and Schmidt (2000) and further if as in their Proposition 4:

$$\frac{k}{n-1} \leq \frac{\mu + \beta - 1}{\alpha + \beta}\,. \tag{1.3}$$

In this case full contribution is even more enforced due to the dynamics: Every unit additionally contributed in earlier periods increases exponentially the income potential. Thus it perfectly makes sense to increase the contribution up to the maximum – if the others do so as well. If a subject deviates from full contribution this creates huge losses as we already argued for the case of the *mm*. The interesting case is the end of the experiment. In the last t' periods the *mm* will not contribute. Condition 1.3 will make the *cc* not deviate from their previous contribution. The dynamics makes it even harder to deviate.

It gets more difficult if the contribution levels are not at maximum and not symmetric. If the subjects would for example not deviate from their initial contribution $c_{i,1}$ with $c_{j=1,1} < c_{j=2,1} < \ldots < c_{j=n,1}$ this means that the subject n with the maximum contribution would run out of resources if $\mu \sum c > e_n - c_n$. In this special case she will not be able to proceed to contribute with the highest absolute amount. This leads to the second most contributor being highest contributor, losing resources and so on. This would, unlike in a non-dynamic experiment, never be a stable equilibrium.

Gürerk et al. (2010b) mentioned an equilibrium that, given a certain number of *cc* and *mm* could lead to equity: The *cc* contribute nothing in the first period and from $t = 2$ mimic the contribution of *mm* in the previous period ($c_{cc,t} = c_{mm,t-1}$). Although this works also in our context, a very high α in the Fehr and Schmidt (1999) model framework would be needed to make this strategy credible for the *cc*'s because the conditional cooperators would destroy a lot of welfare for all participants by withholding endowments from the beginning, making the equalization very costly. The α must be very high to compensate for this.

To sum up: assuming inequity aversion in DYN we might observe cooperative behavior as long as there are enough cooperators with a sufficiently low α so that they care less enough about the deviator(s) but also sufficiently high β so that they maintain their cooperation level. In this case a money maximizer with $\mu + \beta < 1$ chooses to contribute fully until $T - t'$.

If we achieve a cooperative equilibrium in DYN this could mean for HIST that a *mm* could follow the endowment history of a *cc* and the other way round. The *mm* will not contribute in any case. The cooperative equilibrium in DYN means that in HIST there is no inequity in endowments until t' and the Proposition 4 of Fehr and Schmidt (1999) holds. Thus we might observe some *mm* not contributing and some *cc* contributing. If a *cc* in HIST was assigned to the endowment path of a *mm* in DYN she will have a relative high endowment after t'. Her likely contribution if Condition 1.3 is met represents a redistribution, reducing inequity in HIST compared to DYN.

Predictions

Although with social preferences we could observe high contributions we know that in a simple public-good game usually the contribution level is starting at a certain level and deteriorating over time (see Ledyard, 1995; Gächter and Herrmann, 2009; Chaudhuri, 2011). The decrease in the contributions can be largely attributed to imperfect conditional cooperation (Fischbacher and Gächter, 2010; Fischbacher et al., 2001; Keser, 2000). Since the characteristics (except the dynamics) are similar we would expect different contribution levels in the beginning and falling average contributions afterwards as well.

Hypothesis 1. Due to free-riding incentives and imperfect conditional cooperation we expect decreasing contribution levels in all treatments.

There is theoretically a very high benefit of cooperation in the DYN treatment. If all subjects would cooperate their period profit in t is: $\pi_{i,t} = 20\frac{1-1.4^{t+1}}{1-1.4} - 20$ so alone in the last period T the profit would be $\pi_T = 20\frac{1-1.4^{21}}{1-1.4} = 58,497.78$ per subject and the total wealth $\Omega_i = \sum_{t=1}^{T} \pi_{i,t}$ even considerably larger, compared with a $\Omega_{\text{Nash}} = 20 \cdot 20 = 400$. On the other hand, the presence of a single subject with $c = 0$ from $t = 1$ on would be enough to destroy most of the potential gain. We argued in the previous section that this behavior might not be rational but if it is present it will cause serious damage to the cooperation levels. Additionally in case others do contribute much less than maximum, a full cooperator might lose her resources, not being able to contribute again. Thus there is a possibility that subjects are more cautious in the DYN treatment and contribute less from the very beginning. In any case we expect a stronger dynamic toward zero contribution with the DYN treatment than with any other treatment due to the conditional cooperation being triggered even by cooperators running out of resources.

Hypothesis 2. (a) In presence of imperfect conditional cooperation the DYN treatment will more quickly tend toward zero contribution than any other treatment due to either more cautious cooperators or cooperators running out of resources. (b) The cautiousness will lead to lower contributions in DYN at the beginning of the experiment since the threat of exploitation is anticipated.

DYN will create a certain degree of heterogeneity in the endowments that is caused by different conditional-contribution behavior between the subjects. As a consequence the subjects in HIST will experience the same degree of endowment heterogeneity with the difference that they have no influence on their endowment making the heterogeneity exogenous.

Although the literature is not consistent in this regard we support the view of Chan et al. (1999) and assume that HIST should show lower contribution levels compared with CONST. Additionally we would expect a lower overall contribution level if the heterogeneity can be attributed to the subjects' contribution history at least because with exogenous heterogeneity a cooperator has more resources to contribute in the course of the experiment.

Hypothesis 3. We expect lower contribution levels in the HIST treatment than in CONST due to the presence of exogenous heterogeneity but higher contribution levels than in DYN because cooperators will not loose resources in the HIST treatment.

1.5 Results

1.5.1 Dynamics

Figure 1.1 on page 12 reveals the usual picture of a standard linear public-good experiment: Every single treatment shows a negative development in the contribution as a percentage of the endowment. This is supported by the linear mixed-effects regression explaining the per-period and per-group aggregated group contribution level (Table 1.2) by the period number, the squared period number, the optional number of free-riders in the previous period and an optional measure for heterogeneity. The coefficient of the *period* is negative and with one exception significant for all treatments. While the contribution generally deteriorates in the course of the experiment it does so to a much higher extent in the DYN treatment: The coefficient of $period^2$ can be interpreted as a higher curvature in this treatment compared with HIST and CONST. Furthermore we observe a weakly significant difference in the contribution (in percent of the endowment) averaged over all periods between DYN and HIST ($p = 0.078$) as well as between DYN and CONST ($p = 0.0387$ respective $p = 0.0831$ exHC[4]) with a two-sided exact Wilcoxon–Mann–Whitney test (see Table 1.3 for the cross-comparison).

Result 1. *We observe a negative trend in the contribution level over all treatments but the dynamic treatment depicts a higher curvature than the other treatments: The contribution levels deteriorate thus faster than in the other treatments. The overall contribution level is weak significantly lower in DYN than in HIST and CONST.*

[4]Note that we have one exceptional group in the CONST treatment that maintained nearly full cooperation. This is why we compare this treatment mostly separate once with all groups together and on time with those high contributors excluded (CONST ex HC).

Table 1.2: Linear mixed-effects regressions explaining the mean group contribution in percent of the endowment (with bootstrapped standard errors). *L_freerider* is the number of the freerider in $t-1$ and *theil* stands for the Theil-Index of inequality.

	DYN				HIST				CONST				INC	
	1	2	3	4	5	6	7	8	9	10 (exHC)	11	12 (exHC)	13	14
const.	0.4809***	0.4839***	0.4734***	0.4686***	0.4547***	0.4554***	0.4598***	0.4569***	0.3753***	0.3800***	0.3486***	0.3436***	0.5476***	0.5453***
	0.0193	0.0203	0.0257	0.0270	0.0207	0.0213	0.0289	0.0296	0.0399	0.0284	0.0425	0.0355	0.0231	0.0264
period	−0.0520***	−0.0504***	−0.0490***	−0.0500***	−0.0362***	−0.0358***	−0.0348***	−0.0353***	−0.0095	−0.0244***	−0.0045	−0.0173**	−0.0499***	−0.0502***
	0.0038	0.0042	0.0050	0.0049	0.0047	0.0049	0.0061	0.0061	0.0084	0.0061	0.0087	0.0071	0.0047	0.0050
period²	0.0016***	0.0016***	0.0015***	0.0015***	0.0011***	0.0011***	0.0011***	0.0011***	−0.0001	0.0005*	−0.0003	0.0003	0.0014***	0.0014***
	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0003	0.0003	0.0004	0.0003	0.0004	0.0003	0.0002	0.0002
l_freerider			0.0017	0.0030			−0.0268**	−0.0270**			−0.0073	−0.0093		0.0194*
			0.0057	0.0055			0.0126	0.0124			0.0112	0.0112		0.0105
theil		−0.4596*	−0.3764			−0.1152	−0.2131							
		0.2519	0.2717			0.2978	0.3254							
LogLik	251.40	253.07	237.77	236.94	176.97	176.74	162.78	162.82	87.08	108.03	85.35	100.75	175.58	168.37

Bootstrapped standard errors (1000 reps). Significances: *** 1%, ** 5%, * 10%

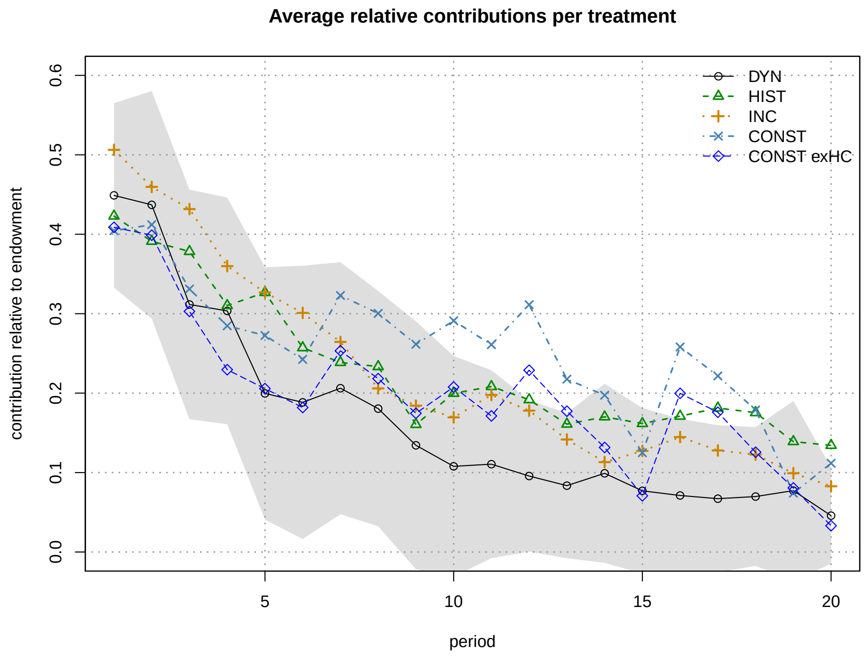

Figure 1.1: Average contribution relative to the endowment per treatment. The grey-shaded area represents the standard deviation of the group means of the DYN treatment.

Table 1.3: P-values of two-sided exact Wilcoxon–Mann–Whitney tests between the treatments over the average relative contribution. Significances: *** 1%, ** 5%, * 10%.

	DYN	HIST	INC	CONST
DYN	-			
HIST	0.0780*	-		
INC	0.1349	0.9699	-	-
CONST	0.0387**	0.8506	0.7984	-
CONST exHC	0.0831*	0.4824	0.4634	-

Interestingly there is no significant difference for the first period's contribution in percent of the endowment with the exact Wilcoxon–Mann–Whitney test in all treatments as depicted in Tables 1.6 and 1.7 (Appendix). The distribution (Figure 1.2, Appendix) of the contribution levels in that period reveals no apparent difference either, except the little spike on the right tail of the distribution which is absent in CONST but is distinctive especially in INC and DYN. However, the differences in the distribution are not significant.[5]

Thus we must reject Hypothesis 2b. Eventually the potential benefits

[5]We conducted several proportion (χ^2) tests, calculating the counts of contribution percentages between the treatments. We got no significances for several proportion tables. For example for ($c < 45\%$, $45\% \leq c < 55\%$, $c \geq 55\%$) which yields the counts 17, 10 and 5 for CONST and 20, 18, 10 for DYN we cannot reject the equal proportion hypothesis ($p = 0.5954$).

Table 1.4: Distribution of the number of above-average contribution in percent of the endowment per subject in the first 10 periods to check for the persistence of the cooperative behavior. Read example: in DYN 6.3% of the subjects have contributed 8 times above the average in the first 10 periods.

Periods	DYN in%	HIST in %	INC in %	CONST in%
10	2.1	0.0	0.0	3.1
9	2.1	2.1	6.3	0.0
8	6.3	8.3	12.5	6.3
7	16.7	6.3	6.3	18.8
6	12.5	20.8	3.1	6.3
5	10.4	18.8	18.8	9.4
4	8.3	16.7	21.9	25.0
3	20.8	10.4	12.5	21.9
2	6.3	10.4	12.5	9.4
1	12.5	6.3	6.3	0.0
0	2.1	0.0	0.0	0.0

just balance with the fear of the exploitation. The "tail behavior" meaning a slightly higher number of above average contributors in DYN is an indication in that direction but there is no statistical evidence however.

Result 2. *There is no significant difference in the contribution levels of the first period between the treatments. We must reject Hypothesis 2b.*

We believe that there could be at least two reasons for a faster contribution deterioration in the DYN treatment. One could be that cooperative subjects run out of resources; the other that high contributors adjust more quickly to lower contribution levels than in the other treatments. We choose two approaches to check for this: First, we compare the behavior of the subjects in the first periods between the and second, conduct further extensive econometric analysis of both group and individual behavior.

In the first five periods we did not observe a significant difference in the contribution in percent of the endowment as we see in the per-period tests of Tables 1.6 and 1.7 (Appendix). The significant differences in contribution levels start from period 10 on. If we count the number of above-average contributions in percent of the endowment per treatment and per subject in the first 10 periods (Table 1.4) in order to check for the persistence of the cooperators – we do not find a significant difference between the treatments (Fisher-Exact test $p = 0.3283$ between DYN and HIST and $p = 0.3961$ for DYN vs. CONST). So we can not say that the cooperators stopped significantly earlier to cooperate in the sense of contribute a high percentage of the endowment in DYN than in the other treatments.

The difference thus is more that by design. In the DYN treatment the cooperators dry out of resources. If we take the average endowment rank

(1 lowest, 4 highest) in the beginning of the 11^{th} period for those subjects that contributed more than average for more than 6 times in the first 10 periods (27% of all subjects in DYN, 17% in HIST, 25% in INC and 28% in CONST, cumulating the first rows of table 1.4) this becomes clear: The cooperators i.e., the subjects that contributed more often above the average indeed have a significantly ($p = 0.01846$, exact Wilcoxon-Mann-Whitney test) worse endowment in DYN compared to HIST. With other words: The cooperators significantly worsen their relative wealth level (as rank) in the first periods and, as a consequence, are not able to contribute

The econometric analysis of the change in contributions relative to the endowment (Table 1.9, Appendix) reveals another factor. The subjects of DYN seem to decrease their contribution level slightly (not significantly) less if they are contributing more than the average (relative to the endowment), but if the subjects contribute below the average, their adjustment is significantly lower than in HIST, CONST and INC. In all treatments the subjects on average increase their contribution level if they are below average. One could regard this as a mean-reversion tendency. In other words: Below-average contributors in DYN are more likely to stay below the average than the subjects in the other treatments.

Result 3. *The stronger decrease of the contribution level in the DYN treatment stems from two different effects. Firstly: More cooperative subjects run out of resources – which they do not in the other treatments. Secondly: Below average contributors are more likely to stay below the average than in the other treatments.*

Although the number of free-riders in the regression of Table 1.9 (Appendix) shows no effect on the contribution level change in DYN, it does have a weakly significant, small impact in other treatments. The appearance of zero contributors seems to have an influence only on the free-riding decision itself (Table 1.10, Appendix). Hence there is some evidence for a two-step decision. In general, free-riding, although most frequently observed in the DYN treatment (23, 24%), seems not to be significantly different from other treatments except for INC which seems to have weakly significant less frequent free-riding (Table 1.5).

1.5.2 Heterogeneity Effect

The Theil-Index of inequality (Theil, 1967): $T_1 = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{x_i}{\bar{x}} \cdot \ln \frac{x_i}{\bar{x}}\right)$, is a dispersion measure comparable to the better known Gini-Coefficient but with the advantage that it is decomposable. The econometric analysis does not reveal a significance for this measure neither in the regression of the contribution changes (Table 1.9, Appendix), nor in the estimate of the per-group aggregated contribution level (Table 1.2). So we can at

Table 1.5: Two-sided exact Wilcoxon-Mann-Whitney test between the treatments over the count of zero contributions (free-riding), *p*-values and the percentage of occurrences of zero contributions. Significances: *** 1%, ** 5%, * 10%.

	DYN	HIST	INC	CONST	Prevalence (%)
DYN	-				23.24
HIST	0.1380	-			13.64
INC	0.0590*	0.6633	-	-	12.03
CONST	0.8650	0.1618	0.0883*	-	19.07
CONST exHC	0.9172	0.1038	0.0502*	-	18.43

least reject that endogenous and exogenous heterogeneity do have a direct (linear) influence on the size or level of contribution. We observe an effect of the Theil-Index on the decision to free-ride (Table 1.10, Appendix) but only for the DYN treatment. This could mean that if the heterogeneity in endowments is endogenous the subjects more likely tend to free-ride but consider the actual contribution choice to a lesser extent.

For the within-group heterogeneity another interesting effect might be the behavior of the 'rich' subjects – subjects that have the highest endowment in their group. But our econometric analysis does not show any significance for 'rich' – they do not necessarily contribute differently nor have a higher probability to free-ride (Table 1.9 and Table 1.10, Appendix).

Result 4. *The heterogeneity in itself is has not enough explanatory power for the differences between DYN and HIST. Regressions show that there is only little influence of Theil's inequality measure. Additionally there is no significant difference in the behavior of the subjects with the highest endowment between the treatments.*

This is also further supported by the non-parametric analysis which neglects any effects on aggregated contribution levels of the exogenous heterogeneity – HIST treatment – except in direct comparison with the DYN treatment: Neither the per-period exact Wilcoxon-Mann-Whitney test (Table 1.6 and Table 1.7, Appendix) nor the same test for the per treatment average relative contribution over all periods (Table 1.3) reveal any significance. Only DYN shows weak differences in the contribution levels which are as we argue more due to the running out of resources than to the heterogeneity itself.

1.5.3 Endowment Effect

CONST has a higher endowment in the first period than all other treatments. We did not find significant differences in the contribution in percent of the endowment conducting the exact Wilcoxon-Mann-Whitney test (Table 1.6 and Table 1.7, Appendix) of this period between CONST and the other treatments. Thus we can clearly reject a general between-treatments endowment effect. An increasing homogeneous endowment as with INC, DYN and HIST seems to have the consequence that the subjects maintain a more or less constant absolute contribution level: in those treatments the subjects started to contribute between 8.5 (HIST) and 10.1 (INC) tokens in the first period and ended between 7.5 (INC) and 11.3 (HIST) (compare also Figure 1.3, Appendix).

1.6 Conclusion

In standard public-good experiments the subjects receive a fresh endowment every period. But this is like fresh restart – sometimes we have to live with the consequences of what we do. If the subjects have to budget their resources they might act very differently because a wrong decision could mean extinction or total failure.

We were interested what might happen if subjects do not receive a fresh endowment in the repetition of the experiment. Thus we created a design where in a repeated public-good experiment the subjects' endowment consisted of the profit in the previous period. This resembles a dynamic public-good which was only introduced quite recently (Gürerk et al., 2010b; Battaglini et al., 2010). Theoretically one might observe an equilibrium where all, even the money-maximizers that will free-ride in a standard public-good, contribute fully until a certain period. In such a case there is the possibility of a huge gain if the group is able to cooperate but there is the danger that a subject loses nearly all resources if she is exploited by others.

We compare this dynamic (DYN) with several other treatments to get to the bottom of the dynamics. In one treatment (HIST) the subjects were randomly assigned to a specific endowment path of DYN to compare exogenous versus endogenous heterogeneity. To control for the heterogeneity itself we had a standard public-good with symmetric and constant endowments (CONST) and a treatment (INC) in which the endowments were symmetric but increasing over time to check for the effect of increasing endowments since we observed increasing average endowments in DYN.

We find that there is no significant difference in the contribution level in the first period though there is a slight anomaly in the distribution of the contributions in DYN. Though not significant it is probably a pointer that some subjects fear to be exploited while others see the enormous income potential by the dynamics.

A difference between the treatments emerges only over the course of the experiment where the contributions deteriorate more quickly in the

DYN treatment compared with all other treatments. The heterogeneity in itself does not have enough explanatory power in the regressions we conducted. However, it is evident that the cooperators that stick with above average contributions quickly lose their resources.

We add another aspect to the literature on heterogeneity. In a certain way our results are representative for the state of the literature: there is no clear evidence that heterogeneity is better or worse for the cooperation. Our dynamic treatment does worse than the others but is that really due to endogenous heterogeneity or just due to the fact that cooperative subjects simply have lost their endowment? Comparing the exogenous heterogeneity to the other treatments reveals no difference. So at the end we would conclude that the heterogeneity does not really make a difference.

It would be interesting to see how other experiments perform without "fresh resource injections" every period. Probably there will be difference as the subjects have to think more about the consequences of their doing, although our results point in a different direction.

1.7 Appendix

1.7.1 Statistics and Graphics

Table 1.6: Per period exact Mann-Whitney-U-test two-sided, high contributors group *included*. p-values.

period	DYN vs. CONST	DYN vs. HIST	DYN vs. INC	CONST vs. HIST	CONST vs. INC	HIST vs. INC
1	0.3431	0.4173	0.4379	0.6924	0.0985*	0.1196
2	0.6640	0.4428	0.4600	0.8355	0.3667	0.2145
3	0.5714	0.2913	0.1153	0.4727	0.1689	0.5714
4	0.5714	0.7987	0.4608	0.5714	0.3141	0.7200
5	0.4727	0.0597*	0.0661*	0.4269	0.2664	0.8949
6	0.3054	0.0887*	0.0938*	0.3431	0.1848	0.3732
7	0.0938*	0.2189	0.1753	0.5082	0.9400	0.5576
8	0.2014	0.2913	0.2703	0.8345	0.8542	0.6784
9	0.1153	0.1432	0.1153	0.9101	0.9591	0.4269
10	0.0073***	0.0684*	0.0691*	0.4269	0.2455	0.9101
11	0.0314**	0.0284**	0.2083	1.0000	0.9591	0.7921
12	0.0252**	0.0519*	0.1153	0.3054	0.3823	0.9101
13	0.0089***	0.0205**	0.1569	0.4603	0.3667	0.7345
14	0.0979*	0.0519*	0.5208	0.7345	0.3823	0.2380
15	0.2618	0.0145**	0.1349	0.2012	0.4875	0.5208
16	0.0201**	0.0145**	0.0552*	0.2703	0.1520	0.7770
17	0.0190**	0.0068***	0.0667*	0.8928	0.2665	0.7769
18	0.0691*	0.0387**	0.2784	0.9699	0.6454	0.8506
19	0.9547	0.0278**	0.2695	0.0551*	0.2908	0.3837
20	0.7772	0.0447**	0.3828	0.1805	0.6444	0.4386

Significances: *** 1%, ** 5%, * 10%

Table 1.7: Per period exact Mann-Whitney-U-test two-sided, high contributors group *excluded*. *p*-values.

period	DYN vs. CONST	DYN vs. HIST	DYN vs. INC	CONST vs. HIST	CONST vs. INC	HIST vs. INC
1	0.3845	0.4173	0.4379	0.8854	0.1786	0.1196
2	0.4197	0.4428	0.4600	0.9507	0.1778	0.2145
3	0.7732	0.2913	0.1153	0.2268	0.0771*	0.5714
4	0.2614	0.7987	0.4608	0.2614	0.1128	0.7200
5	0.7732	0.0597*	0.0661*	0.1673	0.0876*	0.8949
6	0.5358	0.0887*	0.0938*	0.1198	0.0497**	0.3732
7	0.1883	0.2189	0.1753	0.8209	0.7579	0.5576
8	0.3730	0.2913	0.2703	0.8199	0.8401	0.6784
9	0.2268	0.1432	0.1153	0.7732	0.7789	0.4269
10	0.0171**	0.0684*	0.0691*	0.7108	0.4140	0.9101
11	0.0683*	0.0284**	0.2083	0.6504	0.6126	0.7921
12	0.0556*	0.0519*	0.1153	0.5358	0.6126	0.9101
13	0.0207**	0.0205**	0.1569	0.7566	0.5901	0.7345
14	0.1956	0.0519*	0.5208	0.3845	0.6126	0.2380
15	0.4696	0.0145**	0.1349	0.0529*	0.2197	0.5208
16	0.0449**	0.0145**	0.0552*	0.4824	0.2673	0.7770
17	0.0426**	0.0068***	0.0667*	0.7556	0.4460	0.7769
18	0.1422	0.0387**	0.2784	0.7732	0.9551	0.8506
19	0.9014	0.0278**	0.2695	0.1003	0.4145	0.3837
20	0.8844	0.0447**	0.3828	0.0446**	0.3340	0.4386

Significances: *** 1%, ** 5%, * 10%

Table 1.8: Statistics of the contribution in percent of the endowment in the first period.

Treatment	mean in%	st.dev. in %	median %	n
DYN	44.90	21.50	50.00	48
HIST	42.29	22.62	50.00	48
INC	50.63	19.04	50.00	32
CONST	40.45	20.73	40.74	32
CONST ex HC	40.87	20.86	37.96	28

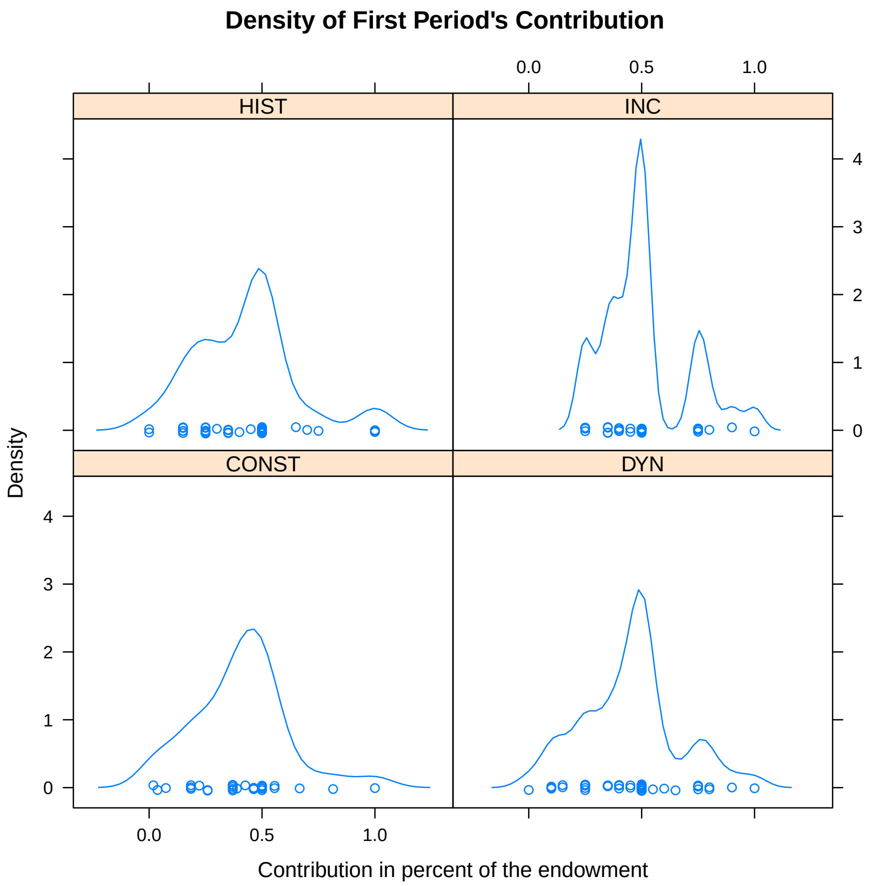

Figure 1.2: Distribution of the first period contribution in percent of the endowment. The circles represent observations, they are slightly jittered to make the distribution of the observations better viewable (reducing the overlaps).

Table 1.9: Random Effects GLS for the change in the contribution in percent of the endowment. L.cp is the contribution in percent of the endowment in t-1, L.avcp the average group contribution in percent of the endowment. L.E and L.Eav are the individual and the group average endowment. Frlag the number of the freeriders. Rich is the dummy for the subject with the highest endowment and theil stands for the Theil-Index of inequality.

	DYN		HIST		CONST		CONST exHC		INC	
Nr.	1	2	3	4	5	6	7	8	9	10
period	0.0033	0.0033	0.0027	0.0027	0.0083	0.0082	0.0175***	0.0171***	0.0059	0.0054
	0.0026	0.0026	0.0036	0.0037	0.0097	0.0092	0.0047	0.0044	0.0049	0.0049
$period^2$	−0.0001	−0.0001	−0.0001	−0.0001	−0.0004	−0.0005	−0.0008***	−0.0008***	−0.0002	−0.0002
	0.0001	0.0001	0.0002	0.0002	0.0004	0.0004	0.0002	0.0002	0.0002	0.0002
L.cp-L.cpav | pos	−1.0603***	−1.0596***	−1.1365***	−1.1454***	−1.1810***	−1.1878***	−1.1955***	−1.2026***	−1.0093***	−1.0014***
	0.1047	0.1054	0.1607	0.1569	0.1611	0.1590	0.1885	0.1854	0.0726	0.0727
L.cpav-L.cp | pos	0.3273***	0.3236***	0.7171***	0.7231***	0.6319***	0.6647***	0.4821***	0.5099***	0.7704***	0.7979***
	0.0980	0.0958	0.1076	0.1075	0.1537	0.1567	0.1649	0.1678	0.0938	0.0951
endpos	0.0015	−0.0094	0.0490	0.1918						
	0.0264	0.0304	0.1484	0.2031						
endneg	−0.1683***	−0.1668***	−0.2033**	−0.2167***						
	0.0423	0.0425	0.0792	0.0807						
Frlag	0.0054	0.0054	0.0209*	0.0217*	0.0299**		0.0217**		0.0142*	
	0.0043	0.0042	0.0117	0.0119	0.0129		0.0108		0.0086	
rich		0.0046		−0.0586						
		0.0085		0.0370						
theil	−0.2124	−0.1940	−0.3134	−0.5313*						
	0.1445	0.1436	0.2214	0.2867						
const.	−0.0258	−0.0261	−0.0280	−0.0206	−0.0080	−0.0130	−0.0447**	−0.0478**	−0.0502**	−0.0513**
	0.0161	0.0164	0.0234	0.0227	0.0426	0.0408	0.0198	0.0206	0.0256	0.0255
R^2 overall	0.3334	0.3335	0.3947	0.3987	0.4010	0.4057	0.4291	0.4323	0.3949	0.3976

Significances: *** 1%, ** 5%, * 10%

Table 1.10: Random effects logistic regression with bootstrapped standard errors, explaining the subjects' decision to free-ride.

	DYN			HIST		CONST		CONST ex HC		INC	
Nr.	1	2	3	4	5	6	7	8	9	10	11
period	0.3106***	0.2464**	0.2478	−0.0249	−0.0252	0.0803	0.0956	0.0839	0.0871	0.2846*	0.1989
	0.1071	0.1026	0.1117	0.1096	0.1128	0.0991	0.1158	0.1147	0.1196	0.1650	0.1855
period^2	−0.0081*	−0.0060	−0.0060	0.0062	0.0062	0.0042	0.0015	0.0039	0.0018	−0.0042	−0.0022
	0.0044	0.0040	0.0043	0.0047	0.0051	0.0042	0.0051	0.0049	0.0053	0.0065	0.0078
L.cp-L.cpav | pos	−0.6802	−3.9234*	−3.9354	1.2667	1.2995	0.8917	0.4237	0.8905	0.4528	0.8951	1.1188
	1.8558	2.2292	2.3684	1.3897	1.3888	1.3599	1.4266	1.2625	1.2352	2.5754	2.2993
L.cpav-L.cp | pos	−2.1271	−1.8082	−1.3389	−1.5412	−1.5491	−3.0717	−2.2334	−4.4211	−3.5268	−21.108***	−17.230**
	2.9780	3.0187	2.9019	2.1689	1.9836	2.2740	1.8759	1.9521	2.3056	5.9336	7.3471
endpos		3.5282***	5.4386	−2.2631	−3.3507						
		1.0923	1.6729	1.4644	2.1424						
endneg		−5.8448***	−6.0723	−0.2560	−0.1172						
		1.2796	1.4168	1.8920	1.6846						
Frlag	0.7543***	0.8416***	0.8538	0.9501***	0.9496***	0.7277***		0.7182***		0.5162*	
	0.1481	0.1496	0.1438	0.2303	0.2450	0.2120		0.2057		0.2732	
rich			−0.7969		0.4531						
			0.5263		0.6128						
theil	−0.5627	−11.5127**	−14.2604***	6.9186	8.3871						
	4.1805	4.6736	5.2185	6.6620	6.4443						
const.	−4.3128***	−4.3974***	−4.3452	−4.0929***	−4.1412***	−3.2973	−3.4836***	−3.0110***	−3.1971***	−5.2349***	−4.7405***
	0.6585	0.6755	0.7374	0.6149	0.6959	0.6368	0.6603	0.6501	0.6737	1.0297	0.9888
Wald χ^2	102.30	94.98	108.00	68.62	69.30	54.46	53.58	48.13	40.05	35.12	37.02

Significances: *** 1%, ** 5%, * 10%

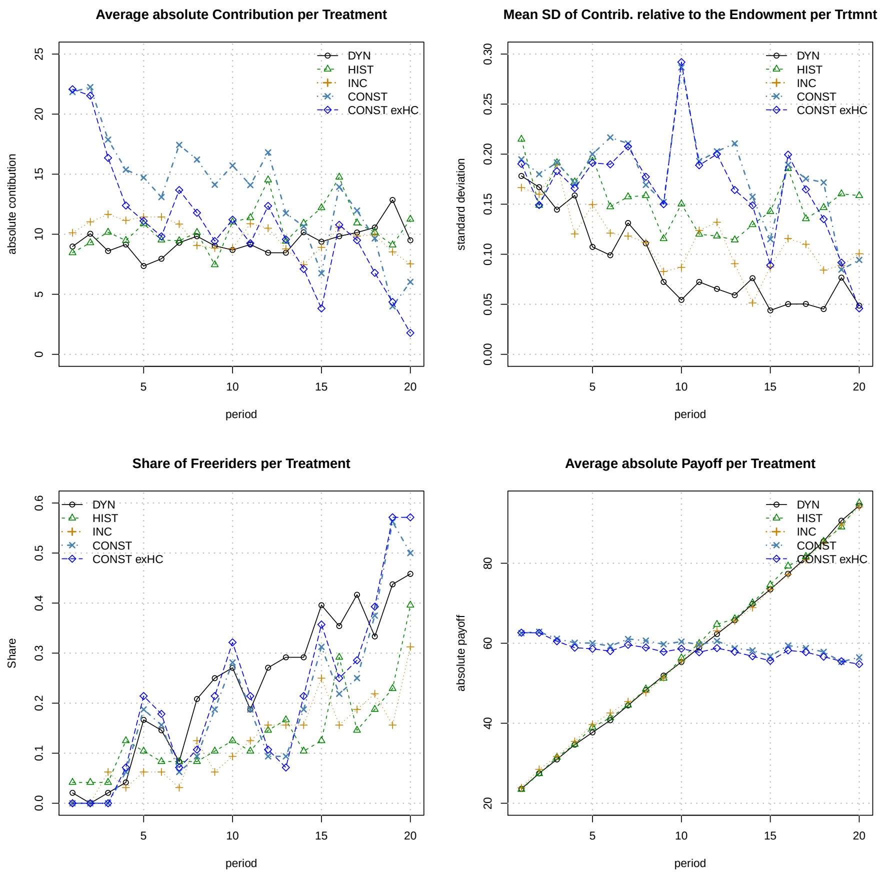

Figure 1.3: Average absolute contribution per treatment. Average standard deviation of the contribution relative to the endowment. Share of the free-riders per treatment. Average absolute per-period payoff per Treatment.

1.7.2 Translation of the Instructions for the INC Treatment

Please read these instructions carefully. If you have any questions please ask the instructor.

From now on you are not allowed to communicate with the other participants!

The course of the experiment

- At the beginning of the experiment you will be randomly assigned to a group of four members.
- You will remain in the same group for the whole experiment.
- Every group is independent of the actions of the other groups.
- The experiment consists of 20 rounds.

Common and private projects

- There exist two projects: a private and a common project fo all four group members.
- In every round you have the opportunity to divide the endowment between the common project and the private project:
    - All four group members may invest in the group project. The income from the group project will be equally shared by all members of the same group.
    - In the private group only you are allowed to invest. The income of this project remains your own.

  You may divide your endowment between both projects or everything in one of the projects.
- The amount you will assign to the common project will be specified in the field *"Your contribution to the group project"*. The difference of your endowment and this contribution will by the investment in the private project.

Round endowment

The endowment is different in every period:

Round	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Endowment	20	24	27	31	35	38	41	44	48	52	55	59	62	66	70	73	77	81	86	91

Calculation of the round profit

- Your round profit consists of two parts:
  - Your contribution to the private project (= Endowment minus contribution to the common project) and
  - the with 0.35 multiplied sum of *all* contributions in your group.

  So all group members receive payout from the common project independent on the size of their individual investment.

Calculation of the round profit:		
Endowment	−	Your Contribution to the common project
	+	0.35 · Sum of contribution of <u>all</u> group members to the common project
	=	Your round profit

Information at the end of each round

After every round each participant will receive the following information:

- Your own contribution to the common project (absolute and in percent of the endowment)
- Sum of all contributions to the common project
- Your own profit of the current round
- The endowment, the contribution (absolute and in percent of the endowment) as well as the the period profit of the other three members of the group.

Calculation of the overall profit

The overall profit is the sum of all 20 period profits. At the end of the experiment the overall profit will be converted into real money with an exchange rate of € 0.25 per experimental currency unit.

Please note that communication between the participants is not allowed during the course of the experiment. If you have a question for the instructor raise your hand out of the cabin.

All decisions made during the experiment as well as the size of the cash you will receive at the end of the experiment are anonymous. No other participant of the experiment will know the identity of her group members. If you have any questions you might ask them now.

Good luck.

2

Monitoring in Teams: Using Laboratory Experiments to Study a Theory of the Firm[1]

Abstract

Alchian and Demsetz's (1972) influential explanation of the classical business firm argues that there is need for a concentrated residual claim in the hands of a central agent, to motivate the monitoring of workers. We model monitoring as a way to transform team production from a collective action dilemma with strong free-riding incentives to a productivity-enhancing opportunity with strong private marginal incentives to contribute effort. In an experiment, we let subjects experience team production without monitoring, team production with a central monitor, and team production with peer monitoring. Then subjects vote on whether to employ the central monitor, who gets to keep a fixed share of the team output, or to rely on peer monitoring, which entails a coordination or free-riding problem. Our subjects usually prefer peer monitoring but they switch to the specialist when unable to successfully self-monitor. We provide evidence for situations in which team members resist the appointing of a central monitor and succeed in overcoming coordination and free-riding problems as well as for a situation in which an Alchian-Demsetz-like firm "grows" in the laboratory.

2.1 Introduction

What accounts for the structure of the capitalist firm, in which equity suppliers or their agents hire and supervise workers given few or no residual claims? In an influential paper about the theory of the firm, Alchian and Demsetz (1972) characterized team production by the following four properties: (*i*) there exist several input providers, (*ii*) the combined output

[1]Part of "Monitoring in Teams: Using Laboratory Experiments to Study a Theory of the Firm" by Stefan Grosse, Louis Putterman and Bettina Rockenbach (2011) and published in the *Journal of the European Association*, Vol.9(4), (Grosse et al., 2011). All authors contributed equally.

is larger than the sum of the outputs that the individual input providers can achieve by working alone, (*iii*) there is an observable team output but no observable output of the individual input provider, and (*iv*) it is possible but costly to measure the amount of input contributed by each individual provider. The central dilemma of team production, they argued, is that the benefits of working as a team (e.g., benefits from economies of scale or of specialization) may be undercut by the incentive to free-ride that each team member has if compensated according to team output rather than personal input. To mitigate this problem, team members' rewards must be tied to their contributions, but that requires another costly input – monitoring – and this in turn gives rise to another collective action problem if monitoring is to be supplied by the team members themselves. The classical capitalist firm solves this problem, they argue, by making one specialized agent the monitor of the other team members who pays them according to their observed inputs. The central agent is motivated to monitor by the fact that he keeps all team revenue above his contractual obligations to the input providers.

We understand Alchian and Demsetz's depiction of team production in the absence of monitoring to be an example of the familiar problem of collective action or incentives in teams that has been studied by experimental economists in recent decades under the heading Voluntary Contribution Mechanism (VCM) or Public Goods Game (PGG). In a VCM or PGG, subjects are grouped with others and each decides how much of a certain endowment to contribute to a group project and how much to hold for herself. Contributions to the project are scaled up by the experimenter, such that there is a social optimum of contributing. However, since the resulting revenues are divided equally among team members, the individual optimum is to contribute nothing. We interpret Alchian and Demsetz as saying that if a sufficient investment is made in monitoring individuals' contributions, then they can be paid according to their contributions, rather than an equal per capita share, as a result of which there will be an incentive to contribute and not to free-ride.

We present a simple theoretical model corresponding to this structure, and we investigate how real decision-makers respond to the structure by having subjects make potentially rewarding / costly decisions under it in a laboratory experiment. In the model and experiment, monitoring can either be done by a specialized agent, who is assigned a fraction of the team's joint output, or by the team members themselves, who are then compensated for their contributions to production but not for their monitoring itself. Suppose that agents care only about increasing their own earnings, know one another to be of the same type, and are rational. Then, if the only monitoring were to be that done by the team members themselves, there would be a considerable possibility that monitoring would not suffice and hence that the production stage of the model would be a simple VCM, for which there is a straightforward prediction of zero contributions. If, instead, a specialist were offered a sufficient fraction of team output and permitted to monitor, it would be in the specialist's interest to monitor enough to make contributing to team production rational for each team member. With appropriate specifications of returns to team production and of the share claimed by the specialist, team members earn more producing together with a specialist monitor than having no monitor and producing individually. If allowed to vote at no cost – a proxy for workers' choice among organizational forms in a market economy – the model predicts that team members will vote to hire the specialist unless they manage to successfully monitor themselves.

We carry out experimental play of such a model. We vary the conditions under which team members and specialists can learn about their tasks by varying the order in which play occurs (*i*) with no monitoring, (*ii*) with monitoring (if any) by team members, and (*iii*) with monitoring (if any) by a specialist, before having several opportunities to vote on which kind of monitoring to use, more periods of play, and opportunities to vote again. We also vary the costliness of monitoring for team members versus specialists.

Ours is the first experiment we are aware of in which a public goods game with its well-known free-rider problem can be converted into a payment for effort environment without free-rider problem by the free choices of subjects. It extends the recent innovation of studying institutional evolution in the laboratory, applying it to a key issue in the theory of economic organization that has not previously been addressed by such methods.

Our results are striking. In four of the six treatments with which we experiment, almost all teams are successful at self-monitoring and thus choose not to hire a specialist. But when we make monitoring by team members more costly than that by the specialist peer monitoring fails in many groups and a trend towards specialist monitoring emerges. Our results thus accord with experimental findings that a large number of subjects attempt cooperation in the lab, but also with the standard experimental finding that in repeated dilemma games without devices such as punishment opportunities or pre-play communication, cooperation tends to flag over time.[2] For this reason, the logic of Alchian and Demsetz's argument is supported in the lab in a particularly clear fashion.

[2]See for example Ledyard (1995) as well as Davis and Holt (1993) for an overview of older public goods experiments, Fehr and Gächter (2000a) for the effect of punishment and Brosig et al. (2003) for the effect of communication.

The structure of the Chapter is as follows: Section 2.2 briefly discusses the theory and literature on the organizational form of production in a market economy. Section 2.3 presents our theoretical model, and Section 2.4 lays out its implementation in our experimental design. Section 2.5 presents the experiment's results. Section 2.6 summarizes and provides additional discussion.

2.2 Literature

Why most firms in market economies exhibit certain common features, and in particular why control rights usually reside in a group of investor / residual claimants, with employees working under the supervision of their employers, has long been a central question of the economics of organization and comparative institutional analysis. Knight (1921) argued that the more confident and less risk-averse individuals become entrepreneurs while others become workers who demand insurance against risk and who accordingly must be supervised, since their fixed wages give rise to moral hazard (see also Kihlstrom and Laffont, 1979). Alchian's and Demsetz's explanation of why workers are supervised by a residual-claiming central monitor was summarized in the introduction. Marglin (1974) argued that capitalists carved out the role of imposing discipline on workers at the expense of workers' welfare, by developing technologies that undercut the positions of independent workers. Holmström (1982) suggested that the monitoring of inputs could be rendered unnecessary by a forcing contract, but the contract envisioned is largely hypothetical and has been argued to suffer from serious moral hazard problems (Eswaran and Kotwal, 1984; MacLeod, 1988). Eswaran and Kotwal (1989) and Banerjee and Newman (1993) explain the assignment of control rights to financiers by reference to unequal wealth and imperfections in credit markets associated with the limited liability of borrowers. Kremer (1997) argued that workers usually do not run firms because control by workers leads to a tendency to redistribute earnings among members, which distorts incentives.

Dow and Putterman (2000) as well as Dow (2003) view Alchian–Demsetz's monitoring hypothesis as one of the leading candidates to explain the conventional employment relationship,[3] alongside theories

[3]See also the references to Alchian's and Demsetz's hypothesis in many of the papers cited in the previous paragraph. One of our referees argued that Alchian and Demsetz's model has no applicability to modern corporations, since in such firms supervision is done by hired personnel and since such personnel could as easily be hired by workers. In fact, Alchian and Demsetz took care to state that their approach is most directly applicable to the "classical capitalist firm" with an owner-manager engaging directly in supervision. They argue, nonetheless, that the corporation system works well partly due to the existence of a market for corporate control in which "control is facilitated by the temporary congealing of share votes into voting blocs owned by one or a few contenders

of worker liquidity constraints and risk aversion, additional financing problems associated with missing membership markets, and potential decision-making problems due to heterogeneity of worker preferences. However, they point out that contrary to the theory's implication that work incentives would be weak without a residual-claiming central monitor, most evidence on worker-owned and profit-sharing firms, as well as that on self-managing teams, suggests that they achieve higher-than-average effort levels with less-than-average numbers of supervisors (Estrin et al., 1987; Weitzman and Kruse, 1990; Craig and Pencavel, 1995). Incentives appear to be a strength rather than a weakness of profit-sharing, with a frequently mentioned theme being its encouragement of mutual monitoring.

In a recent experimental study of work organization and incentives Potters et al. (2009) compare laboratory manager-less teams that play a standard public goods game with teams having managers who can decide how much to pay the other members. They find that managers are able to elicit higher effort from team members than is forthcoming in the PGG, by linking pay to effort somewhat in the manner suggested by Alchian and Demsetz. While the performance of their "managerial" firms is remarkable, their manager-less firms may be a poor representation of self-managing teams, since linkage of pay to effort is ruled out in such teams under their experimental design.

Another attempt to experimentally compare self-managed teams and centrally managed teams has been undertaken by Frohlich et al. (1998). They designed a real-effort experiment wherein they observed higher productivity, greater perceived fairness in pay and lower need of supervisory efforts for employee owned firms compared to the "conventionally owned" firms. Another experimental study incorporating different group incentive mechanisms is Nalbantian and Schotter (1997). They compared revenue sharing, forcing contracts, competition between teams, profit sharing and monitoring. Monitoring in their context was a probability of being observed and getting fired when one's effort is too low. This kind of monitoring was successful but only if the probability is high enough; thus, successful monitoring is expensive.[4]

. . . a transient resurgence of the classical firm" (p. 788). A full discussion of the applicability of Alchian and Demsetz's theory to real world firm organization would take us beyond the scope of this Chapter.

[4]The numerous social dilemma experiments beginning with Fehr and Gächter (2000a) or Carpenter et al. (2009), in which subjects can punish those who contribute too little to a public good, can also be viewed as studying alternative incentive mechanisms for group production. In these experiments, the public good always remains public, whereas we allow its public character to be eliminated by monitoring.

2.3 A Model of Team Production with Monitoring

We model a team consisting of N members who play a finitely repeated game for T periods. In each period, a team member receives an endowment e, which we assume to be identical for all members. Team member i chooses an amount c_i with $0 \leq c_i \leq e$ to contribute to a team production process, leaving $e - c_i$ for private production. The sum of the team members' contributions (denoted by $C = \sum_{i=1}^{N} c_i$) generates a team profit of $R \cdot C$ with $1 < R < N$. The division of the team profit among the team members depends upon the monitoring technology applied to identify the individual team contributions, which is a result of a simultaneous investment process prior to the contribution decision. Each team member invests $m_i \in [0, \ldots, 1]$ into the monitoring technology at a linear cost $\kappa \cdot m_i$ (with the marginal monitoring cost $\kappa \geq 1$). The total investment in monitoring $M = \sum_{i=1}^{N} m_i$ determines the "accuracy" of the monitoring technology and thus the proportion of the team profit which is divided according to the individual contribution. $M = 0$ allows no identification of the individual contributions and hence the team profit is divided equally among the team members. The higher M, the higher is the proportion of the team profit which is allocated according to the individual contributions. $M = N$ allows a perfect identification of the team members' contributions and hence the team profit is allocated according to the individual contributions. The general rule for team member i's profit is:

$$\pi_i = e - \kappa \cdot m_i - c_i + \frac{N - M}{N} \cdot \frac{R}{N} \cdot C + \frac{M}{N} \cdot R \cdot c_i \,. \tag{2.1}$$

The monitoring technology changes the nature of the team problem. Without any monitoring ($M = 0$) team production is a classical linear public good provision problem with free-rider incentives due to $\pi_i = e - c_i + R/N \cdot C$. However, if each team member fully invests in the monitoring technology ($M = N$), team production is a private investment task with $\pi_i = e - \kappa - c_i + R \cdot c_i$. The positive interest rate $R - 1$ provides incentives for full contributions. Intermediate values of M lead to linear combinations of the public and the private good provision. If, for example, half of all team members fully invest in monitoring, i.e., $M = N/2$, then half of the team output is allocated according to the private contribution and the other half is distributed equally among the team members, i.e.,

$$\pi_i = e - \kappa \cdot m_i - c_i + \frac{1}{2} \cdot \frac{R}{N} \cdot C + \frac{1}{2} \cdot R \cdot c_i \,.$$

Thus, the model reflects Alchian and Demsetz's idea that without monitoring team members have incentives to free-ride on others' effort provision, however with a sufficient investment in monitoring individuals' contributions, team members can be paid according to their contributions as a result of which there will be an incentive to contribute and not to free-ride.[5] In the model, the level of monitoring and hence the quality of the pay-effort link is determined and made known before the team members select their production effort, which is necessary if monitoring is to influence effort as Alchian and Demsetz assume it to do.[6] For simplicity, our model compresses the choice of monitoring investment, the observation of effort, and the translation of observations into payment shares into a single step, much as Alchian and Demsetz (p. 778) cover both measurement and apportionment by the term "metering".[7] Hence, team members may choose their effort in response to the remuneration scheme in place, including whether their pay will be linked to their effort, something that might be possible only with monitoring.

For the analysis of the subgame perfect equilibria of the game it is convenient to restructure (2.1) as:

$$\pi_i = e - \kappa \cdot m_i - c_i + \beta \cdot c_i + \gamma \cdot C_{-i} \tag{2.2}$$

where $C_{-i} = \sum_{j=1;j\neq i}^{N} c_j$ denotes the sum of the others' contributions, the weight $\beta = R/N^2 \cdot (N - M + N \cdot M)$ denotes the team member's individual return from her own investment and the weight $\gamma = R/N^2 \cdot (N - M)$ denotes the team member's return from the investment of the others.

With no monitoring $\beta = \gamma = R/N$, meaning that all team members profit equally from each unit of contribution, while with perfect monitoring $\beta = R$ and $\gamma = 0$, meaning that only the contributor profits from the own contribution. Obviously, it is individually rational to contribute the entire

[5]Notice that the production function in our model is – in contrast to Alchian and Demsetz – additively separable in individual inputs. We decided on this for two reasons: firstly, to keep the model simple and understandable to the subjects and secondly that with this kind of production function we can link up to the literature on public good provision, which uses the same kind of modeling. Our set-up nevertheless captures the key Alchian and Demsetz idea that it is socially efficient to contribute to team production, but individual input is not costlessly distinguishable from the inputs of others, unless costly monitoring is provided. Even with a separable production function like the one that we use, individual reward may not be well linked to individual effort without costly monitoring because the inputs of the individual team members might be costly to discern.

[6]The level of supervision or probability of detection of shirking is also assumed known when effort choices are made in efficiency wage models like Shapiro and Stiglitz (1984) and Bowles and Gintis (1990).

[7]In real organizations, there is first a commitment of resources to the monitoring process, then the process itself is carried out as effort is being exerted, and finally the information obtained is used to adjust rewards, including by making promotion and firing decisions. We capture all of this in the simplest possible manner, as a decision on monitoring resources that automatically determines observational accuracy and thereby the degree to which payment follows an equal-sharing versus a proportionate-to-effort formula. Whereas more complex models in which monitoring's accuracy or its translation into payments are stochastic or even subject to intra-organizational conflict (see, e.g. Kremer, 1997) are possible, our model is kept simple so as to focus on the central idea that the members of the team must have a sense of how much monitoring will be in place and how the information obtained is to be used if it is to affect their effort choices.

endowment when $\beta \geq 1$, because each token invested has an individual return of at least 1. $\beta \geq 1$ is satisfied if and only if

$$M \geq \frac{N}{N-1} \cdot \left(\frac{N}{R} - 1\right) =: \widetilde{M}\,.$$

Equilibrium investment in monitoring and contributions to the team project

The game consists of two stages. In the first stage players simultaneously invest in monitoring. After having learned the total investment M players simultaneously decide on their contribution to the team project. We analyze the game by backward induction identifying the subgame perfect equilibria under the assumption that each team member is solely motivated by the maximization of her monetary payoff. Consider the subgames of the contribution to the team project (after the amount M was made public). It suffices to distinguish three classes of subgames: those with $\beta < 1$, those with $\beta > 1$, and those with $\beta = 1$. For $\beta < 1$ the individual return from the individual contribution is lower than the cost of contributing and hence in the equilibria of these subgames all team members choose $c_i = 0$. If, however, $\beta > 1$ each team member individually gains from contributing and hence will choose $c_i = e$ in equilibrium. For $\beta = 1$ players are indifferent between contributing and keeping the entire endowment or parts of it and hence each contribution $0 \leq c_i \leq e$ may be part of a subgame perfect equilibrium. Now turn to the investment in monitoring. The subgame has multiple equilibria. There is a symmetric Nash equilibrium in pure strategies in which no player invests in monitoring ($m_i = 0$) and there is a multiplicity of equilibria in which the critical investment level $\widetilde{M}$ (which is necessary to make full contribution to the public good individually rational) is exactly met. One of these is a symmetric equilibrium in pure strategies in which each player invests the N-th part of $\widetilde{M}$ (i.e., $m_i = \widetilde{M}/N$). In addition, there is an infinite number of asymmetric pure strategy equilibria of the subgame which are all characterized by investments m_i satisfying $M = \widetilde{M}$ and additionally there are symmetric and asymmetric mixed strategy equilibria.

Hence the "good news" that the public good dilemma of team production may be "resolved" in the monitoring phase prior to it comes along with the "bad news" that the investment in monitoring is vulnerable to severe coordination failures due to a multiplicity of equilibria.[8]

[8] See also Marx and Matthews (2000).

Specialists monitoring

To overcome the problem of multiple equilibria in the monitoring phase, team members may hire a specialist to take the monitoring decision. The substitution of peer monitoring by specialist monitoring has the advantage that the specialist is a single decision maker who (in equilibrium) chooses an incentive compatible level of monitoring without any coordination problems. The drawback is that she has to be paid a share of the team output in order to have the proper incentives.[9]

Let the specialist be entitled to a share $S \leq 1$ of the team profit $R \cdot C$. Suppose that the specialist has an endowment e_S which enables her to invest at least $\widetilde{M}$ units of monitoring. Thus, the payoff functions under specialist monitoring are as follows:

$$\pi_S = \quad e_S - \kappa_S \cdot m_S + S \cdot C \cdot R \qquad \text{for the specialist} \tag{2.3}$$
$$\pi_i^S = \quad e - c_i + \beta^S \cdot c_i + \gamma^S \cdot C_{-i} \quad \text{for team member } i \tag{2.4}$$

with the adjusted weight $\beta^S = (1 - S) \cdot \beta$ denoting the team member's individual return from the own investment after deduction of the specialist's share and the adjusted weight $\gamma^S = (1 - S) \cdot \gamma$ denoting the team member's return from the investment of the others after deduction of the specialist's share.

Full contribution of the team members is individually rational if and only if

$$\beta^S \geq 1 \Leftrightarrow M \geq \frac{N}{N-1} \cdot \left(\frac{N}{(1-S) \cdot R} - 1 \right) =: \widetilde{M}^S. \tag{2.5}$$

If the specialist invests less than $\widetilde{M}^S$, team members in equilibrium contribute a total of zero units of effort to team production, so the specialist's earnings from team production will be $S \cdot 0 = 0$. If the specialist invests at least $\widetilde{M}^S$ in monitoring, each team member in equilibrium contributes her full endowment of e to team production, so the specialist's earnings from team production will be $S \cdot N \cdot e \cdot R$. Hence, for reasonable costs κ_S the specialist will in equilibrium choose the lowest monitoring level for which it is individually rational for the team members to fully contribute their endowment – that is $\widetilde{M}^S$ – and gain a total profit of $\pi_S = e_S - \kappa_S \cdot \widetilde{M}^S + S \cdot N \cdot e \cdot R > e_S$.

To recap, we presented a formal model of team production in the spirit of Alchian and Demsetz. The elegance of the model is that it allows a

[9] Alchian and Demsetz never spell out where the residual earnings of the central monitor come from, simply asserting that the monitor pays team members the estimated value of their marginal products and keeps the residual. Our model assigns to the monitor a fraction of the output because with average and marginal product equal, there is no residual above the sum of marginal products. We implement the model with sufficiently large R so that both monitor and team members can profit from centrally monitored team production.

continuous transformation of the team problem with free-riding incentives into a profitable private investment problem through the actions of the team members and/or the decision of the specialist monitor. Due to the multiplicity of equilibria, it is difficult – though not impossible – that team members manage the transformation on their own. In contrast, the specialist is unambiguously predicted to carry out the transformation if parameters are consistent with $\pi_S > e_S$ when $M = \widetilde{M}^S$, since she can accomplish this by a single individual decision. The drawback to the team members of hiring is its cost, albeit they are – in equilibrium – more than compensated compared with full free-riding.

A discrete version of the model

For the experimental implementation of the game we choose a discrete version of the payoff function and a binary choice in the investment in peer monitoring $m_i \in \{0,1\}$ to facilitate comprehension by subjects. We exogenously introduce two different thresholds of monitoring T_1 and T_2 with $T_1 < T_2 \leq N$. If the total investment in monitoring $M < T_1$ all team members equally profit from all contributions, for $T_1 \leq M < T_2$ half of the team profit is allocated equally and the other half according to individual contributions, and finally, for $T_2 \leq M \leq N$ each team member solely profits from her own contribution. Notice, that players are informed both on the achieved monitoring threshold (but not on the exact value of M) and on the corresponding payoff function:

$$\pi_i = \begin{cases} e - \kappa \cdot m_i - c_i + \frac{1}{N} \cdot R \cdot C, & 0 \leq M < T_1; \\ e - \kappa \cdot m_i - c_i + \frac{1}{2} \cdot \frac{1}{N} \cdot R \cdot C + \frac{1}{2} \cdot R \cdot c_i, & T_1 \leq M < T_2; \\ e - \kappa \cdot m_i - c_i + R \cdot c_i, & T_2 \leq M \leq N. \end{cases} \quad (2.6)$$

In terms of β and γ this means:

$$\begin{cases} \beta = R/N, & \gamma = R/N \quad 0 \leq M < T_1; \\ \beta = \frac{R(N+1)}{2N}, & \gamma = \frac{R}{2N} \quad T_1 \leq M < T_2; \\ \beta = R, & \gamma = 0 \quad T_2 \leq M < N. \end{cases}$$

Example

The following example illustrates the model and uses functional forms and parameters that will also be used in our experiment. Let $N = 5$ be the number of team members with an endowment $e = 10$, marginal monitoring costs $\kappa = \kappa_S = 1$, a multiplier $R = 3$, the specialist's endowment

$e_S = 5$ and the specialist's share $S = 0.25$. Then

$$\begin{cases} \beta = 0.6, & \gamma = 0.6, & \beta^S = 0.45, & \gamma^S = 0.45 & 0 \leq M < T_1; \\ \beta = 1.8, & \gamma = 1.5, & \beta^S = 1.35, & \gamma^S = 1.125 & T_1 \leq M < T_2; \\ \beta = 3.0, & \gamma = 0, & \beta^S = 2.25, & \gamma^S = 0 & T_2 \leq M < N. \end{cases}$$

Hence for $M \geq T_1$ full contribution to the team project is individually rational, because the individual return from investment β is greater than 1. In the subgame perfect equilibrium without peer monitoring ($m_i^* = 0$), contributions to the team project are 0 ($c_i^* = 0$), leading to team members' payoffs of 10. However, there are also equilibria in which monitoring takes place. The simplification of the model by choosing discrete values of monitoring and thresholds restricts the number of those equilibria. Nevertheless, there are still $\binom{N}{T_1}$ subgame perfect pure strategy equilibria, characterized by exactly T_1 team members investing in monitoring in addition to mixed strategy equilibria (see below).

In the experiment we used two treatments in which $T_1 = 2$ and three in which $T_1 = 4$. Because team members are restricted to integer investments, a symmetric equilibrium with monitoring is not achievable. This means that the only symmetric pure strategy equilibrium prescribes no investment in peer monitoring. All the pure strategy equilibria with monitoring are asymmetric and hence very vulnerable to coordination failure. In case of $N = 5$ and $T_1 = 2$, the game has 10 pure strategy equilibria in which exactly 2 out of the 5 players have to invest in monitoring and in case of $N = 5$ and $T_1 = 4$, the game has 5 pure strategy equilibria in which exactly 4 out of the 5 players have to invest in monitoring.

In a symmetric mixed strategy equilibrium each player has to be indifferent between investing and not investing in monitoring, which means that the expected excess gain from investing in monitoring has to equal the cost of investing in monitoring. To achieve this, each team member assesses the probability that he/she is pivotal, meaning that by her investment in monitoring the threshold to incentive compatible contributions would be reached.[10] Only in the case of being pivotal, a player may have an incentive to invest in monitoring. Appendix 2.7.1 provides the exact calculations for the symmetric mixed strategy equilibrium. In fact, it can be shown that in our parameterization there are two symmetric mixed strategy equilibria in each of the parameterizations, one with a low probability of investment in monitoring and another one with a high probability of investment in monitoring, as listed in Table 2.1. For each of the three parameterizations the table lists the probability of investing in monitoring in each of both equilibria. In the first parameterization, for example, each player independently chooses to invest in monitoring with probability of

[10] See also Gradstein and Nitzan (1990) and Offerman et al. (1996).

Table 2.1: Symmetric mixed strategy equilibria.

Parameterization	Probability of investment in monitoring in: "low probability" equilibrium	"high probability" equilibrium
$T_1 = 2$, $T_2 = 5$, $\kappa = 1$	1.30%	74.39%
$T_1 = 4$, $T_2 = 5$, $\kappa = 1$	25.61%	98.70%
$T_1 = 4$, $T_2 = 5$, $\kappa = 3$	39.60%	95.72%

1.3% in the low probability equilibrium, whereas each player independently chooses to invest in monitoring with roughly 75% probability in the high probability equilibrium. These numbers allow us to calculate the likelihood with which we can expect the different investment levels if subjects are playing the respective equilibrium strategies. In Figure 2.3 we will report these numbers and contrast them to the observed frequencies.

If team members are able to self-organize (i.e., achieve $M \geq T_1$) each team member earns 30 minus the investment in monitoring (if individually applicable).[11] In the equilibrium of specialist monitoring the specialist invests T_1 in monitoring and the team members contribute their entire endowment. Hence, the team members earn $0.75 \cdot 30 = 22.5$ and the specialist earns her endowment (of 5) minus the monitoring investment plus $0.25 \cdot 150 = 37.5$.

Obviously, it would be most profitable for the team members to play one of the equilibria with positive peer-monitoring. Then each member earns 29 or 30, dependent on whether he/she invested in monitoring or not. However, failing to reach the sufficient level of monitoring leads to drastically lower individual payoffs of 9 and 10, dependent on whether the individual invested in monitoring or not.[12] Facing this risk, team members may decide to hire a specialist to make the monitoring decision and achieve a payoff 7.5 lower than the highest equilibrium payoff, but 13.5 higher than the worst payoff in case of coordination failure without sufficient monitoring.

2.4 Experimental Design

We conducted an experiment consisting of five treatments corresponding closely to the model above. In each session of the experiment, subjects were randomly and anonymously assigned to groups of six, with one subject

[11]Note, that the monitoring cost is paid out of end-of-round earnings; thus, contributing to monitoring does not prevent a subject from still contributing a full 10 units to team production.

[12]The other form of coordination failure in the form of over-provision of monitoring is "less disastrous" because it just leads to more players earning 29 instead of 30, than in equilibrium.

Table 2.2: Step structures.

Division rule	**Step structure 2-5**	**Step structure 4-5**
equal division ("EQUAL")	$0 \leq M < 2$	$0 \leq M < 4$
half divided equally, half according to contributions ("HALF/HALF")	$2 \leq M < 5$	$4 \leq M < 5$
Division according to individual contributions ("ATIC")	$M = 5$	$M = 5$

randomly assigned the role dubbed "observer" (corresponding to the theory section's "specialist") and the other five the role "team member". We implemented the discrete version of the game described above with the parameters of the example above. Subjects were told at the outset that they would engage in thirty rounds of decisions in the same roles and with the same anonymous group members. The two *step structures* 2-5 ($T_1 = 2$ and $T_2 = 5$) and 4-5 ($T_1 = 4$ and $T_2 = 5$) specify two sets of parameters for the thresholds T_1 and T_2, which in turn generate three possible incentive regimes for team production henceforth referred to as EQUAL, HALF/HALF, and ATIC ("according to individual contribution") (see Table 2.2).

In step structure 2-5 at least two units have to be invested in monitoring to make contributions to the team project individually rational, while in step structure 4-5 at least 4 units have to be invested. Each group of subjects was assigned to either one structure or the other throughout their session, with no knowledge of the other structure.

The 30 rounds of a session were divided into six phases, with 5 rounds each. In every session, Phase I consisted of 5 rounds with no monitoring – i.e., a standard 5 round VCM condition. Phases II and III consisted of 5 rounds with monitoring (if any) by the observer and 5 rounds of monitoring (if any) by peers, with the order in which observer and peer monitoring occurred varying among sessions (see Table 2.3). In OP sessions, the observer made the monitoring decisions in Phase II and the team members made the monitoring decisions in Phase III; in PO sessions, the order was reversed.

To avoid boredom and unnecessary inequalities and to motivate the observer to learn about incentives in team production, we assigned the observer a task to perform in those periods in which he or she was not permitted to monitor and earn a 25% share of team project revenue. The observer's task was to estimate the period's sum of contributions C in her group. As an incentive for accuracy, the observer earned more the

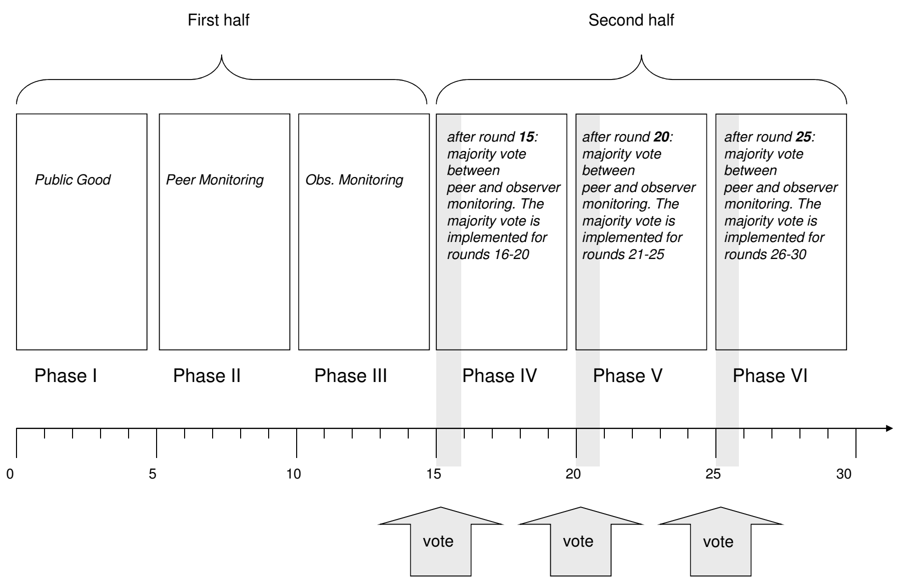

Figure 2.1: Schematic representation of the course of the interaction for PO.

closer was her guess to the actual C, which was revealed to him/her at the end of the period.[13] Note that the observer might learn something about how team members' contributions respond to monitoring by observing peer monitoring phases, and accordingly sessions using the PO ordering might be expected to be more conducive than those with ordering OP to successful decision-making by the observer when in the monitoring role.

In each session, each of the last three phases could have either observer or peer monitoring, depending on how the members of the team in question voted. Before rounds 16, 21, and 26, each team member was asked to vote for either observer or peer monitoring. The group was informed of the majority vote (without a breakdown of the number of votes) and began to play five rounds according to the chosen institution. A schematic representation of the course of the interaction in the PO ordering is given in Figure 2.1. Phases I to III form the first half of the experiment, and phases IV to VI the second half.

The alternatives of the PO or OP ordering and of the 2-5 or 4-5 monitoring structure give rise to a 2×2 design with four treatments: PO25, OP25, PO45, OP45. Due to the unexpected nature of the results of those treatments, which are discussed in the next section, we conducted sessions with an additional treatment that is otherwise like the PO45 treatment but in which the marginal cost of monitoring was made three times higher for a team member than in the other four treatments, while the cost of

[13]The formula for the observer's profit during phases in which he did not play a monitoring role, such as Phase I, was: $\pi = 30/(1 + 0.05|C - \mathit{Guess\ of\ C}|)$

Table 2.3: Treatment description.

Treatment	Phase Sequence		Step	Monitoring unit cost	
	Phase II	Phase III	strc.	Peer κ	Observer κ_S
PO25MC1	Peer	Observer	2-5	1	1
OP25MC1	Observer	Peer	2-5	1	1
PO45MC1	Peer	Observer	4-5	1	1
OP45MC1	Observer	Peer	4-5	1	1
PO45MC3	Peer	Observer	4-5	3	1

monitoring for the observer was left unchanged (i.e., $\kappa = 3$ and $\kappa_S = 1$). We distinguish the two treatments by referring to them as OP45MC1 and OP45MC3, with the other three treatments also sharing the MC1 designation. Table 2.3 provides an overview of the five treatments.

In each treatment we have 6 groups (from two sessions of three groups each) each containing 6 subjects (5 team members and 1 observer). Hence we had 180 subjects in the experiment who were randomly allocated to the different treatments. Each subject sat in a separate compartment in the experimental lab at the University of Erfurt and interacted anonymously via the computer interface with the other subjects in her group. The identities of the other group members were not revealed and could not be deduced, because three groups were playing in parallel. There was no possibility of verbal communication with others; the only information transmitted was on the actual choices. Subjects were first read aloud and followed on their screens instructions explaining the structure of the entire session, worked through examples, and asked the experimenter questions, if any. The instructions are provided in the Appendix.[14] The subjects were students who were recruited at the University of Erfurt using the Orsee System (Greiner, 2004).[15] The experiment was conducted with the z-tree Software package (Fischbacher, 2007). Subjects were paid privately after the end of the experimental session with an exchange rate of € 3 for each 100 experimental currency units. They earned on average € 21.

2.5 Results

Evaluation of the data shows that there are no significant effects associated with whether the OP or the PO order is used in phases II and III, in particular the investments in monitoring and the contribution levels are

[14]See also http://data.cereb.eu/monitoring.html for additional materials including the tables used to show payoffs in the treatment discussed in Section 3.

[15]http://www.orsee.org/

not significantly different.[16] Therefore, we analyze the pooled treatments PO25MC1 and OP25MC1 as 25MC1 and the pooled treatments PO45MC1 and OP45MC1 as 45MC1. In each of the pooled treatments we now have 12 independent observations. Discussion of treatment PO45MC3 is postponed to Section 2.5.3.

2.5.1 Effects of monitoring in the first half

The results of the first phase of play, in which subjects interact in a classical public goods environment, are well in line with the observations from numerous previous experimental studies of VCMs. Average contributions start off at about half of the endowment – on average 5.5 over all groups – and decrease from there on. In all four treatments we observe a negative trend in contributions over time[17] with an average contribution of 3 in the last period of Phase I. This deterioration of contributions is in line with past experiments[18] and illustrates Alchian's and Demsetz's intuition about free-riding if monitoring is absent, yet departs (as is typical in VCMs) from the strict theoretical prediction of zero contributions assuming payoff-maximizing agents.

Which division rules were implemented? Peers managed to supply incentive-imparting division rules 95% of the pooled cases of the two treatments, while the observers did so in 82.5% (see Figure 2.2a).[19]

Result 1. *Failure to achieve a division rule providing incentives to contribute the full endowment occurred less often in the peer than in the observer monitoring phases of the first half.*

At first glance, it comes as a surprise that despite their coordination problem team members succeeded more often in achieving an incentive compatible allocation rule than observers. But high "success rates" as observed in our experiment are well in line with observations in binary step-level/threshold public good experiments:[20] van de Kragt et al. (1983)

[16]The difference in average contributions and monitoring between OP and PO are not different at 10% level (exact Wilcoxon–Mann–Whitney test over the independent observations, two-sided) with one exception: OP25MC1 vs. PO25MC1 is significantly different in the monitoring investments of the observer in the observer monitoring phase ($p = 0.0065$), but has no significant effect in contributions. The average monitoring was 2.2 in PO25MC1 and 3.4 in OP25MC1.

[17]A linear regression shows a negative time trend in contributions for Phase I. The regression is performed with the average (per group) contributions of Phase I as the dependent variable. The time coefficient is significantly negative with at least 5% for all treatments (robust, Huber–White standard errors).

[18]See again Ledyard (1995) as well as Davis and Holt (1993) for a review of the literature on VCM experiments.

[19]This difference is significant ($p = 0.044$, two-sided exact Wilcoxon signed rank test over the independent observations).

[20]See Croson and Marks (2000) for a review.

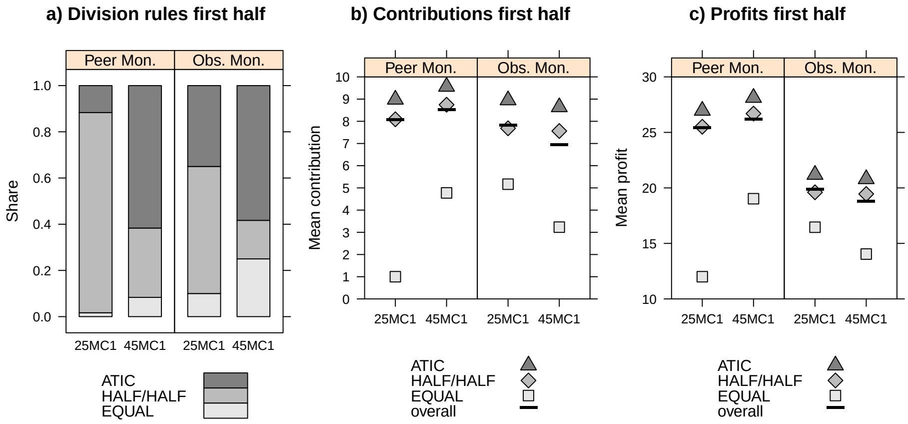

Figure 2.2: a) Frequency of implemented division rules; displayed are averages over the observations in the peer and observer monitoring phases of the first half in the 25MC1 and 45MC1 treatments. **b)** Average contributions over the observations in the peer and observer monitoring phases of the first half in the 25MC1 and 45MC1 treatments. **c)** Average team member payoffs in the peer and observer monitoring phases of the first half in the 25MC1 and 45MC1 treatments. The team member payoffs are net payoffs, meaning that the individual team member's monitoring costs are already deducted in the peer monitoring phase and the observer's share is deducted in the observer monitoring phase.

for example noticed a 72% rate in the treatment with a step return[21] like our 45MC1 treatment. Offerman et al. (1996) observed lower success-rates with step returns quite similar to 25MC1 and 45MC1. The observation of Croson and Marks (2000) that subjects respond to the size of the step-return may explain the difference in the success rates of 25MC1 and 45MC1.

How did the teams manage the high investments in monitoring? The only symmetric equilibria with positive investments in monitoring are the two mixed strategy equilibria (see Table 2.1). Figure 2.3 shows the distributions of total monitoring investments M in each of both symmetric mixed strategy equilibria and the actually observed distribution, for each of the three parameterizations in the first as well as in the second half.

In each of the six cases, displayed in Figure 2.3, the generalized Fisher exact test rejects the hypotheses that the observed distribution coincides with the equilibrium distribution at $p = 0.0001$ for the low probability equilibrium as well as for the high probability equilibrium. The low probability equilibrium rarely meets the threshold, while the high probability equilibrium leads to frequent overprovision. The team members seem to be better in "managing" the provision problem, because they meet the

[21] The step return is a concept analyzed in Croson and Marks (2000). The step return is the fraction of the aggregated group payoff over the total contribution threshold (step return = aggregated group payoff/total contribution threshold). Assuming the subjects would fully contribute in HALF/HALF we would have a step-return of 10 in 25MC1 and a step-return of 5 in 45MC1.

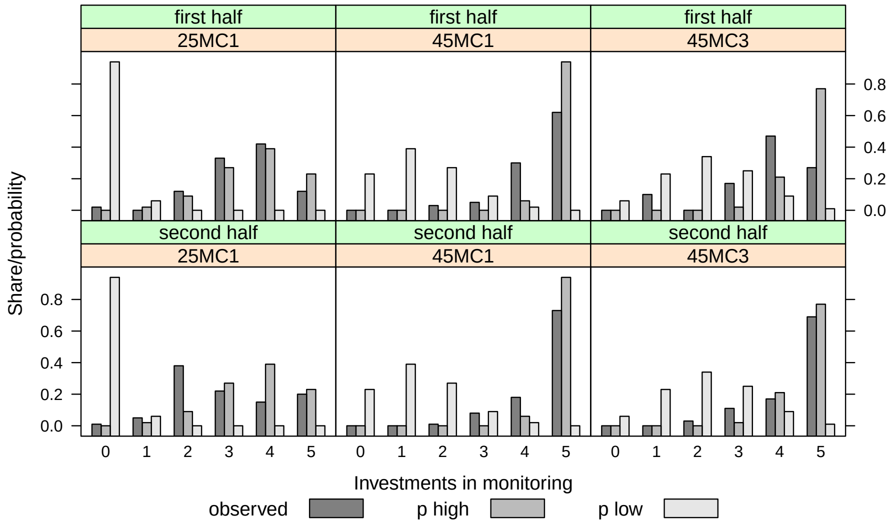

Figure 2.3: Distributions of total peer monitoring investments M in each of both symmetric mixed strategy equilibria and the actually observed distribution, for each of the three parameterizations in the first as well as in the second half.

threshold more frequently than in the low probability equilibrium, avoiding overprovision more often than in the high probability equilibrium.[22] Although team members do not seem to play one of the symmetric mixed strategy equilibria (an observation that is shared with Offerman et al. (1998)), we cannot definitely exclude that they play an asymmetric mixed strategy equilibrium. The frequencies displayed in Figure 2.3 also show that team members do not play any asymmetric pure strategy equilibrium, because then they would meet the threshold (2 and 4, respectively) and neither over- nor under-provide in monitoring.

A possible explanation of the observed behavior is that a certain fraction of subjects is guided by "non-standard" or social preferences, for example, that some subjects are conditional cooperators[23] for whom the (subjective) payoffs in a VCM may resemble those of an assurance or stag hunt game more than those of a prisoners' dilemma. Their presence could help to explain the higher-than-predicted contributions in Phase I, and likewise would account for propensities to contribute to monitoring even if coordination is difficult or if no equilibrium strategies existed, for payoff-maximizers.[24] Evidence that subjects with preference-based inclinations

[22]The theoretically expected overprovision is 89.0% for 25MC1, 93.7% for 45MC1 and 80.4% for 45MC3 in the high probability equilibrium, whereas we observe levels of 86.7% (25MC1), 61.7% (45MC1) and 26.7% (45MC3), respectively.

[23]In the sense of Fehr and Gächter (2000b) and Fischbacher et al. (2001)

[24]Duffy et al. (2007) find that subjects are not much more likely to complete a public project of fixed size when a final payoff jump causes equilibrium strategies in positive contributions to exist than when absence of such a jump makes a positive giving equilibrium

to cooperate account both for some contributions and some monitoring could be found in a significant correlation between contributions especially in the first period of Phase I, and average monitoring during a peer monitoring phase. We checked the correlation at individual subject level between monitoring investment during the exogenous peer monitoring phase and first period contribution in Phase I. Pooling the data for the two MC1 treatments, we found a significant positive correlation, meaning that the subjects with high contributions also tend to invest in monitoring (asymptotic Spearman correlation test, stratified by treatment, $p = 0.016$).

How did subjects' contributions respond to the various division rules? With peer monitoring (in Phase II) and observer monitoring (in Phase III), subjects responded to HALF/HALF and ATIC division rules with a considerable increase in average contributions compared with EQUAL division as can be seen in Figure 2.2b).[25] A comparison of HALF/HALF and ATIC exhibits two mild surprises: First, subjects contributed moderately but significantly[26] more under ATIC than under HALF/HALF. This is remarkable because a payoff-maximizing subject should contribute her full endowment under both division rules (recall that in both cases one unit of contribution is repaid by more than one, for all possible actions of the other team members). Two explanations, both invoking bounded rationality, come to mind. Firstly, although under both schemes each unit of contribution is repaid by more than 1, the private marginal return under HALF/HALF is just 1.8 while it is 3.0 under ATIC. The difference in contributions may be explained by subjects' concern for marginal returns[27] and/or social preferences. Under ATIC only the contributing team member profits from her contribution, whereas under HALF/HALF all other team members also profit (at least partly). Although it maximizes the individual payoff, a team member may (for example, due to fairness concerns) withhold contribution in order to reduce a potential free-rider's benefit from her contributions. The second surprise in comparing HALF/HALF and ATIC is that subjects tend to contribute somewhat less when the observer monitored than when the team members do.[28] Both explanations

theoretically non-existent, a result that might also be explained by the presence of some conditional willingness to cooperate. Nevertheless, the presence of a payoff jump in this chapter's linear design but its absence in our later QUAD treatment (see Chapter 3) may explain some of the difference between behaviors in these treatments.

[25]For observer monitoring contributions under EQUAL differ significantly from the ones of HALF/HALF ($p = 0.023$ two-sided exact Wilcoxon signed-rank test, 25MC1 and 45MC1 pooled). For peer monitoring there are too few occurrences of the EQUAL rule for a meaningful test.

[26]The difference is significant in both phases (Phase II: $p = 0.006$ and Phase III: $p = 0.006$, both according to the two-sided exact Wilcoxon signed-rank test).

[27]See also Isaac and Walker (1988a).

[28]The difference is significant for HALF/HALF ($p = 0.074$) as well as for ATIC ($p = 0.025$), both according to the two-sided exact Wilcoxon signed-rank test.

offered above – i.e., the sensitivity to differences in MPCRs, although exceeding 1, and fairness considerations – may also be valid to explain this phenomenon.

Result 2. *Contributions in Phase I (EQUAL division) are consistent with those in the experimental literature on the voluntary contribution mechanism. In Phases II and III, contributions are higher under ATIC than under HALF/ HALF and – ceteris paribus – contributions are higher under peer than under observer monitoring.*

In light of Result 2 and the higher costs under observer monitoring, it comes as no surprise that team members' earnings were significantly lower under observer monitoring than under peer monitoring.[29]

Result 3. *Team members earn less under observer monitoring than under peer monitoring.*

2.5.2 Voting results and consequences

One of our main research focuses is on the endogenous monitoring choice after subjects gained experience with peer as well as with observer monitoring. The voting results draw a very clear picture: the observer was never chosen by majority vote in the 36 voting rounds of treatment 45MC1 and chosen only once in the same number of votes in treatment 25MC1 (see Table 2.4). The experience from the first half of successful peer monitoring at higher payoffs may well explain why team members voted to implement peer rather than observer monitoring in the second half of their sessions. Of course, if teams had then failed to achieve sufficient monitoring to sustain contributions in later phases, they might be expected to have switched to voting for observer monitoring (see Section 2.6). But no team experienced more than one period of incentive failure during phases IV and V, so their ongoing preference for peer monitoring is not surprising.[30]

Result 4. *In the four MC1 treatments, the observer is almost never chosen by the majority vote of the team members.*

[29] $p < 0.01$, two-tailed Wilcoxon–Mann–Whitney test over the independent observations.

[30] We can find no explanation for the one instance in which three of five team members voted for observer monitoring after Phase IV, occurring in OP25MC1. Although the team in question had achieved HALF/HALF monitoring in four of five periods of Phase II with a bare two subjects monitoring (achieving ATIC one time), team members have no way to know whether 2, 3 or 4 monitored, and their earnings were higher under peer (Phase II) than observer (Phase III) monitoring in every period. Non-parametric tests for differences between the antecedents of that vote and others in the MC1 treatments are impossible since the case in question is singular.

Table 2.4: Choice of Observer or Peer Monitoring in the second half.

	Number of choices of Observer monitoring	Peer monitoring
25MC1	1 (3%)	35 (97%)
45MC1	0 (0%)	36 (100%)

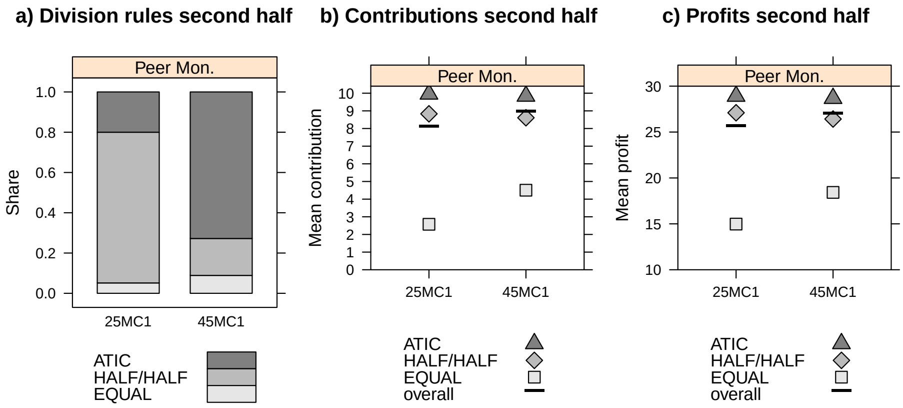

Figure 2.4: a) Frequency of implemented division rules; **b)** Contributions; **c)** Payoffs (incl. monitoring costs); displayed are averages over the observations in the second half.

How did the teams voting for peer monitoring perform? In the majority of cases team members failed to reach the exact level of monitoring necessary to provide incentives for full contributions. In 25MC1 an investment in monitoring of exactly 2 was reached in 37.7% of the cases, while in 45MC1 the equilibrium level of exactly 4 units of monitoring was only reached in 18.3% of all cases. This demonstrates the high vulnerability of monitoring to coordination failure. Nevertheless, as in the first half of the experiment in the two MC1 treatments the peer monitored groups were very successful in implementing a division rule in which full contribution to the team project is individually rational (see Figure 2.4a). They implemented HALF/HALF or ATIC in 93% of the cases, either by providing the exact level of investment necessary to give incentives for full contribution or by overinvestment (compared with the equilibrium prediction). Figures 2.4b and 2.4c additionally show that contributions as well as payoffs under both sharing rules are extremely high. However, we still observe the interesting difference between HALF/HALF and ATIC: contributions under ATIC are on average 9.9, whereas contributions under HALF/HALF are on average 8.7. The difference is significant.[31]

[31] $p < 0.001$, two-tailed exact Wilcoxon test conducted over the independent observations.

Result 5. *In the MC1 treatments peer monitoring performs extremely well: in 93% of the cases a rule capable of eliciting full contributions is reached; contributions are near 100% of endowments and payoffs are high.*

An interesting finding is that the average payoffs in 45MC1 are weakly significantly higher than in 25MC1,[32] although 4 instead of 2 units of monitoring are required to make full contribution individually rational. The reason is the extremely high number of implementations of ATIC in 45MC1 accompanied by high contributions in ATIC (see above). A likely reason is that by "overinvestment" in monitoring the risk of coordination failure is reduced at a low cost. Given the lack of verbal communication it seems practically impossible to play an asymmetric equilibrium. Thus, most team members seem to have decided to monitor every period. Not only is the average cost of over-monitoring to each subject (statistically speaking) only one unit every five periods, but in practice that cost is not "wasted", given that subjects respond to ATIC with more effort than to HALF/HALF. The histograms of the total investment M in Figure 2.3 show systematic "overinvestment" in monitoring.

2.5.3 Raising the bar – a further test

As we have seen above, team members seem to reduce the risk of coordination failure by "overinvestment" in monitoring, which is a less costly way of achieving an incentive compatible division rule than "hiring" the observer. In the light of these results we extended our analysis by conducting a new treatment PO45MC3 which is identical to PO45MC1, with the only exception that for the peers the cost of one unit of monitoring is raised to 3 (i.e., $\kappa = 3$). This raises the bar for peer monitoring: it increases the cost of implementing the HALF/HALF rule from 4 to 12, triples the cost of implementing ATIC from 5 to 15, and it also triples the cost to the individual of adhering to an "overinvestment" strategy.[33] Notice that the observer's cost remains at 1 per unit of monitoring (i.e., $\kappa_S = 1$). We collected six independent observations in this treatment. Through this change monitoring by the observer should become more attractive.

Indeed, we observe a sharp increase in voting results implementing observer monitoring. The observer was voted for by a majority in 61% of the 18 votes. Figure 2.5a) shows that the observer implements ATIC in the majority of cases. In response to this, team members make high contributions and receive payoffs which are diluted by the observer's share of 25%. Interestingly, in those groups and phases in which peer monitoring

[32] $p = 0.043$ one-sided Wilcoxon–Mann–Whitney test conducted over the independent observations.

[33] As before, the monitoring charge is still paid out of end-of-round earnings, so it is possible to pay 3 to monitor, yet still contribute 10 to team production.

was the voting choice, team members manage to achieve HALF/HALF or ATIC in almost 90% of periods. Hence, when peer monitoring is voted by the majority of the team, the team is quite successful in providing enough units of monitoring to provide incentives for making full contributions individually rational, despite the higher costs and continued, perhaps even exacerbated, coordination problem.

What is it that makes the observer model more appealing to subjects in PO45MC3? Figure 2.5 shows the differences in the first half between those groups voting for the observer later on (vote O) and those who did not (vote P). It is clear from figure 2.5a that there were more failures to achieve HALF/HALF or ATIC under exogenous peer (observer) monitoring, in groups that eventually voted for observer (peer) monitoring. Those groups that voted for peer monitoring experienced higher average contributions under peer monitoring in the first phase, while those who voted for observer monitoring experienced higher contributions under observer monitoring in the first phase (see Figure 2.5b). The same tendency is observed when looking at profits (see Figure 2.5c).

Result 6. *If the unit cost of peer monitoring is raised to* 3*, the majority of teams vote for observer monitoring. However, almost* 40% *still vote for peer monitoring and perform well, out-earning those who hire the observer.*

2.6 Conclusion

We modeled team production as a process that varies in incentive features from a pure public goods game with free-riding incentives to a privately profitable opportunity with payment in proportion to contribution. Thus, the incentive to contribute was a function of costly investment in a process denoted monitoring. We compared two institutions: in observer monitoring the monitoring is provided by a specialist who is compensated with a share of the team output. In peer monitoring the monitoring is provided by the production team members, who benefit from providing monitoring insofar as the better incentives it brings about lead to more contributions to production and hence to higher earnings. We investigated the claim that monitoring is usually provided by a residual-claiming specialist because team members have insufficient incentives and/or ability to coordinate on the provision of monitoring, and thus fail to provide adequate incentives to contribute effort to team production.

In our main model and experiment, incentives for peer monitoring are potentially adequate, but there exists a severe coordination problem. These conditions make success in peer monitoring at least improbable. Our experimental subjects were surprisingly successful in peer monitoring,

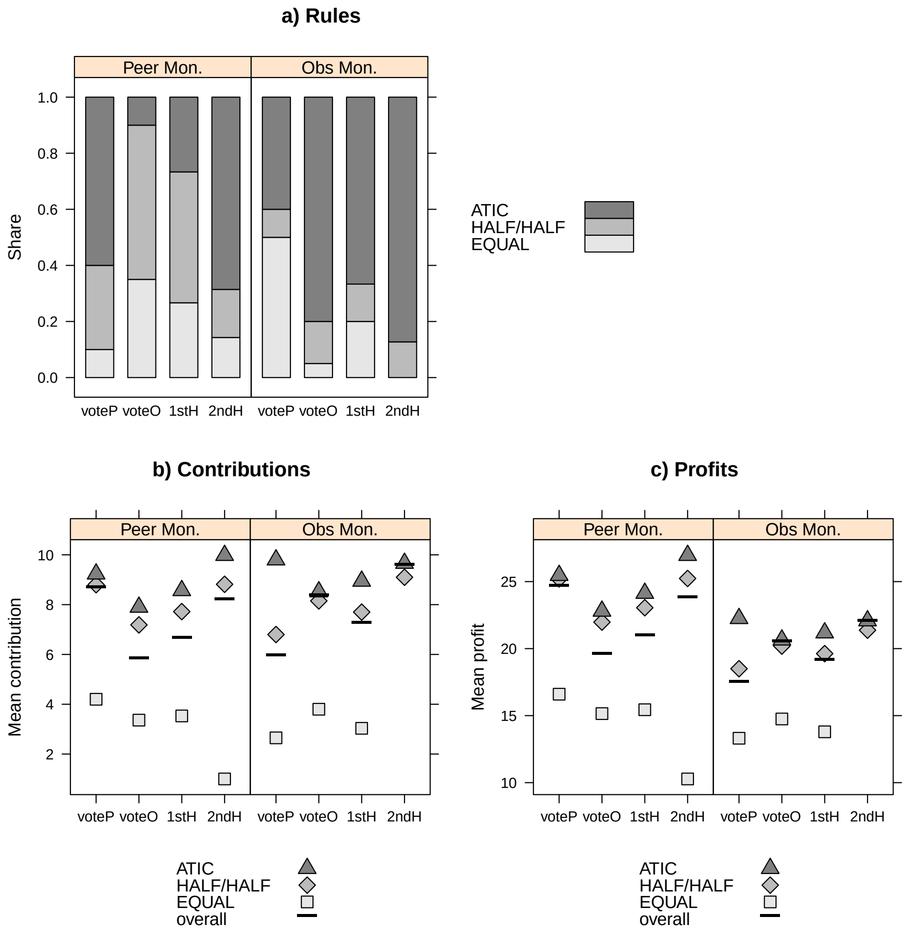

Figure 2.5: a) Frequency of implemented division rules; displayed are averages over the observations in the first half of PO45MC3 comparing those groups voting for the observer (vote O) with those who voted for peer monitoring (vote P), the overall share of the rules in the first half (1st H) and the second half (2nd H). **b)** and **c)** Averages of the contributions resp. profits over the first half of those groups voting later for the peer monitoring (vote P) or observer monitoring (vote O) plus the averages of the first and second half without the voting decision distinction.

eschewing the opportunity to use a specialist monitor almost every time they choose between the two options in treatments in which monitoring was equally costly to both peers and observer. Only when monitoring costs of team members were raised dramatically were there a substantial number of peer monitoring failures and thus votes for a specialist monitor. Even in the treatment with higher monitoring costs for team members, some groups succeeded in peer monitoring and earned substantially more than those using a specialist, despite the higher cost. In the pure public good case, subjects showed hesitation to resort to specialist monitoring, but there were clear signs of evolution in that direction, rendering our laboratory a faithful incubator of firms with residual-claiming central monitors like those in Alchian's and Demsetz's theory.

Our experiment is the first to nest the VCM or public goods game within a set of team incentive conditions, and to make the choice of organizational form or incentive regime an endogenous one. Our subjects behaved rationally in that they usually voted for the institution that gave them the highest earnings. However, their success at peer monitoring seems unlikely to be explained by individually rational behavior of the kind modeled in standard theory. Given the severe difficulty of coordinating an efficient monitoring strategy, many subjects seemed to adopt an "over-provision" strategy which should in theory invite free-riding but may not have done so in practice to the extent expected because of conditional willingness to cooperate. Conditional cooperation has been found among many subjects in recent VCM experiments, and it may have been enhanced in the present experiment by the desire to avoid ceding a significant share of output to a specialist monitor.

While our results cannot explain why mutual monitoring and profit-sharing is usually not relied upon as the main method of eliciting effort from workers in most actual firms, they are consistent with the fact that when profit-sharing is introduced, it is often successful at raising productivity (Weitzman and Kruse, 1990; Craig and Pencavel, 1995). A typical claim of writers on the topic is that despite the free-riding incentives that some associate with profit-sharing (Baker et al., 1988), workers in many firms respond to it by mutually monitoring one another's effort and working harder (Kruse, 1993) either because of a psychological identification with the firm's "bottom line", or to avoid the reproach of fellow workers (Kandel and Lazear, 1992). Thus, although one of our treatments succeeded in validating the logic of Alchian and Demsetz, the conditions under which mutual monitoring fails may be somewhat special, and the facts that most residual claims are not held by workers and that firms employ substantial amounts of top-down monitoring may be better explained by factors other than an inclination of workers to free-ride in the provision of monitoring.

2.7 Appendix

2.7.1 Symmetric mixed strategy equilibrium in the discrete version of the model

Let M^* be the lowest threshold of monitoring with $\beta > 1$ (i.e., individual rationality of full contributions) and p be the probability of investing in monitoring. Thus for any player the probability that exactly k other players invest is $\binom{N-1}{k} p^k (1-p)^{N-1-k}$.

Case 1: If $M < M^* - 1$ even by the investment of one player the threshold is not met. Then the payoff from investing in monitoring $\pi_{invest} = e - \kappa$ and the payoff from not investing in monitoring $\pi_{not\ invest} = e$ just differ by the costs of monitoring, because in both cases all players contribute $c^* = 0$ in the equilibrium of the contribution subgame. Thus, $\pi_{invest} - \pi_{not\ invest} = -\kappa$. The probability of $M < M^* - 1$ is given by

$$Pr(M < M^* - 1) = \sum_{k=0}^{M^*-2} \binom{N-1}{k} p^k (1-p)^{N-1-k}.$$

Case 2: If $M > M^* - 1$ the threshold is already reached. An extra investment of another team member would decrease the team member's profit by the cost of investing in monitoring, because even without the extra investment all players already contribute their entire endowment in the contribution subgame. Thus $\pi_{invest} = e - \kappa + (R-1)e = R \cdot e - \kappa$, $\pi_{not\ invest} = e + (R-1)e = R \cdot e$ and $\pi_{invest} - \pi_{not\ invest} = -\kappa$. The probability of $M > M^* - 1$ is

$$Pr(M > M^* - 1) = \sum_{k=M^*}^{N-1} \binom{N-1}{k} p^k (1-p)^{N-1-k}.$$

Case 3: If $M = M^* - 1$ then an investing player is pivotal, i.e., by her investment the player enables the team to exactly meet the threshold. Without the pivotal player's investment, all team members contribute zero in the equilibrium of the contribution subgame, while with the pivotal player's investment all team members contribute the complete endowment in the equilibrium of the contribution subgame. Therefore, $\pi_{invest} = R \cdot e - \kappa$, $\pi_{not\ invest} = e$ and $\pi_{invest} - \pi_{not\ invest} = (R-1)e - \kappa$. The probability of $M = M^*$ is

$$Pr(M = M^* - 1) = \binom{N-1}{M^*-1} p^{M^*-1} (1-p)^{N-M^*}.$$

For the symmetric mixed strategy equilibrium the expected payoff from investing has to equal the expected payoff from not investing, i.e.:

$$\begin{aligned}
& \sum_{k=0}^{M^*-2} \binom{N-1}{k} p^k (1-p)^{N-1-k} \cdot (-\kappa) + \cdots \\
+ \quad & \sum_{k=M^*}^{N-1} \binom{N-1}{k} p^k (1-p)^{N-1-k} \cdot (-\kappa) + \cdots \\
+ \quad & \binom{N-1}{M^*-1} p^{M^*-1} (1-p)^{N-M^*} \cdot ((R-1)e - \kappa) = 0 \,. \qquad (2.7)
\end{aligned}$$

This equivalently transforms into

$$\sum_{k=0}^{N-1} \binom{N-1}{k} p^k (1-p)^{N-1-k} \cdot \kappa = \binom{N-1}{M^*-1} p^{M^*-1} (1-p)^{N-M^*} \cdot (R-1)e$$

which reduces further to

$$\frac{1 \cdot \kappa}{(R-1)e} = \binom{N-1}{M^*-1} p^{M^*-1} (1-p)^{N-M^*} . \qquad (2.8)$$

2.7.2 Translation of the Instructions for Treatment PO25MC1

Instructions to the Experiment

General Information
At the beginning of the experiment you will be randomly assigned to one of 4 **groups**. During the whole experiment you will interact only within your group. Each group contains 5 **members** and 1 **observer**. You will be informed whether you are a member or an observer. You keep your role during the whole experiment. The instructions explain the actions and payoffs of both the members and the observer.

Course of Action
The experiment consists of 30 **rounds** which are divided into 6 phases of 5 rounds each.

Payoff
The total payoff from the experiment consists of the sum of round payoffs of 30 rounds. At the end of the experiment your total payoff will be converted at an exchange rate of €3 per 100 tokens.

Please note
Communication is not allowed during the whole experiment. If you have some question please raise your hand out of the cabin. All decisions are made anonymously, i.e., no other participant learns the identity of the other decision makers. The payment is anonymous and takes place immediately after the experimental session.

Phase 1 Group Project

Member
In each round of Phase 1, every group member receives an **endowment of** 10 tokens. You have to decide how many of these 10 tokens you contribute to the group project. The remaining tokens are assigned to your private account. The tokens contributed to the group project are **tripled** and equally distributed among all group members, while the tokens in the private account solely benefit the member.

Calculation of a *member's round* payoff in phase 1
A member's round payoff consists of two parts:

- **earnings from the group project** = 3 x sum of the contributions of all group members/ number of group members

- **tokens in the private account** = endowment – member's contribution to the project

Round **payoff of a member in phase 1**:
10 – member's contribution to the project +
+ 3 x sum of the contributions of all group members / 5

Here is an **example**:

Member's contributions	Sum of contributions	3x sum of conributions/5	Member's payoff = 10 – contribution + 3 x sum contributions of all groups members/5 (equal shares)
0	25	15	25
3	25	15	22
5	25	15	20
7	25	15	18
10	25	15	15

Observer

In each round of phase 1 the observer is asked to guess the sum of the contributions of "her" group.

Calculation of the *Observer's Round* Payoff in Phase 1

The observer's payoff depends on the correctness of her guess. The closer the observer's guess is to the actual sum of contributions the higher is the observer's payoff.

The following example illustrates this:

Sum of contributions	25	25	25	25	25	25	25	25
Observer's guess	0	5	10	20	25	30	35	50
Observer's guessing error	25	20	15	5	0	5	10	25
Observer's profit	13.33	15.00	17.14	24.00	30.00	24.00	20.00	13.33

Notice that the maximum which the observer can earn is 30 token and that the amount that the observer loses for guessing errors is the same for guessing too high a number as for guessing too low a number.

Information at the End of the Round

At the end of the round each member as well as the observer is informed about the sum of the group members' contributions. The members are additionally informed about their individual payoff. The observer is additionally informed about the error in her guess and the resulting payoff.

Phase 2 Investment in Verification by Group Members

Member

In each round of phase 2 every group member receives an endowment of 10 tokens. Prior to the contribution decision each member may invest zero or one token in **verification** which will be deducted at the end of the period. The total investment in verification of all group members influences the way in which the earnings from the group (the tripled sum of contributions) project will be divided. There are three possibilities:

Sharing rule 1 "Equal shares": As in Phase 1 the earnings from the group project will be divided **equally among all members if** 0 or 1 member invested in verification (like Phase 1)

Sharing rule 2 "Half/half": Half of the earnings of the group project will be divided **equally among the group members** and the other **half in proportion to the individual contributions** of the group members **if** 2, 3, or 4 members invested in verification

Sharing rule 3 "Proportionate": The **entire earnings** from the group project will be allocated **in proportion to the individual contributions if** all 5 members invested in verification

After each member has completed her investment in verification each member as well as the observer is informed on the sharing rule to be applied. Then group members decide how many of their endowment of 10 tokens to invest in the group project (as in Phase 1).

Calculation of a *Member's* Round Payoff in Phase 2

A member's payoff in Phase 2 is calculated according to:

$$\begin{array}{l} 10 \\ -\ \text{contribution} \\ -\ \text{verification investment} \end{array} + \begin{cases} \text{Equal shares of the tripled sum of contributions } \textbf{if} \text{ total investment in verification: } \mathbf{0} \text{ or } \mathbf{1} \\ \text{half in equal shares + half in proportionate } \textbf{if} \text{ total investment in verification: } \mathbf{2}, \mathbf{3} \text{ or } \mathbf{4} \\ \text{member's contribution tripled (propotionate) } \textbf{if} \text{ total investment in verification: } \mathbf{5} \end{cases}$$

Here is an **example**:

Observer

In each round of Phase 2 the observer has to decide how many tokens he/she wants to invest in verification of "her" group. The calculation of the observer's payoff is as in Phase 1.

Member's contributions	3× sum of contributions/5 (Equal share)	Sharing rule 1 "Equal-shares"	Payoff* according to Sharing rule 2 "Half/half"	Sharing rule 3 "Proportionate"
0	15	25	17.5	10
3	15	22	19.0	16
5	15	20	20.0	20
7	15	18	21.0	24
10	15	15	22.5	30

* the payoff still has to be reduced by the members' investment in verification

Information at the End of the Round

At the end of the round each member as well as the observer is informed about the sum of the group members' contributions. The members are additionally informed about their individual payoff. The observer is informed about the error in her guess and the resulting payoff.

Phase 3: Verification by the *Observer*

In Phase 3 the **observer decides upon the level of verification**. The observer receives an endowment of 5 tokens and may invest from 0 up to 5 tokens in verification and keep the rest. Unlike the previous phases, the observer is not asked to guess the sum of the members' contributions. Instead, the observer receives **25**% of the **tripled** amount of the sum of all members' contributions. The remaining **75**% of the **tripled** amount of the sum of all members' contributions are distributed among the group members.

As in Phase 2 the investment in verification influences the way in which the earnings from the group project (that is, 75% of the tripled amount, or $2\frac{1}{4}$ times the sum of the contributions) will be divided among the members. Again the three sharing rules described above are possible:

Sharing rule 1 "Equal shares"	**if**	the observer invests 0 or 1 tokens in verification
Sharing rule 2 "Half/half"	**if**	the observer invests 2, 3, or 4 tokens in verification
Sharing rule 3 "proportionate"	**if**	the observer invests 5 tokens in verification

Member

The members are informed on the sharing rule to be applied according to the observer's investment in verification. The members decide how many of their 10 tokens they invest into the group project.

Calculation of a *Member's* Round Payoff in Phase 3

$$10 - \text{contribution} + \begin{cases} \frac{3}{4} \text{ of equal shares of the tripled sum of contributions } \textbf{if} \text{ the observer invested } \mathbf{0} \text{ or } \mathbf{1} \text{ in verification} \\ \frac{3}{4} \text{ (half in equal shares + half in proportionate share) } \textbf{if} \text{ the observer invested } \mathbf{2, 3} \text{ or } \mathbf{4} \text{ in verification} \\ \frac{3}{4} \text{ of member's contribution tripled (propotionate share) } \textbf{if} \text{ the observer invested } \mathbf{5} \text{ in verification} \end{cases}$$

Here is an **example**:

	3× sum of		Payoff according to	
Member's	contribu-	Sharing rule 1	Sharing rule 2	Sharing rule 3
contribu-	tions/5	"Equal-	"Half/half"	"Proportio-
tions	(Equal share)	shares"		nate"
0	15	21.25	15.63	10.00
3	15	18.25	16.00	13.75
5	15	16.25	16.25	16.25
7	15	14.25	16.50	18.75
10	15	11.25	16.88	22.50

Observer

In each round of Phase 3 the observer has to decide how many tokens he/she wants to invest in verification of "her" group.

Calculation of an *Observer's* Round Payoff in Phase 3

The payoff of the observer in Phase 3 consists of the endowment of 5 minus her investment in verification plus 25% of the tripled amount of the sum of all members' contributions:

Round **payoff of an observer in Phase 3**:
$= 5 - \text{investment in verification} + \frac{1}{4} \times 3 \times \text{sum of member's contributions}$

Information at the End of the Round

At the end of the round each member as well as the observer is informed about the sum of the group members' contributions. The members are additionally informed about their individual payoff. The observer is additionally informed about the sum of the profits and the resulting payoff.

Phase 4–6: Voting for the Verification System

At the beginning of Phases 4, 5, and 6 group members vote for one of two possibilities: verification by the group members (as in Phase 2) or verification by the observer (as in Phase 3). For the next 5 rounds the verification type that was preferred by the majority of the group members is implemented. In other words, group members have the choice whether they want to implement the mechanism of Phase 2 (that is, provide their own verification, if any) or the mechanism of Phase 3 (that is, "hire the observer", who provides verification, if any) again for each of these sets of 5 rounds. The observer cannot vote.

We wish you success!

3

The Zero Monitoring Equilibrium in Team Production[1]

Abstract

We introduce a non-linear production function to a team-production monitoring experiment similar to that of Chapter 2. The non-linearity changes the monitoring problem into a second-order linear public-good. This prevents the existence of positive peer monitoring equilibria and should favor central monitoring. In this new experimental setting the peer monitoring, although preferred in the beginning, quickly loses out against the residual-claiming agent.

3.1 Introduction

In Chapter 2 we presented a model for team production along the lines of Alchian and Demsetz (1972). In team production there exists a free-riding incentive if the team members are compensated according to the team and not the individual output. The collective action problem in this context should be mitigated by the employment of a residual claimant. This central agent takes over the decision of the investment into a monitoring mechanism that makes the individual output attributable and in turn incentivizes cooperation. Of course, this investment could be made by the team members themselves; however the monitoring investment in itself is either a coordination problem like the one we modeled in the previous chapter or as a public-good problem in itself. We refer to Section 2.2 for a review of the relevant literature of the team production context.

If the monitoring is a step-level public-good and the team production modeled through a linear production function, then there exist equilibria

[1]This chapter is part of "Monitoring in Teams: Using Laboratory Experiments to Study a Theory of the Firm" by Stefan Grosse, Louis Putterman and Bettina Rockenbach (2011) and published in the *Journal of the European Association*, Vol.9(4), (Grosse et al., 2011). All authors contributed equally.

with positive levels of monitoring as we did show in Section 2.3. The intuition behind this is that there exist strategy combinations in which a single (pivotal) player's investment in monitoring causes an increase of the marginal per capita return above 1, transferring the public-good into a private-good and making the investment profitable and, as a consequence, making positive investments in monitoring individually rational. Observing peer monitoring might thus be the result of equilibrium play or the attempt to do so.

These equilibria seem to be one reason that make the peer monitoring more attractive than the central agent monitoring in Chapter 2 since it saves the peers the residual claim, and indeed the self monitoring was very successful in that context. Most teams consequently did not choose the central agent but managed the monitoring themselves. This somewhat resembles the results of Frohlich et al. (1998), who observed that employee owned firms can be more efficient than "conventional" firms in an real effort experiment. Contrary to that Potters et al. (2009) found that managers might induce a higher effort. Their manager-less experiment resembles a pure public-good experiment, which gives the team members hardly a chance to mitigate the free-riding problem.

We wondered whether the tendency of team members to pay by themselves for monitoring like they did in Chapter 2 – despite the temptation to let others do the job – would survive a yet harder challenge: a situation in which the only equilibrium in monitoring involves no monitoring at all. To model peer monitoring as a pure public-good problem, we need a specification in which the gains from monitoring lack the discrete jump that can make the marginal unit of investment into monitoring privately profitable. To achieve this we develop a model of team production in which the investment in monitoring itself is linear public-good problem with free-riding incentives and thus lacks any equilibria with positive monitoring expenses. This cannot be done with linear costs of contributions, thus we reformulated the team production problem of Chapter 2 into one with a quadratic contribution cost model. Quadratic cost functions in the context of public-good experiments were used for example in Isaac and Walker (1988b), Irlenbusch and Ruchala (2008), Keser (1996), Sefton and Steinberg (1996). As a second change, we eliminate the step-like relationship of the previous Chapter between payoffs and monitoring in the sense that a certain number of investors is needed to get the monitoring implemented, now allowing the share of the privately attributable production increase incrementally by the number of team members that decided to invest in the monitoring mechanism.

Intuitively, the quadratic cost function causes the marginal return to effort to decline as monitoring induces more effort, rendering monitoring individually unprofitable at the margin despite the fact that an outcome with more monitoring and higher effort would be collectively preferable – a classic social dilemma. Hence, our game consists of the sequence of two social dilemmas: firstly the dilemma in monitoring and secondly the dilemma in contributing. Overcoming the monitoring-dilemma might change the contribution stage into an incentive compatible investment problem. However, in an equilibrium of profit-maximizing agents the players refrain from monitoring as well as from contributing.

3.2 A Model of Team Production with a Non-Linear Production Function

We model team production with the following individual profit π_i for player i:

$$\pi_i = e - \kappa \cdot m_i - \frac{c_i^2}{f} + \beta \cdot c_i + \gamma \cdot C_{-i}\,. \tag{3.1}$$

with e as the endowment, κ the cost of the binary ($m_i \in \{0, 1\}$) monitoring decision, c_i the contribution of player i, f as a cost function parameter and $C_{-j} = \sum_{j=1, j\neq i}^{N} c_j$. As prior to introducing the discrete version of the model of Section 2.3, γ and β are again $\gamma = R/N^2 \cdot (N - M)$ and $\beta = R/N^2 \cdot (N - M + N \cdot M)$ where M is the number of the N group members that invested in monitoring and R represents the team production multiplicator. During the observer monitoring regime the β and the γ in Equation (3.1) are replaced by their by the residual claim S reduced counterparts β^S and γ^S.

The joint payoff level of all team members is:

$$\Pi = \sum_{i=1}^{N} \pi_i = N \cdot e - \sum_{i=1}^{N} \frac{c_i^2}{f} - \kappa \cdot M + R \cdot C \tag{3.2}$$

leading to player i's socially optimal contribution of

$$c_i^{so} = \frac{fR}{2}\,. \tag{3.3}$$

What about the individual incentives for monitoring and contributing? For given investments in monitoring m_j we derive the individually optimal contribution c_i as:

$$c_i^*(M) = \frac{f}{2} \cdot \frac{N + M(N-1)}{N^2} R = \frac{f}{2} \beta \tag{3.4}$$

Hence, the individually optimal contribution only depends on the sum of all monitoring expenses M. Obviously, for $N = M$ the socially and the individually rational contribution levels coincide ((3.3) equals (3.4)). Hence, full monitoring ensures that a payoff-maximizing subject contributes the socially optimal amount.

But is it in the self-interest of individuals to invest in peer monitoring? An individual invests in peer monitoring if – ceteris paribus – the payoff difference between the optimal contribution level with and without monitoring investment is positive, i.e.:

$$\begin{aligned}
\Delta\pi_{invest} &= \pi_i(c_i^*(M_0+1)) - \pi_i(c_i^*(M_0)) \\
&= -\kappa - \frac{f}{4}R^2\left[\left(\frac{N+(N-1)(M_0+1)}{N^2}-1\right)^2-1\right]+\dots \\
&\quad +\frac{f}{4}R^2\left[\left(\frac{N+(N-1)M_0}{N^2}-1\right)^2-1\right] \\
&= -\kappa - \frac{f}{4}R^2\left[\left(\frac{N+(N-1)(M_0+1)}{N^2}-1\right)^2-\dots\right. \\
&\quad \left.-\left(\frac{N+(N-1)M_0}{N^2}-1\right)^2\right] \\
&= -\kappa - fR^2\frac{(N-1)^2(2M_0-2N+1)}{4N^4} \\
&= -\kappa + \frac{fR^2(N-1)^2(2N-1)}{4N^4} - \frac{fR^2(N-1)^2}{2N^4}M_0. \qquad (3.5)
\end{aligned}$$

If $\Delta\pi_{invest} \leq 0$ no team member invests in monitoring and there will be no monitoring in equilibrium. If the level of monitoring is zero, then the equilibrium contribution is $c_i^N = (fR)/(2N)$, lower than socially optimal.

What are the incentives of the observer in this model? The payoff of the observer during the observer monitoring phase is

$$\pi_o = e_o - \kappa_o \cdot M_o + S \cdot R \cdot C.$$

The peers' payoff under observer monitoring is

$$\pi_i^o = e - \frac{c_i^2}{f} + (1-S)(\beta \cdot c_i + \gamma \cdot C_{-i}).$$

For a given monitoring level M_o the team members' optimal contribution level is

$$c_i^{o*}(M_o) = \frac{f}{2}\frac{N+M_o(N-1)}{N^2}(1-S)R = \frac{f}{2}\beta^S.$$

This contribution level gives the observer the payoff

$$\pi_o(c_i^{o*}(M_o)) = e_o - \kappa_o \cdot M_o + S \cdot R \cdot N \cdot c_i^{o*}(M_o).$$

Is it in the self-interest of the observer to invest in peer monitoring? The observer will invest in monitoring if – ceteris paribus – the payoff difference between the optimal contribution level with and without monitoring investment is positive. The one unit increase of the monitoring investment causes a difference in the observer's payoff of

$$\begin{aligned} \Delta_{\pi_o(M_o)} &= \pi_o(c_i^*(M_o+1)) - \pi_o(c_i^*(M_o)) \\ &= -\kappa_o + \frac{1}{2} f R^2 (S - S^2) \left(1 - \frac{1}{N}\right). \end{aligned} \quad (3.6)$$

As we see, if $\kappa_o < fR^2/2(S - S^2)\,(1 - 1/N)$ the observer will choose full monitoring (i.e., $M = N$), while with $\kappa_o > fR^2/2(S - S^2)\,(1 - 1/N)$ the observer will not monitor at all (i.e., $M = 0$).

If the observer fully invests in monitoring, the individually rational contribution levels of the team members are higher than in the equilibrium of peer monitoring with $\Delta\pi_{invest} \leq 0$: $c_i^{o*}(M_o = N) = (1 - S)\frac{fR}{2} > c_i^N = (fR)/(2N)$ as long as $(1 - S) > 1/N$ (equivalently $S < (1 - 1/N)$). For $\Delta\pi_{invest} \leq 0$ the efficiency under peer monitoring in the Nash equilibrium is lower than the efficiency under observer monitoring as long as the same condition $S < (1 - 1/N)$ holds. The efficiency under peer monitoring is

$$\frac{4e + fR^2\left(-1/N^2 + 2/N\right)}{4e + fR^2},$$

while the efficiency under observer monitoring (including the observer) is

$$\frac{4e + fR^2(1 - S^2)}{4e + fR^2}.$$

The team member Nash equilibrium profit in peer monitoring would be

$$\pi_i^N = e - \frac{fR^2}{4} \cdot \left(-\frac{1}{N^2} + \frac{1}{N}\right)$$

and for the observer monitoring

$$\pi_i^o = e - \frac{fR^2}{4} \cdot (1 - S)^2.$$

Thus the team members' profit is higher with observer monitoring than with peer monitoring as long as $S < 1 - (\sqrt{2N - 1})/N$.

To sum up, with the parameters of our experimental study of the quadratic model the team members have no incentive to monitor and thus team production remains a voluntary contribution problem. In the subgame with observer monitoring, however, there will be full monitoring in equilibrium. This results in team members' payoffs which are – despite the observers' share – higher than under peer monitoring. Hence team members have an incentive to enter the subgame, i.e., to hire the observer.

3.3 Experimental Setup and Prediction

The experimental design, which we dub QUAD for its quadratic cost function, is quite similar to that of Chapter 2. In each session the subjects were randomly assigned to a group of six. One subject was randomly determined as "observer" the other 5 as "members" for the whole experiment. For the first five periods the subjects were experiencing the public-good problem without a monitoring possibility. This was followed by a peer monitoring and an observer monitoring phase with a length of five periods as well. Since we did not find a significant effect of the order of the phases in Chapter 2, we did not alternate the order again. In the last 15 periods the subjects could vote three times on the monitoring mode for 5 consecutive periods. Thus we followed the schematic representation of Figure 2.1 on page 40.

The profit was calculated according to equation 3.1. A group consisted again of five members. The cost κ of the investment was 3.5 for the peers and the team production factor $R = 3$. In the observer monitoring phase the observer had to bear the (lower) costs but received a share of $S = 20\%$ of the team production. The cost function parameter f was specified with 6.377. The number had to be such a non-integer number in order to achieve that full contribution is socially optimal and to prevent larger jumps in the profit levels in dependence on the monitoring investments.

We collected eight independent observations with a total of 48 subjects. The experiment was conducted at the elab-laboratory of the University of Erfurt using the Orsee-system (Greiner, 2004) for the recruitment of the subjects and z-tree software (Fischbacher, 2007) for the interaction.

As has been shown in Section 3.2 assuming that the subjects act as money maximizing individuals they should never invest in monitoring because with our parameters $\Delta\pi_{\text{invest}} \leq 0$. Further the observer is incentivized to invest in the monitoring mechanism. With the observer receiving 20% of the group production during those phases in which she is exogenously assigned or chosen by vote as the monitor, she is predicted to maximize earnings by selecting $M^* = 5$. These monitoring levels imply that team members maximize their individual earnings by each selecting effort levels $c_i^* = 2$ under peer and $c_i^* = 8$ under observer monitoring, with earnings of 15.37 and 19.16 respectively. The observer, in turn, earns a maximum of 12.5 when choosing $M = 0$ and 30 when choosing $M^* = 5$, assuming that team members respond in privately optimal fashion. Because the subjects earn more with the observer employed they will vote for the observer monitoring mechanism.

Due to the results of Chapter 2 we expect that opposite to the theoretical assumption the peers will at least start with certain monitoring levels but we expect that as in common public-good experiments the monitoring investments will deteriorate due to the free-riding incentives leading to the employment of the observer.

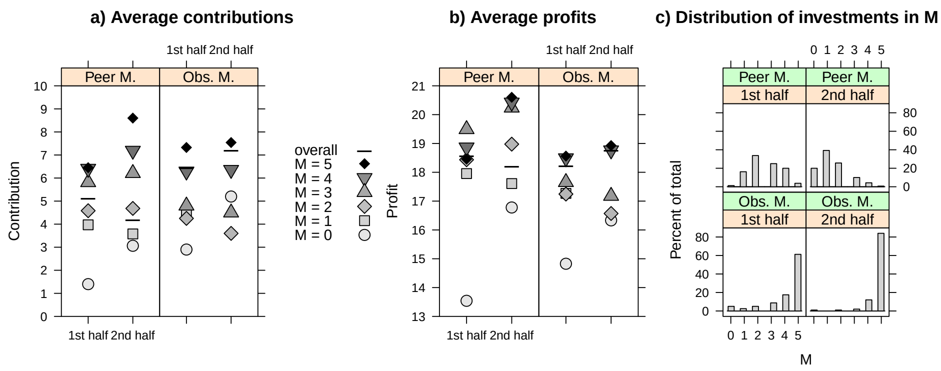

Figure 3.1: a) average contribution and **b)** profit by monitoring investment level and overall; **c)** distribution of the monitoring investment levels.

3.4 Results

Even though the parameters and quadratic cost function make monitoring more costly compared with the experiments we conducted in Chapter 2, reduce the gains from team production, and generate a pure free-rider problem, the peers still manage to achieve high monitoring levels during the peer monitoring phase of the first half (periods 6 – 10). This can be seen in Figure 3.1c. The peers' investments in monitoring ensures still a higher contribution level compared to those occurrences with no monitoring at all (compare figure 3.1). Since the observers did invest significantly more in the monitoring mechanism they achieved higher contribution levels: The mean contribution is 5.19 under first half peer monitoring, versus 6.51 under first half observer monitoring and 3.72 in the no monitoring periods of Phase I. The differences in contributions between the phases with peer (Phase II) or observer (Phase III) monitoring and the phase without monitoring (Phase I) are statistically significant, as are those between the two first-half monitoring phases (II and III):[2] Although observer monitoring in the first half achieves significantly higher contributions, the observers' claim of 20% of the team production reduces the profits. This explains why the team members' profits are on average not significantly higher with Phase II peer than with Phase III observer monitoring as within group comparisons show ($p = 0.641$ exact Wilcoxon signed-rank test conducted over the independent observations). Only three of the eight groups have a higher payoff under observer monitoring. Together with potentially some dislike of sharing with the observer this might explain why only one out of eight groups vote for the observer mechanism on the first vote round in QUAD.

[2] $p = 0.008$ for public-good vs. first half peer phase, $p = 0.008$ public-good vs. first half observer phase and $p = 0.055$ for the comparison between first half peer and first half observer monitoring. (exact Wilcoxon signed-rank test (two-sided)).

However, after the first vote the monitoring level of the peers declines considerably, as shown in Figure 3.1c.[3] This causes considerably lower contributions and profits[4] and leads to a growing tendency to choose the observer monitoring mechanism: four out of eight groups vote for the observer in the fifth phase and five out of eight groups do so in the last phase.

Result 1. *Although the introduction of a quadratic cost function, which generates a pure free-riding problem, makes cooperation more difficult, peer monitoring is still substantial in Phase II and is favored by seven of eight groups in their initial votes. The peers' monitoring investment declines in the second half of the experiment, leading to a more frequent choice of the observer as the monitoring decision maker.*

The break-down of peer monitoring with repetition resembles the decline of contributions to a public-good found in ordinary voluntary contribution experiments, which is not surprising since in the QUAD treatment peer monitoring is precisely such a public-good. Thus, while the prediction of free-riding from the outset is not supported either in the QUAD treatment or in standard finitely-repeated VCM experiments, a trend towards increased free-riding over time can be observed, which in this case leads to an increased choice of the observer as a monitor, representing the specialized residual-earning agent. The tendency that appears to emerge closely resembles that discussed by Alchian and Demsetz – i.e., insufficient incentives to engage in peer monitoring lead to the choice to organize the firm around a residual-claiming specialist monitor.[5]

3.5 Summary

We created an experimental design that should validate the emergence of the firm in a sense similar to Alchian and Demsetz (1972). This would imply that the drawback of collective action, the incentive to free-ride, could be overcome by the monitoring of a central agent who is an insider. We offer the possibility that the team production members invest in monitoring or alternatively "outsource" this decision to the central agent.

[3]Exact Wilcoxon signed-rank test: $p = 0.016$ (two-sided).

[4]Exact Wilcoxon signed-rank tests: for the differences between contributions under peer in first half and those under peer in second half: $p = 0.0555$, between contributions under peer versus contributions under observer in second half $p = 0.016$, and the two parallel tests for profits: $p = 0.383$ and $p = 0.031$, respectively (all two-sided).

[5]It has been shown elsewhere that the decline in contributions to a public-good can be prevented, delayed, or slowed by devices such as (a) permitting costly punishment of free-riders (Fehr and Gächter, 2000a; Page et al., 2005; Gürerk et al., 2006) and (b) allowing pre-play communication (e.g., Brosig et al., 2003). If such devices also slow or prevent the decline in peer monitoring, they would perhaps prevent observer monitoring from coming to be favored in the long run.

In the previous chapter we observed that the employment of this central agent, the "observer" can not be taken for granted. Her residual claim makes the centralization of the monitoring decision expensive. If the monitoring investment is cheap enough for the peers and the team is able to solve the coordination problem they will prefer to monitor themselves.

In this chapter we have transformed the monitoring investment from a step-level public-good – hence the coordination problem – into a standard public-good problem. This renders the monitoring decision individually unprofitable opposite to a linear production function as in Chapter 2. With linear production costs the decision can be profitable for a pivotal subject since her investment makes all subjects increase their contribution from zero to full, hence the coordination problem.

Though the monitoring investment is a second-level public-good for the peers they still have a preference for self-monitoring in the beginning. In the second half of the experiment, as peer monitoring decreases, the central agent proves to be more successful and in consequence gets employed more frequently. Thus we finally found a design that makes the emergence of the firm in the Alchian/Demsetz sense more likely.

3.6 Appendix

3.6.1 Translation of the Instructions

General Information

At the beginning of the experiment you will be randomly assigned to one of **4 groups**. During the whole experiment you will interact only within your group. Each group has **5 members** plus **one observer**, 6 in total. You will be informed whether you are a member or an observer. You keep your role during the whole experiment.

Course of Action

The experiment consists of 60 periods which are divided into 6 phases of 10 periods each.

Payoff

The total payoff from the experiment consists of the sum of all 60 period payoffs. At the end of the experiment your total payoff will be converted at an exchange rate of € 1 per 100 tokens.

Please note:

During the whole experiment communication is not allowed. Please switch off your cell phone. If you have any question, please raise your hand out of the cabin. All decisions are made anonymously, i.e., no other participant gets to know the identity of the other decision makers. Payments are made anonymously and take place immediately after the experimental session.

Phase 1: Group Project

Team Member

In each period of Phase 1 each group member receives an **endowment of** 10 **tokens**. A discretionary amount out of these tokens can be contributed to the group project. Every contribution to the group project causes a **cost** for the contributing team member which is calculated as follows:

Contributions	0	1	2	3	4	5	6	7	8	9	10
Total costs of contributions:	0.0	0.16	0.63	1.41	2.51	3.92	5.65	7.68	10.04	12.7	15.68

If you contribute for example 2 this will cost 0.63 token and you will keep $10 - 0.63 = 9.37$ token of the original endowment of 10. The costs can be lower or higher than the contribution. The tokens that are contributed to the group project will be **multiplied by 3** and divided equally among the group members, notwithstanding whether they contributed or not.

Calculation of a *Group Member's* Period Payoff in Phase 1

	Endowment	10
minus	**Costs of contribution**	According to the table above
plus	**Earnings from the group project**	3× sum of contributions of all group members/number of group members
=	Group member payoff	

You can gather more detailed payoffs from the attached Table 1 (Phase 1).

Observer

In each period of Phase 1 the observer will be asked to guess the sum of contributions of his group.

Calculation of the *Observer's* Period Payoff of Phase 1

The observer's payoff **depends on the accuracy of his/her guess**. The closer the observer's guess is to the actual sum of contributions the higher is the observer's payoff. The following examples illustrate this:

Sum of contributions	25	25	25	25	25	25	25	25
Observer's guess	0	5	10	20	25	30	35	50
Error of the guess	25	20	15	5	0	5	10	25
Payoff	6.7	7.5	8.6	12.0	15.0	12.0	10.0	6.7

Notice that the maximum the observer can earn is 15 and that the amount of the payoff is symmetric: it does not matter if the guessing was too high or too low, only the size of the error matters.

Information at the End of the Period

At the end of each period, the group members and the observer will be informed about the sum of contributions of all group members, as well as their respective payoffs.

Quiz

1. a) What is your payoff if you contribute 2 and the other team members contribute 2 tokens on average? b) What is your payoff if you contribute 7 instead?

2. If the other group members contribute 10 tokens on average, a) what is your payoff when contributing 10, and b) what is your payoff when you contribute 2?

Phase 2: Investment in Verification by the Group Members

Team Members

In each period of Phase 2 each group member receives an **endowment of 10 tokens**. There are two consecutive decisions each member has to make:

1. Whether the group member invests in **verification of the contributions of the other group members** or not. The investment costs 3.5 tokens which will be deducted at the end of the period.

2. How many of the 10 tokens the group member will contribute to the group project (analogous to Phase 1).

The total investment of all group members influences how the payment from the group project will be distributed. There are two "pots": the "*equal-distribution-pot*" and the "*according-to-contribution-pot*". The contents of the *equal-distribution-pot* will be divided in equal shares among the group members. The contents of the *according-to-contribution-pot* will be distributed directly according to the individual contribution.

Importance of the Verification Mechanism The investment in the verification mechanism determines the division of the payoffs from the group between the two pots:

- If no one invests in the verification mechanism all earnings from the team project go to the *equal-distribution-pot*.

- If all group members invest in the verification mechanism all earnings from the team project will go in the *according-to-contribution-pot*.

- In all other cases the earnings from the group project will be divided between the pots according to the investments in the verification mechanism (number invested/total number = share of the *according-to-contribution-pot*). For example if 2 out of 5 group members invest in the verification mechanism $2/5 = 40\%$ of the payoff will go to the *according-to-contribution-pot* and the remaining 60% to the *equal-distribution-pot* (for all).

You can gather more detailed payoffs from the attached Table 2 (Phase 2). **NOTE** that you have to deduct costs of 3.5 tokens for the investment in verification in case you have invested in the verification mechanism. In other words, the table shows the earnings before deducting the cost of your investment in verification; if you are one of those who make that investment, your earnings are 3.5 tokens less than the number in the table.

Observer

In each period of Phase 2, the observer is asked – as in Phase 1 – to guess the sum of the group contributions. Her payoff is calculated accordingly.

Information at the End of the Period

At the end of the period each member and the observer will be informed about the number of investments in the verification mechanism, the sum of contributions, the distribution to the pots and their period payoff. The observer will be additionally informed about her guess and the resulting payoff.

Quiz

1. Compare Table 2 with no investment in verification with Table 1 – is there a difference?

2. If 3 group members invest in the verification mechanism, what percentage is divided equally and what percentage is divided according to individual contributions?

3. Suppose all members invest in the verification mechanism. What is your payoff as a team member a) if you contributed 5 and the others contributed 5 on average? What is your payoff b) if you contributed 10 and the others contributed 5 on average?

4. Now, suppose the others contributed 7 on average, how do the payoffs in Question 3 change?

Phase 3: Investment in the Verification Mechanism by the Observer

In Phase 3 the observer will not be asked for her guess. Instead **the observer** determines by her investment the degree of verification of the contributions of the group members. For this purpose she will receive an endowment of 5 token. Her investment thus specifies the **division between the two pots**. The more the observer invests in the verification mechanism, the greater is the share of the *according-to-contribution-pot* of the earnings. The observer will receive 20% of the group projects' earnings. The remaining 80% will be distributed between the pots for the distribution among the group members according to the observer's investment in the verification:

- If she does not invest in the verification mechanism, all earnings from the team project go to the *equal-distribution-pot*.
- If she invests her complete endowment in the verification mechanism, all earnings from the team project will go to the *according-to-contribution-pot*.
- In all other cases, the earnings from the group project will be *divided between the pots* according to the investments in the verification mechanism. For example, if she invests 2 tokens in the verification mechanism $2/5 = 40\%$ of the payoff will go to the *according-to-contribution-pot* and the remaining 60% to the *equal-distribution-pot*.

You can gather more detailed payoffs from the attached Table 3 (Phase 3).

Team Members

Team members will be informed about the investment of the observer. The observers' investment in verification determines the division of the payoffs from the project. Each team member decides how much of his or her 10 tokens endowment to invest in the group project. Contributions lead to the same costs as in Phase 1.

Observer

The observer decides in each period of Phase 3 how many of his tokens she will invest in the verification mechanism of her group. The observer receives an endowment of 5 tokens. She is allowed to spend any amount between 0 and 5. The observer will keep tokens she does not invest. The observer will receive a share of 20% of the group project payoff.

Information at the End of the Period

At the end of each period each member and the observer will be informed about the sum of contributions. The members additionally will be informed about the composition of their payoff. The observer additionally will be informed about the resulting payoff.

Quiz

1. If the observer invests 3 tokens in the verification mechanism, a) which percentage share will a group member receive from the *according-to-contribution-pot* and b) which share from the *equal-distribution-pot*?

2. Suppose the observer invests 5 tokens in the verification mechanism. What is the payoff of a group member if the others invest on average 5 tokens and she invests 10 tokens? What if a group member invests 5 tokens?

Phase 4–6: Vote on the Verification System

At the beginning of Phases 4, 5, and 6, group members vote for one of two possibilities:

- Verification of the contributions by the group members (like Phase 2).
- Verification of the contributions by the observer (like Phase 3).

For the following 10 periods the verification system will be the one chosen by the majority of group members.

With other words: The group members decide whether they want either the system of Phase 2 (investment in verification by group members) or the system of Phase 3 (investment in verification by observer) for 10 consecutive periods. This will be repeated twice. The observer cannot vote.

Payoff tables

General Explanation:

The entries in the cells show your payoff conditional on your contribution and the average contribution of the others.

The rows represent your contribution choice while the column represents the average contribution of the others.

The color indicates the size of the payoff for each column: blue represents low payoff and red stands for a high payoff.

Numbers that appear in bold typeface are for cases in which your contribution is the same as the average contribution of the others in your group.

Phase I: Group project

	Average contribution of the others										
	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
My Contribution	**My payoff**										
0	**10.00**	12.40	14.80	17.20	19.60	22.00	24.40	26.80	29.20	31.60	34.00
1	10.44	**12.84**	15.24	17.64	20.04	22.44	24.84	27.24	29.64	32.04	34.44
2	10.57	12.97	**15.37**	17.77	20.17	22.57	24.97	27.37	29.77	32.17	34.57
3	10.39	12.79	15.19	**17.59**	19.99	22.39	24.79	27.19	29.59	31.99	34.39
4	9.89	12.29	14.69	17.09	**19.49**	21.89	24.29	26.69	29.09	31.49	33.89
5	9.08	11.48	13.88	16.28	18.68	**21.08**	23.48	25.88	28.28	30.68	33.08
6	7.95	10.35	12.75	15.15	17.55	19.95	**22.35**	24.75	27.15	29.55	31.95
7	6.52	8.92	11.32	13.72	16.12	18.52	20.92	**23.32**	25.72	28.12	30.52
8	4.76	7.16	9.56	11.96	14.36	16.76	19.16	21.56	**23.96**	26.36	28.76
9	2.70	5.10	7.50	9.90	12.30	14.70	17.10	19.50	21.90	**24.30**	26.70
10	0.32	2.72	5.12	7.52	9.92	12.32	14.72	17.12	19.52	21.92	**24.32**

Figure 3.2: Payoff table for the public-good phase.

Phase II: Investment in verification by group members

Note! If you invest in verification, costs of 3,5 tokens will additionally be deducted

Members investing in verification	My Contribution	Average contribution of the others: 0	1	2	3	4	5	6	7	8	9	10
		My payoff										
0	0	**10.00**	12.40	14.80	17.20	19.60	22.00	24.40	26.80	29.20	31.60	34.00
	1	10.44	**12.84**	15.24	17.64	20.04	22.44	24.84	27.24	29.64	32.04	34.44
	2	10.57	12.97	**15.37**	17.77	20.17	22.57	24.97	27.37	29.77	32.17	34.57
	3	10.39	12.79	15.19	**17.59**	19.99	22.39	24.79	27.19	29.59	31.99	34.39
	4	9.89	12.29	14.69	17.09	**19.49**	21.89	24.29	26.69	29.09	31.49	33.89
	5	9.08	11.48	13.88	16.28	18.68	**21.08**	23.48	25.88	28.28	30.68	33.08
	6	7.95	10.35	12.75	15.15	17.55	19.95	**22.35**	24.75	27.15	29.55	31.95
	7	6.52	8.92	11.32	13.72	16.12	18.52	20.92	**23.32**	25.72	28.12	30.52
	8	4.76	7.16	9.56	11.96	14.36	16.76	19.16	21.56	**23.96**	26.36	28.76
	9	2.70	5.10	7.50	9.90	12.30	14.70	17.10	19.50	21.90	**24.30**	26.70
	10	0.32	2.72	5.12	7.52	9.92	12.32	14.72	17.12	19.52	21.92	**24.32**
1	0	**10.00**	11.92	13.84	15.76	17.68	19.60	21.52	23.44	25.36	27.28	29.20
	1	10.92	**12.84**	14.76	16.68	18.60	20.52	22.44	24.36	26.28	28.20	30.12
	2	11.53	13.45	**15.37**	17.29	19.21	21.13	23.05	24.97	26.89	28.81	30.73
	3	11.83	13.75	15.67	**17.59**	19.51	21.43	23.35	25.27	27.19	29.11	31.03
	4	11.81	13.73	15.65	17.57	**19.49**	21.41	23.33	25.25	27.17	29.09	31.01
	5	11.48	13.40	15.32	17.24	19.16	**21.08**	23.00	24.92	26.84	28.76	30.68
	6	10.83	12.75	14.67	16.59	18.51	20.43	**22.35**	24.27	26.19	28.11	30.03
	7	9.88	11.80	13.72	15.64	17.56	19.48	21.40	**23.32**	25.24	27.16	29.08
	8	8.60	10.52	12.44	14.36	16.28	18.20	20.12	22.04	**23.96**	25.88	27.80
	9	7.02	8.94	10.86	12.78	14.70	16.62	18.54	20.46	22.38	**24.30**	26.22
	10	5.12	7.04	8.96	10.88	12.80	14.72	16.64	18.56	20.48	22.40	**24.32**
2	0	**10.00**	11.44	12.88	14.32	15.76	17.20	18.64	20.08	21.52	22.96	24.40
	1	11.40	**12.84**	14.28	15.72	17.16	18.60	20.04	21.48	22.92	24.36	25.80
	2	12.49	13.93	**15.37**	16.81	18.25	19.69	21.13	22.57	24.01	25.45	26.89
	3	13.27	14.71	16.15	**17.59**	19.03	20.47	21.91	23.35	24.79	26.23	27.67
	4	13.73	15.17	16.61	18.05	**19.49**	20.93	22.37	23.81	25.25	26.69	28.13
	5	13.88	15.32	16.76	18.20	19.64	**21.08**	22.52	23.96	25.40	26.84	28.28
	6	13.71	15.15	16.59	18.03	19.47	20.91	**22.35**	23.79	25.23	26.67	28.11
	7	13.24	14.68	16.12	17.56	19.00	20.44	21.88	**23.32**	24.76	26.20	27.64
	8	12.44	13.88	15.32	16.76	18.20	19.64	21.08	22.52	**23.96**	25.40	26.84
	9	11.34	12.78	14.22	15.66	17.10	18.54	19.98	21.42	22.86	**24.30**	25.74
	10	9.92	11.36	12.80	14.24	15.68	17.12	18.56	20.00	21.44	22.88	**24.32**
3	0	**10.00**	10.96	11.92	12.88	13.84	14.80	15.76	16.72	17.68	18.64	19.60
	1	11.88	**12.84**	13.80	14.76	15.72	16.68	17.64	18.60	19.56	20.52	21.48
	2	13.45	14.41	**15.37**	16.33	17.29	18.25	19.21	20.17	21.13	22.09	23.05
	3	14.71	15.67	16.63	**17.59**	18.55	19.51	20.47	21.43	22.39	23.35	24.31
	4	15.65	16.61	17.57	18.53	**19.49**	20.45	21.41	22.37	23.33	24.29	25.25
	5	16.28	17.24	18.20	19.16	20.12	**21.08**	22.04	23.00	23.96	24.92	25.88
	6	16.59	17.55	18.51	19.47	20.43	21.39	**22.35**	23.31	24.27	25.23	26.19
	7	16.60	17.56	18.52	19.48	20.44	21.40	22.36	**23.32**	24.28	25.24	26.20
	8	16.28	17.24	18.20	19.16	20.12	21.08	22.04	23.00	**23.96**	24.92	25.88
	9	15.66	16.62	17.58	18.54	19.50	20.46	21.42	22.38	23.34	**24.30**	25.26
	10	14.72	15.68	16.64	17.60	18.56	19.52	20.48	21.44	22.40	23.36	**24.32**
4	0	**10.00**	10.48	10.96	11.44	11.92	12.40	12.88	13.36	13.84	14.32	14.80
	1	12.36	**12.84**	13.32	13.80	14.28	14.76	15.24	15.72	16.20	16.68	17.16
	2	14.41	14.89	**15.37**	15.85	16.33	16.81	17.29	17.77	18.25	18.73	19.21
	3	16.15	16.63	17.11	**17.59**	18.07	18.55	19.03	19.51	19.99	20.47	20.95
	4	17.57	18.05	18.53	19.01	**19.49**	19.97	20.45	20.93	21.41	21.89	22.37
	5	18.68	19.16	19.64	20.12	20.60	**21.08**	21.56	22.04	22.52	23.00	23.48
	6	19.47	19.95	20.43	20.91	21.39	21.87	**22.35**	22.83	23.31	23.79	24.27
	7	19.96	20.44	20.92	21.40	21.88	22.36	22.84	**23.32**	23.80	24.28	24.76
	8	20.12	20.60	21.08	21.56	22.04	22.52	23.00	23.48	**23.96**	24.44	24.92
	9	19.98	20.46	20.94	21.42	21.90	22.38	22.86	23.34	23.82	**24.30**	24.78
	10	19.52	20.00	20.48	20.96	21.44	21.92	22.40	22.88	23.36	23.84	**24.32**
5	0	**10.00**	10.00	10.00	10.00	10.00	10.00	10.00	10.00	10.00	10.00	10.00
	1	12.84	**12.84**	12.84	12.84	12.84	12.84	12.84	12.84	12.84	12.84	12.84
	2	15.37	15.37	**15.37**	15.37	15.37	15.37	15.37	15.37	15.37	15.37	15.37
	3	17.59	17.59	17.59	**17.59**	17.59	17.59	17.59	17.59	17.59	17.59	17.59
	4	19.49	19.49	19.49	19.49	**19.49**	19.49	19.49	19.49	19.49	19.49	19.49
	5	21.08	21.08	21.08	21.08	21.08	**21.08**	21.08	21.08	21.08	21.08	21.08
	6	22.35	22.35	22.35	22.35	22.35	22.35	**22.35**	22.35	22.35	22.35	22.35
	7	23.32	23.32	23.32	23.32	23.32	23.32	23.32	**23.32**	23.32	23.32	23.32
	8	23.96	23.96	23.96	23.96	23.96	23.96	23.96	23.96	**23.96**	23.96	23.96
	9	24.30	24.30	24.30	24.30	24.30	24.30	24.30	24.30	24.30	**24.30**	24.30
	10	24.32	24.32	24.32	24.32	24.32	24.32	24.32	24.32	24.32	24.32	**24.32**

Figure 3.3: Payoff table for the peer monitoring phase.

Phase III: Investment in verification by observer

Observer investment in verification	My Contribution	0	1	2	3	4	5	6	7	8	9	10
		My payoff (Average contribution of the others)										
0	0	**10.00**	11.92	13.84	15.76	17.68	19.60	21.52	23.44	25.36	27.28	29.20
	1	10.32	**12.24**	14.16	16.08	18.00	19.92	21.84	23.76	25.68	27.60	29.52
	2	10.33	12.25	**14.17**	16.09	18.01	19.93	21.85	23.77	25.69	27.61	29.53
	3	10.03	11.95	13.87	**15.79**	17.71	19.63	21.55	23.47	25.39	27.31	29.23
	4	9.41	11.33	13.25	15.17	**17.09**	19.01	20.93	22.85	24.77	26.69	28.61
	5	8.48	10.40	12.32	14.24	16.16	**18.08**	20.00	21.92	23.84	25.76	27.68
	6	7.23	9.15	11.07	12.99	14.91	16.83	**18.75**	20.67	22.59	24.51	26.43
	7	5.68	7.60	9.52	11.44	13.36	15.28	17.20	**19.12**	21.04	22.96	24.88
	8	3.80	5.72	7.64	9.56	11.48	13.40	15.32	17.24	**19.16**	21.08	23.00
	9	1.62	3.54	5.46	7.38	9.30	11.22	13.14	15.06	16.98	**18.90**	20.82
	10	-0.88	1.04	2.96	4.88	6.80	8.72	10.64	12.56	14.48	16.40	**18.32**
1	0	**10.00**	11.54	13.07	14.61	16.14	17.68	19.22	20.75	22.29	23.82	25.36
	1	10.71	**12.24**	13.78	15.32	16.85	18.39	19.92	21.46	23.00	24.53	26.07
	2	11.10	12.64	**14.17**	15.71	17.24	18.78	20.32	21.85	23.39	24.92	26.46
	3	11.18	12.72	14.25	**15.79**	17.32	18.86	20.40	21.93	23.47	25.00	26.54
	4	10.95	12.48	14.02	15.55	**17.09**	18.63	20.16	21.70	23.23	24.77	26.31
	5	10.40	11.94	13.47	15.01	16.54	**18.08**	19.62	21.15	22.69	24.22	25.76
	6	9.54	11.07	12.61	14.15	15.68	17.22	**18.75**	20.29	21.83	23.36	24.90
	7	8.36	9.90	11.44	12.97	14.51	16.04	17.58	**19.12**	20.65	22.19	23.72
	8	6.88	8.41	9.95	11.48	13.02	14.56	16.09	17.63	**19.16**	20.70	22.24
	9	5.07	6.61	8.15	9.68	11.22	12.75	14.29	15.83	17.36	**18.90**	20.43
	10	2.96	4.49	6.03	7.57	9.10	10.64	12.17	13.71	15.25	16.78	**18.32**
2	0	**10.00**	11.15	12.30	13.46	14.61	15.76	16.91	18.06	19.22	20.37	21.52
	1	11.09	**12.24**	13.40	14.55	15.70	16.85	18.00	19.16	20.31	21.46	22.61
	2	11.87	13.02	**14.17**	15.32	16.48	17.63	18.78	19.93	21.08	22.24	23.39
	3	12.33	13.48	14.64	**15.79**	16.94	18.09	19.24	20.40	21.55	22.70	23.85
	4	12.48	13.63	14.79	15.94	**17.09**	18.24	19.39	20.55	21.70	22.85	24.00
	5	12.32	13.47	14.62	15.78	16.93	**18.08**	19.23	20.38	21.54	22.69	23.84
	6	11.84	12.99	14.15	15.30	16.45	17.60	**18.75**	19.91	21.06	22.21	23.36
	7	11.05	12.20	13.36	14.51	15.66	16.81	17.96	**19.12**	20.27	21.42	22.57
	8	9.95	11.10	12.25	13.40	14.56	15.71	16.86	18.01	**19.16**	20.32	21.47
	9	8.53	9.68	10.83	11.99	13.14	14.29	15.44	16.59	17.75	**18.90**	20.05
	10	6.80	7.95	9.10	10.25	11.41	12.56	13.71	14.86	16.01	17.17	**18.32**
3	0	**10.00**	10.77	11.54	12.30	13.07	13.84	14.61	15.38	16.14	16.91	17.68
	1	11.48	**12.24**	13.01	13.78	14.55	15.32	16.08	16.85	17.62	18.39	19.16
	2	12.64	13.40	**14.17**	14.94	15.71	16.48	17.24	18.01	18.78	19.55	20.32
	3	13.48	14.25	15.02	**15.79**	16.56	17.32	18.09	18.86	19.63	20.40	21.16
	4	14.02	14.79	15.55	16.32	**17.09**	17.86	18.63	19.39	20.16	20.93	21.70
	5	14.24	15.01	15.78	16.54	17.31	**18.08**	18.85	19.62	20.38	21.15	21.92
	6	14.15	14.91	15.68	16.45	17.22	17.99	**18.75**	19.52	20.29	21.06	21.83
	7	13.74	14.51	15.28	16.04	16.81	17.58	18.35	**19.12**	19.88	20.65	21.42
	8	13.02	13.79	14.56	15.32	16.09	16.86	17.63	18.40	**19.16**	19.93	20.70
	9	11.99	12.75	13.52	14.29	15.06	15.83	16.59	17.36	18.13	**18.90**	19.67
	10	10.64	11.41	12.17	12.94	13.71	14.48	15.25	16.01	16.78	17.55	**18.32**
4	0	**10.00**	10.38	10.77	11.15	11.54	11.92	12.30	12.69	13.07	13.46	13.84
	1	11.86	**12.24**	12.63	13.01	13.40	13.78	14.16	14.55	14.93	15.32	15.70
	2	13.40	13.79	**14.17**	14.56	14.94	15.32	15.71	16.09	16.48	16.86	17.24
	3	14.64	15.02	15.40	**15.79**	16.17	16.56	16.94	17.32	17.71	18.09	18.48
	4	15.55	15.94	16.32	16.71	**17.09**	17.47	17.86	18.24	18.63	19.01	19.39
	5	16.16	16.54	16.93	17.31	17.70	**18.08**	18.46	18.85	19.23	19.62	20.00
	6	16.45	16.83	17.22	17.60	17.99	18.37	**18.75**	19.14	19.52	19.91	20.29
	7	16.43	16.81	17.20	17.58	17.96	18.35	18.73	**19.12**	19.50	19.88	20.27
	8	16.09	16.48	16.86	17.24	17.63	18.01	18.40	18.78	**19.16**	19.55	19.93
	9	15.44	15.83	16.21	16.59	16.98	17.36	17.75	18.13	18.51	**18.90**	19.28
	10	14.48	14.86	15.25	15.63	16.01	16.40	16.78	17.17	17.55	17.93	**18.32**
5	0	**10.00**	10.00	10.00	10.00	10.00	10.00	10.00	10.00	10.00	10.00	10.00
	1	12.24	**12.24**	12.24	12.24	12.24	12.24	12.24	12.24	12.24	12.24	12.24
	2	14.17	14.17	**14.17**	14.17	14.17	14.17	14.17	14.17	14.17	14.17	14.17
	3	15.79	15.79	15.79	**15.79**	15.79	15.79	15.79	15.79	15.79	15.79	15.79
	4	17.09	17.09	17.09	17.09	**17.09**	17.09	17.09	17.09	17.09	17.09	17.09
	5	18.08	18.08	18.08	18.08	18.08	**18.08**	18.08	18.08	18.08	18.08	18.08
	6	18.75	18.75	18.75	18.75	18.75	18.75	**18.75**	18.75	18.75	18.75	18.75
	7	19.12	19.12	19.12	19.12	19.12	19.12	19.12	**19.12**	19.12	19.12	19.12
	8	19.16	19.16	19.16	19.16	19.16	19.16	19.16	19.16	**19.16**	19.16	19.16
	9	18.90	18.90	18.90	18.90	18.90	18.90	18.90	18.90	18.90	**18.90**	18.90
	10	18.32	18.32	18.32	18.32	18.32	18.32	18.32	18.32	18.32	18.32	**18.32**

Note! The share for the observer is already deducted.

Figure 3.4: Payoff table for the observer monitoring phase.

4 Limited Punishment[1]

Abstract

In public-good experiments with a punishment option, "anti-social" sanctioning is common: high contributors can get punished. We measure the effect of this type of punishment by introducing a punishment constraint in an endogenous public-good experiment and by comparing the results with the data of Gürerk et al. (2010a) without the constraint. The results show that "anti-social" punishment has social costs that go beyond the direct costs caused for the punishing as well as the punished players. The sum of the distributed punishment token is even lower than expected. Uncontrolled sanctioning thus leads to a substantial efficiency loss.

4.1 Introduction

Public-good games are one of the main tools in the cooperation research of experimental economics because they model in a mathematically simple way the conflict of own versus joined payoff maximization. This conflict leads to the break down of cooperation in the repeated experimental play if no counter-measures are taken such as revealing the identity of the players or allowing sanctions.

However, the public-good experiments with punishment option have an efficiency issue. Experiments with decentralized[2] sanction possibilities are prone to "anti-social" punishments. Like Herrmann et al. (2008) we consider sanctions of pro-social cooperators as "anti-social". We chose the micro-perspective for the definition of "anti-social". With that we mean that no one should punish someone who is behaving better than oneself.

The "anti-social" sanctions are costly for the society because firstly, especially if aimed at maximum contributors – they directly burn welfare

[1]This part is not published yet.

[2]With decentralized we mean that every group member is allowed to sanction any other group member.

and do not help in increasing the overall cooperation level; secondly, they cause more costs if the penalized contributors increase their punishment level when they got sanctioned; thirdly, if the punished contributors would react with lowering contributions they could reduce the welfare even further. And in fact most "anti-social" contributions are aimed at the maximum contributors, as can be seen in the plot of the punishment acts of the data of Gürerk et al. (2010a) Figure 4.5 (Appendix, page 101). In their data we found 33.2% "anti-social" sanction occurrences in the punishment institution, 29.4% punishments of more and equal contributing subjects in the data of the *punvcm*[3] treatment of Nikiforakis (2008) and even 37.9% in *punvcm* of Fehr and Gächter (2000a).

One cause of those "anti-social" punishments that the subjects of Fehr and Gächter (2000a) stated in a questionnaire is vengeance. Vengeance can be defined as "[...] an act designed to harm someone else, or a social group, in response to feeling that oneself has been harmed by that person or group" (Frijda, 1994 as cited in Gollwitzer, 2004). Some low contributing subjects did not like being sanctioned by others since they presumed correctly they were sanctioned by the highest contributors, they punished those to retaliate. Another reason could be "do-gooders derogation" (Monin, 2007): the dislike of people that behave better in moral aspects. Hopfensitz and Reuben (2009) point at the large emotional component in the sanctioning behavior of the subjects.

In order to improve the efficiency of the punishment design it has been proven successful to give people the choice whether they want to join a group with punishment option or remain in a group without that option (Gürerk et al., 2006, 2010a). But although this endogenous-institution-choice design proves successful in the course of the experiment and leads to almost full cooperation with high efficiency, the punishment group is highly unpopular in the beginning of the experiment with less than a third of the subjects choosing that option. One cause for this could have been that the subjects fear "anti-social" sanctions, another could have been that they anticipated the very low efficiency that was observed during the initialization phase.

In this paper we want to explore two things: First, whether the low popularity of the punishment institution in Gürerk et al. (2010a) could be attributed to the "anti-social" punishment and second, the effects on the efficiency of the punishment institution by mitigating them in this "voting-by-feet" experiment. We prevent "anti-social" sanctions mechanically by definition in so far as we simply do not allow punishments of higher and

[3]*punvcm* means that the subjects first experienced a design where sanctions where allowed and afterwards a sanction-less public-good game was played while with *vcmpun* for example it was the other way around.

equal contributors. We compare the data with the punishment constraint with the data of Gürerk et al. (2010a).[4]

The results of our experiment show that there is no difference in the subjects' choice of the institution, the punishment institution still is unpopular. The second result of the experiment is that we observe that the limit on the punishment indeed makes the sanctioning institution much more efficient. The sanction point revenue (the sum of all sanction points) is far below that of Gürerk et al. (2010a) even if all "anti-social" sanctions from their study are calculatory not taken into account. Hence, we could reduce revenge activities. As a consequence, the higher efficiency leads to a faster migration in the course of the experiment into the punishment institution. The higher efficiency of such an institution is an argument when thinking about the evolutionary competitiveness of a society. Speculatively this could have been a reason of the evolvement of some religious rules like "[...], avenge not yourselves, [...]" (Romans 12:19).

In the next section we will give an overview on the experimental literature that deals with cooperation in general and with punishment in particular. Thereafter we will describe the design of the experiment, provide the research hypothesis before we explain the results.

4.2 Related Literature

Public-good experiments are the work horse of experimental research on cooperation. The first public-good experiments have been conducted more than 30 years ago, early examples are Smith (1979) and Isaac et al. (1984). A survey of the older experiments can be found in Ledyard (1995). Gächter and Herrmann (2009) provide an excellent and very recent overview of public-good and related experiments and the determinants of cooperation, as does Chaudhuri (2011). We will give a shorter and more selective overview here.

In standard public-good experiments with anonymity of the subjects the cooperation breaks down due to several reasons. One explanation is that people are conditionally cooperative in the sense that they cooperate only if others cooperate as well (Fischbacher et al., 2001; Gächter, 2007; Fischbacher and Gächter, 2010). People often are imperfect conditional cooperators meaning that they cooperate slightly less than others. As Fischbacher and Gächter (2010) show this consequently leads to the decrease of cooperation. This behavior is common even in cross-cultural comparison (Herrmann et al., 2008).

[4] We thank the authors hereby for the permission to use their data.

One successful way of improving the cooperation is lifting the usual anonymity constraint or allowing communication between the participants (Isaac and Walker, 1988b; Brosig et al., 2003; Bochet et al., 2005). This could be related to reputation building (Milinski et al., 2006; Rockenbach and Milinski, 2006), which has also been proven to be effective.

The provision of sanctioning possibilities has been shown to be another successful measure to alleviate the free-riding problem and to enhance the cooperation (Ostrom et al., 1992; Fehr and Gächter, 2000a, 2002). This has been confirmed in many further experiments and modifications and even across cultures.[5] Although it seems harsh to say that human beings appear to need a certain threat and disciplinary sanctions, one has to keep in mind that this would not necessarily mean that people need punishment in the sense of causing a damage. Instead, it could mean some non-monetary costs are imposed so that the sanction could have a signaling-only character. Even symbolic punishment with no cost at all seems to improve the cooperation level to a certain extent (Masclet et al., 2003; Noussair and Tucker, 2005). The reason probably is that the sanctions induce negative feelings like guilt (Hopfensitz and Reuben, 2009). Interestingly, rewards for cooperators seem to be much less efficient than punishment of non-cooperators (Gürerk et al., 2006; Sefton et al., 2007; Sutter et al., 2008).

Since some subjects in experiments are willing to pay for disciplinary purposes with no direct (monetary) benefit, it has been dubbed "altruistic" punishment (Fehr and Gächter, 2000a). One explanation for the altruistic sanctions is offered by inequity-aversion models like those of Fehr and Schmidt (1999) and Bolton and Ockenfels (2000) that assume that people have a certain dislike of inequality. In that case they might use sanctions to equalize payoffs. Furthermore, it seems that humans experience some sort of a positive emotion if they punish unfair behavior (de Quervain et al., 2004; Hopfensitz and Reuben, 2009). People respond to unfair behavior even if it costs payoff as well as when there is no opportunity to build up a reputation as for example in experiments with just a single period (Gächter and Herrmann, 2009).

The punishment should not be too expensive. The higher the leverage of the punishment (the lower the costs thereof), the better usually is the enhancing effect on the contribution (Anderson and Putterman, 2006; Carpenter, 2007; Nikiforakis and Normann, 2008; Egas and Riedl, 2008; Masclet and Villeval, 2008).

As we already mentioned in the introduction, one shortcoming of punishment in general is that it can be costly for society. A potential con-

[5]For example: Masclet et al. (2003); Carpenter et al. (2004); Page et al. (2005); Masclet and Villeval (2008); Herrmann et al. (2008).

sequence is that efficiency is not higher than in a comparable public-good experiment without punishment (Fehr and Gächter, 2000a; Herrmann et al., 2008; Masclet and Villeval, 2008). In Herrmann et al. (2008), for example, the efficiency of the non-punishment treatment yielded a higher average payoff in most subject pools. Of course, the effect can depend on the length of the interaction. Gächter et al. (2008) showed that extending the length of the experiment from the usual 10 up to 50 periods yielded much better comparative efficiency of the punishment option.

It also has been proven highly successful to give people the choice whether to join (Gürerk et al., 2006) or to vote for (Sutter et al., 2008; Ertan et al., 2009) either a specific punishment institution or a non-punishment institution. Though the punishment institution in Gürerk et al. (2006) is chosen only by a minority of subjects in the beginning and there is high punishment leading to low payoffs, the endogenous punishment institution quickly proves highly efficient and almost all subjects join and cooperate in the end.

One widespread phenomenon throughout these public-good experiments with punishment option has been the punishment of high-contributors. This was noted in early experiments like Fehr and Gächter (2000a) and has become subject to more recent research (Cinyabuguma et al., 2006; Denant-Boemont et al., 2007; Nikiforakis, 2008; Herrmann et al., 2008; Ertan et al., 2009). The behavior has been dubbed "anti-social" (Nikiforakis, 2008; Herrmann et al., 2008) or "perverse" (Cinyabuguma et al., 2006). There is some difference in the exact definition namely the reference point of "anti-social" acts – whether the benchmark is the average contribution level (Nikiforakis, 2008; Cinyabuguma et al., 2006) or the individual contribution level (Herrmann et al., 2008). For us the individual perspective is the reference point. No one will be able to punish someone who is behaving better, i.e., contributes more than oneself. The actual difference between the approaches is minimal since the majority of "anti-social" punishments is directed towards the maximum contributors.

How are these "anti-social" punishments motivated? Fehr and Gächter (2000a) categorize the responses of their questionnaire as follows: *spiteful revenge* which are those punishments that are in the very same period exercised against high contributors. This spiteful revenge is possibly related to a dislike of "do-gooders"(Monin, 2007), but also partially the punishment of non-conformist behavior (Carpenter and Matthews, 2005), especially if there are only few high contributors. Furthermore *Blind revenge* occurs according to Fehr and Gächter (2000a) if someone was punished a period before and takes revenge now. The subjects also stated *relative strategic advantage* as a reason – if they want to earn more relative to the others they can lower the income of others by sanctions. Furthermore there is also the possibility of *errors*. In cross-cultural comparison the extent of "anti-social" punishments can be related to sociological measures like norms of civic cooperation which appear to be negatively correlated with the "anti-social" punishments (Herrmann et al., 2008).

The so-called counter-punishment might be seen as one solution against "anti-social" punishment. Furthermore since punishment in itself represents a second-level public-good – everyone benefits if the punished subject increases the contribution but there is the incentive to shirk on the punishments as it is costly – another punishment stage is also a mechanism that possibly prevents this second-level free-riding. This could be set up in such a way as the subjects enter a second (or more) punishment stage(s) in order to be able to punish those high contributors that refrained from punishing (who free-ride on the sanctions of others) or the low contributors with their "anti-social" punishments (Nikiforakis, 2008; Denant-Boemont et al., 2007; Nikiforakis and Engelmann, 2010; Nicklisch and Wolff, 2011). However with the counter-punishment possibility the cooperation could break down (Nikiforakis, 2008) since the enforcing cooperators are threatened by revenge sanctions and refrain from sanctioning. On the other hand Nicklisch and Wolff (2011) report no apparent "anti-social" punishment in their first stage of their basic treatment. However, opening several punishment stages offers the possibility of retaliation and raises efficiency issues.

Another way to reduce the "anti-social" sanctions is letting people choose the punishment rule, as did Ertan et al. (2009). Their experiment is closely related to ours. They let the subjects vote on whether they want to implement a rule that does not allow to punish high contributors. Contrary to their approach we implemented a rule that does only allow to punish someone contributing less. Further, our voting is "by feet" meaning that the subjects are free to join an institution with punishment or an institution without. But if subjects opt for the sanctioning institution they have to live with the limit on punishment.

Decker et al. (2003) also evaluated different punishment rules. They compared different collective rules like choosing the punishment by majority with an individual rule. In their experiments the costs for punishment were divided equally while in most other papers the sanction costs are individually imposed. A voting experiment approach on different central sanction schemes was conducted by Putterman et al. (2010).

4.3 Experimental Design

The experiment has 30 periods each consisting of 3 stages: the institution-choice stage ($S0$), the voluntary-contribution stage ($S1$) and the punishment stage ($S2$).

In the *institution-choice* stage $S0$ the subjects individually can choose between the two mechanisms to join, either *pun* – a public-good game where punishment is allowed, or *free* which resembles a classical public-good game. So the number of subjects in an institution can differ from period to period.

In the *voluntary-contribution* stage $S1$, the participants receive an endowment of 20 tokens and have to decide on their "contribution to the team project", c_i. The payoff of the "team project" – the sum of contributions multiplied by the factor 1.6 – is divided equally among the members in the same mechanism.

In the third and last stage, the *punishment* stage $S2$, the subjects choosing the punishment institution receive an extra endowment of 20 tokens as punishment budget. A list with the individual contributions without revealing the identity[6] of the other group members in *pun* is given. The tokens not spent on sanctioning are kept and thus add to the periods' profit. A distributed punishment token causes the receiver a loss of 3 tokens. If the subject chooses the *free* mechanism they receive an additional profit of 20.

To sum up, the profit π_i of a subject i is:

$$\pi_i = \underbrace{20 - c_i + \frac{1.6}{n}\sum_{j=1}^{n} c_j}_{S1} + \underbrace{\begin{cases} 20, & n = 1 \text{ or } \textit{free} \\ 20 - 3T_{-i} - T_i, & n > 1 \text{ and } \textit{pun} \end{cases}}_{S2} \tag{4.1}$$

with n as the number of members in the same mechanism, T_i denoting the sum of allocated punishment tokens and T_{-i} representing sum of the received punishment tokens. Both are of course 0 when the subject chooses *free*.

Our "constraint" or "limit" means that in our punishment institution (*pun*) subject i was only allowed to punish subject j if $c_i > c_j$ in contrast to our comparison data of Gürerk et al. (2010a) where it was allowed to sanction any other *pun* institution member. Our limit should prevent "anti-social" punishments.

Our experiment had 8 groups with 12 members each, making a total of 96 subjects. The experiment took place at the Erfurt Laboratory for Experimental Economics (elab) using the z-tree software (Fischbacher, 2007). The subjects, students of the University of Erfurt, were recruited using the Orsee system (Greiner, 2004).

[6]The subjects are not identifiable by a tracking number. Only a list of the contributions is shown. The order of the list of contributions is randomized each period. The sanctions can only be assigned to the specific contribution. This is common knowledge for all subjects.

4.4 Theoretical Considerations and Research Hypotheses

4.4.1 Standard Nash Equilibrium

Traditional economic theory assumes all actors conform to the so-called *homo oeconomicus* model – a strict money-maximizer – and predicts only zero contributions in both institutions and treatments without any punishment: Since $\frac{\partial \pi_i}{\partial c_i} < 0$ – the return of a unit of contribution is smaller than one – no one will contribute. Punishment is costly and thus resembles a public-good problem itself. Thus, no one will punish. By using backward induction for the repeated game this represents the subgame-perfect Nash equilibrium: no punishment and no contribution. Because of the identical outcome in both institutions the *homo oeconomicus* would be indifferent in the choice of the institution. Furthermore there is no difference whether there is a limit or not.

4.4.2 Inequity Aversion

There exist several approaches to improve the traditional economic theory since people generally do not behave like a *homo oeconomicus* (see for example Henrich et al. (2006)). One alternative are inequity aversion models, like those of Bolton and Ockenfels (2000) and Fehr and Schmidt (1999). We will focus on the latter to discuss the predictions.

It has been shown by Fehr and Schmidt (1999) that there exist cooperative equilibria in a public-good with punishment possibility under certain circumstances. A similar analysis in the context of endogenous institutional choice has been conducted by Gürerk et al. (2010a). It is possible to predict the success of the sanctioning institution using that kind of analysis.

While this model can explain why cooperation might emerge in public-good experiments with punishment it fails to explain "anti-social" punishment. A low contributor punishing a high contributor would even make the difference in profits larger but $\alpha, \beta > 0$ by definition in Fehr and Schmidt (1999) (p. 822). In other words both parameters represent a dislike of inequality but a low contributor that punishes a high contributor increases the inequality even further. By definition there can only be a disutility by the sanction and by the increased inequality. It would be different if β were allowed to be negative, in this case there would be a positive utility of higher inequity. But since this is not the case, the limit on the sanctions as in our design should not make a difference.

Relaxing the assumption of perfect information could probably explain some of the "anti-social" sanctions imposed on the people who contribute equally high: If there is no knowledge about who punishes at what level then we have a coordination problem among the enforcers that likely leads to low contributors being "over-punished". Those might be inclined after the contribution adjustment to reduce the resulting inequality. However this should be very unlikely since they would not free-ride if they cared much for inequality.

4.4.3 Alternatives

In order to predict the effect of the constraint on the "anti-social" punishment we need an approach that forecasts the occurrence of this punishment. Spite – subjects that feel good doing the bad – would be one model that does so (Saijo and Nakamura, 1995; Fehr et al., 2008). Another could be intention based theories like that of Dufwenberg and Kirchsteiger (2000) that assume a reciprocal motivation – people being kind if they are treated kindly and hostile if treated unfriendly. An interesting extension is the model of Falk and Fischbacher (2006). They combine inequity aversion with intentions: The subjects do not only evaluate the distribution but also the intention behind the action of the others. If the intention is taken into account it is possible that money-maximizers perceive their received sanctions as an unfriendly act and reciprocate accordingly. Even with the intention based models it is not easy to explain the sanctions of low contributors towards high contributors in the same and especially in the first period. Possibly it is necessary to extend the inequity models towards different behavior at least for some of the subjects.

As already mentioned in the introduction, there is also the work of Hopfensitz and Reuben (2009). Although it does not offer a technical model it reminds us that there is an emotional component that has to be taken into account. The emotions lead to even harder actions and possibly revenge. The "anti-social" sanctions could be attributed to negative emotions toward others, acting more morally than oneself. This was called "do-gooders derogation" by Monin (2007).

4.4.4 Hypotheses

As we know from Gächter and Herrmann (2009) punishment can be an effective measure to enhance cooperation and Gürerk et al. (2010a) did show that the endogenous institutional choice induces a higher efficiency mitigating the efficiency issue on punishment. We expect a similar success of the *pun* institution in our treatment with limit as well since our design is comparable. This is in line with the theoretical analysis assuming the Fehr and Schmidt (1999) model.

Hypothesis 1. We expect that the *pun* institution maintains a higher cooperation level as the *free* institution.

As we know from Herrmann et al. (2008) and as the sunflower plot of Figure 4.5 demonstrates for Gürerk et al. (2010a), "anti-social" sanctions are very present. Since we prevent those sanctions in our design by not allowing the distribution of punishment tokens toward higher or equal contributors, we expect that this should bring the number of sanctions down to at least the amount of Gürerk et al. (2010a) with a calculatory sanction cap (i.e., not counting the "anti-social" sanctions). Assuming that the subjects act reciprocal as with the theory of Falk and Fischbacher (2006) the number should be lower. The same holds if we assume that there is an emotional component in the motivation of the subject.

Hypothesis 2. If the distribution of sanction points is limited to subjects who contribute less than the punishing subject, the punishment level will not be higher than the hypothetical level of Gürerk et al. (2010a) with the "anti-social" punishments deducted.

High contributors that receive "anti-social" punishment could possibly react in two ways: either to lower their contributions to not expose themselves and/or to increase their punishment as well. In both cases the limit should introduce a higher efficiency of the punishment institution. Furthermore free-riders could be less likely to increase their contribution when they must fear being punished as high contributors as well.

Hypothesis 3. The punishment constraint leads compared to the data of Gürerk et al. (2010a) to a significantly higher efficiency in the sanctioning institution due to a) higher contributions and b) a lower punishment. This should lead to a higher attractiveness for the subjects and thus a faster migration towards the *pun* institution with the limit imposed.

In Gürerk et al. (2010a) only a third of the subjects joined the sanctioning institution in the first period. One possible explanation for this could be that the subjects fear being punished regardless of their contribution level. Furthermore they might expect counter-punishment: high contributors might be punished by free-riders either in the very same period or in the period after they had received sanctions. Since the identity of the subjects was not common knowledge they would use the contributions as an orientation and thus direct their own punishment towards higher contributors: a behavior that was observed in Gürerk et al. (2010a).

If the Hypotheses 2 and 3 are valid, the attractiveness of the punishment institution with the punishment constraint should be higher from the beginning: cooperative subjects that might have feared being punished even when they are contributing are now protected by the limit. Furthermore as we expect less punishment as well as higher contributions this could lead to more cooperative subjects willing to join the sanctioning institution.

Hypothesis 4. We expect that a greater share of subjects chooses the sanctioning institution in the treatment with the limited punishment from the first period on.

4.5 Results

4.5.1 Contributions

As in the endogenous punishment institution experiments of Gürerk et al. (2006) and Gürerk et al. (2010a) we observe that the contribution level in the institution without punishment (*free*) converges towards the Nash equilibrium. Furthermore we confirm an increasing level in the punishment institution (*pun*) to almost full contribution as can be seen in Figure 4.1. Thus we can validate Hypothesis 1:

Result 1. *The contribution level in the endogenously chosen punishment institution quickly emerges to almost 100% while it is deteriorating in the non-punishment institution.*

In contrast to our assumption and to the impression that could arise from Figure 4.1 there is no significant difference in the contributions between the treatment with the limit on punishment and the data of Gürerk et al. (2010a) without the limit: A per-period exact Wilcoxon–Mann–Whitney test (Appendix, Table 4.6) shows that there is not a single period with a significant difference. The same holds for the contribution level aggregated over all 30 periods ($p = 0.8$).

Result 2. *We must reject Hypothesis 3a: contributions are not significantly higher when there is a limit on the punishment.*

To give a first idea how the punishment activity affected the contributions we tabulated the specific reactions, i.e., the change in contribution after a punishment was received in Table 4.1. As we clearly see most subjects increased their contribution if they were punished. Interestingly, without the limit (Gürerk et al., 2010a) the number of above average contributors that decreased their contribution compared to those who increased their contribution is significantly larger than in the treatment with the limit on punishment (13 vs. 33 respective 1 vs. 18; $p = 0.04$, test for equality of proportions). Of those 13 occurrences of contribution

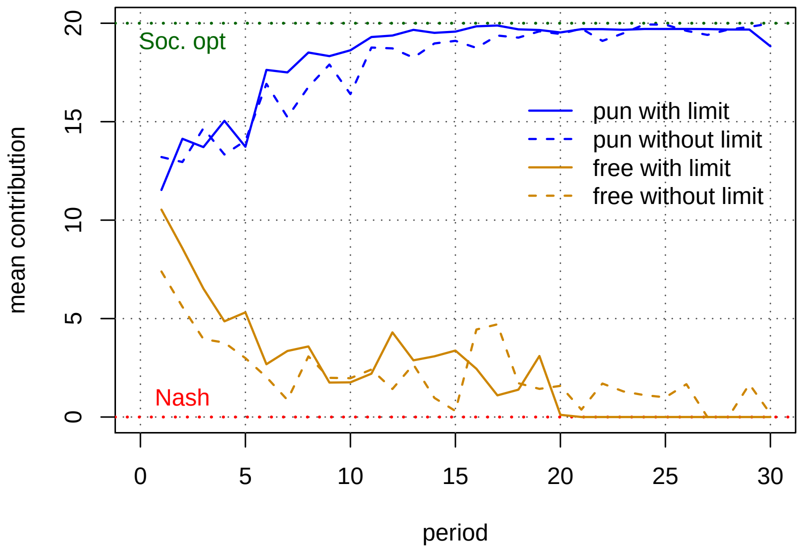

Figure 4.1: Average contributions. The data of Gürerk et al. (2010a) are denoted as "Without limit".

Table 4.1: Subjects' contribution reactions *if punished* in $t-1$. We separated the counts by below and above average contributors and since in the second half of the experiment most subjects in the punishment institution fully cooperated we excluded the cases where the punished subjects were the maximum contributors in $t-1$ in the lower half of the Table for comparative reasons. The data of Gürerk et al. (2010a) are denoted as "no limit".

	$c_{i,t-1} < \bar{c}_{t-1}$		$c_{i,t-1} \geq \bar{c}_{t-1}$		together	
contribution	no limit	limit	no limit	limit	no limit	limit
reduced	8	7	13	1	21	8
unchanged	9	22	217	7	226	29
increased	77	112	33	18	110	130
excluding if $c_{i,t-1} = \max(c_{t-1})$:						
reduced	8	7	1	1	9	8
unchanged	9	22	8	7	17	29
increased	77	112	30	18	107	130

decreases without the limit, 12 were maximum contributors as the comparison with the lower part of the table shows. This supports the Hypothesis 3a that "anti-social" punishment could cause high contributors to lower their contributions. But the impact of this is too small to have an effect on the overall contribution level as we know from Result 2.

Below-average contributors increased their contribution in 82% of the cases without the limit and in 79% with the limit employed. The generalized Cochran–Mantel–Haenszel test could not indicate any significant difference between the treatments ($p = 0.26$ for below, $p = 0.78$ for above average and $p = 0.45$ jointly, all maximum contributors excluded).

The Tobit estimates of Table 4.2 explain the contribution changes of the subjects in the punishment institution in $t - 1$ compared to t. We estimated positive and negative changes separately. The changes are explained by the period number (t and t^2), the positive or negative contribution difference to the average group contribution and those in the previous period received sanction points. The estimates support the assumption that the sanctions mainly increased the contributions. The coefficients for $recpun_{t-1}$ are both positive and significant. Interestingly, the punishment in the limited treatment had a greater effect: the coefficient is 0.4304 for each punishment point received while without limit the coefficient was only 0.1852 – in other words: one sanction point was more efficient with the limit causing a faster contribution adjustment of the punished subject. Beside this we find evidence for conditional cooperation since the subjects appear to adjust to the mean cooperation.

Result 3. *The sanctions with the limit imposed show a higher efficiency: One sanction point causes a faster contribution adjustment.*

Remarkable in this regard is that the subjects seem to clearly understand that punishment is a tool against free-riding. This is more than evident from the fact that the number of total free-riders with $c_i = 0$ in the punishment institution in the first period is either 0 without limit and 1 with the limit while there is a remarkable number in the other institution as can be seen in Table 4.3 which depicts the number of total free-riders versus non total free-riders and the counts of the above and below average contributors per treatment and institution in the first period.

4.5.2 Institution Choice

In the first period there is no significant difference in the institution choice ($p = 0.77$, exact Wilcoxon–Mann–Whitney test). The majority – about 70% in both treatments – of subjects chose the *pun* institution. Thus we must reject Hypothesis 4 at least partially which means that the subjects do not

Table 4.2: Robust Tobit estimates of contribution changes in the punishment world. We conducted a separate regression for each treatment and further ones for negative respective positive contribution changes. The data of Gürerk et al. (2010a) are denoted as "no pun limit".

increased contribution	**no pun limit**		**with pun limit**	
$(c_{i,t} \geq c_{i,t-1})$	Coef.	rob. s.e.	Coef.	rob. s.e.
const	-2.7920^{**}	1.0247	2.8958^{***}	0.9735
t	-0.2094	0.1450	-0.8114^{***}	0.1529
t^2	-0.0019	0.0038	0.0120^{***}	0.0044
$\bar{c}^{pun}_{t-1} - c_{i,t-1}$ if positive	1.2790^{***}	0.1889	0.3550^{*}	0.0291
$c_{i,t-1} - \bar{c}^{pun}_{t-1}$ if positive	0.0295	0.0268	-1.0108^{***}	0.3082
$recpun_{t-1}$	0.1852^{**}	0.0927	0.4304^{***}	0.1451
Wald χ^2	279.63	$p < 0.001$	128.04	$p < 0.0001$
decreased contribution	**no pun limit**		**with pun limit**	
$(c_{i,t} \leq c_{i,t-1})$	Coef.	rob. s.e.	Coef.	rob . s.e.
const	19.7648^{***}	4.2374	12.4437^{***}	2.9029
t	-0.4494	0.3865	0.4102	0.3477
t^2	0.0177^{*}	0.0107	-0.0078	0.0007
$\bar{c}^{pun}_{t-1} - c_{i,t-1}$ if positive	-1.0294	0.8050	3.2624	3.2557
$c_{i,t-1} - \bar{c}^{pun}_{t-1}$ if positive	-2.1790^{***}	0.4271	-2.8556^{***}	0.4621
$recpun_{t-1}$	-0.4057	0.2534	-2.0870	1.5842
Wald χ^2	62.02	$p < 0.0001$	54.91	$p < 0.0001$

Signif. codes: *** 1%, ** 5%, * 10%

Table 4.3: First period contribution behavior (counts): number of total free-riders with zero contribution vs. non free-riders and number of below average contributors vs. above average contributors. The data of Gürerk et al. (2010a) are denoted as "no limit". (total: 96 subjects per treatment)

Treatment	**Institution**	$c_i = 0$	$c_i \neq 0$	$c_i < \bar{c}_i$	$c_i \geq \bar{c}_i$
no limit	free	54	12	37	29
no limit	pun	30	0	14	16
limit	free	61	5	40	26
limit	pun	29	1	16	14

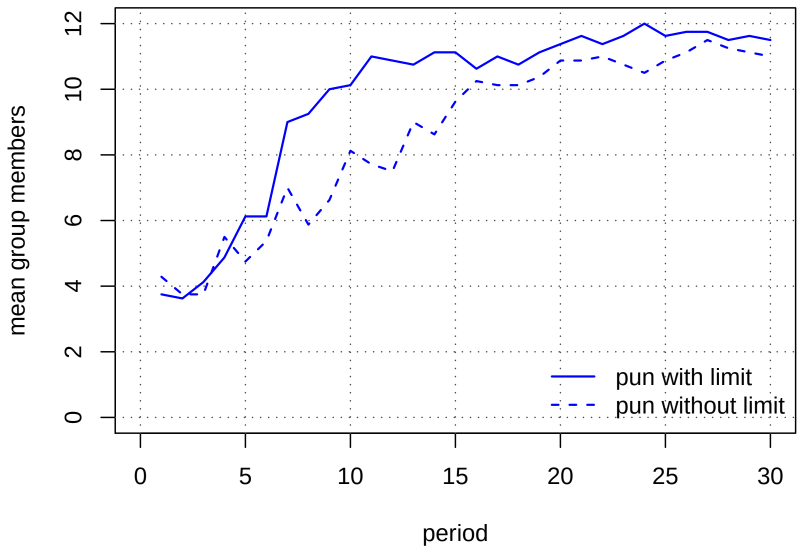

Figure 4.2: Average number of members in the punishment institution. Without limit is the data of Gürerk et al. (2010a).

anticipate a difference resulting from the design. People might thus have a general dislike of punishment and care less about the possible threat to receive unjustified punishment.

Nevertheless, there is a difference in the course of the experiment as Figure 4.2 shows. The dynamic from period 5 on turns remarkably towards the punishment institution our treatment with the punishment constraint compared to Gürerk et al. (2010a). Between period 5 and 15 we observe that in 6 of the 10 periods there are significantly more people on average in the punishment institution with the limit than without (Gürerk et al., 2010a) (see Table 4.6, Appendix, p. 100, for the tests). In the second half of the experiment the subjects' institution choice converges in both treatments as by then most subjects joint the punishment institution.

Result 4. *The punishment limit itself does not make the punishment institution more attractive in the first period: 70% of the subjects chose the free institution which is the same number as in Gürerk et al. (2010a). But with the limit they change the institution more likely in the course of the experiment.*

So what exactly makes then the punishment institution so much more appealing with the limit? One reason is very obvious: subjects earn more when there is the punishment limit imposed as Figure 4.3 reveals. Without the limit the profits of the first period were not only lower than in *free* but much lower than the Nash equilibrium. Opposite to that first-period profits in *pun* were much higher in presence of the limit. The profit differences between our data and Gürerk et al. (2010a) are significant at the 5%-level in 6 of the first 10 periods as the two-sided exact Wilcoxon–Mann–Whitney

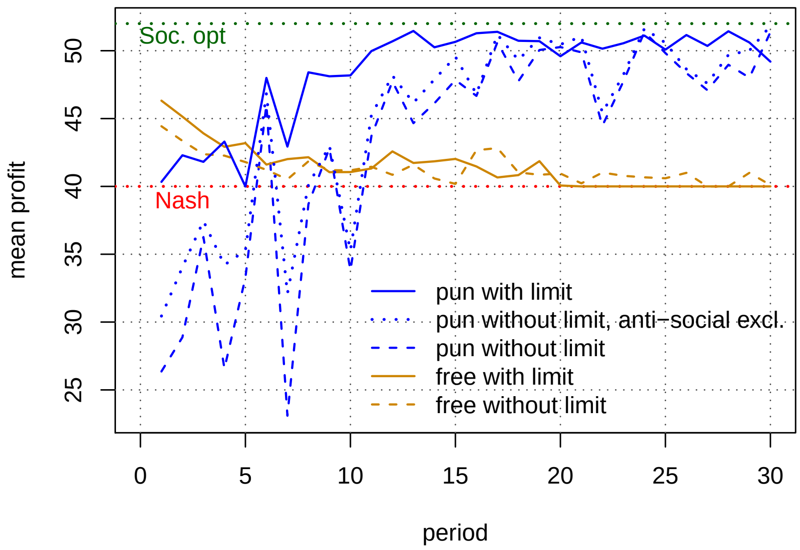

Figure 4.3: Average group profits. Without limit is the data of Gürerk et al. (2010a).

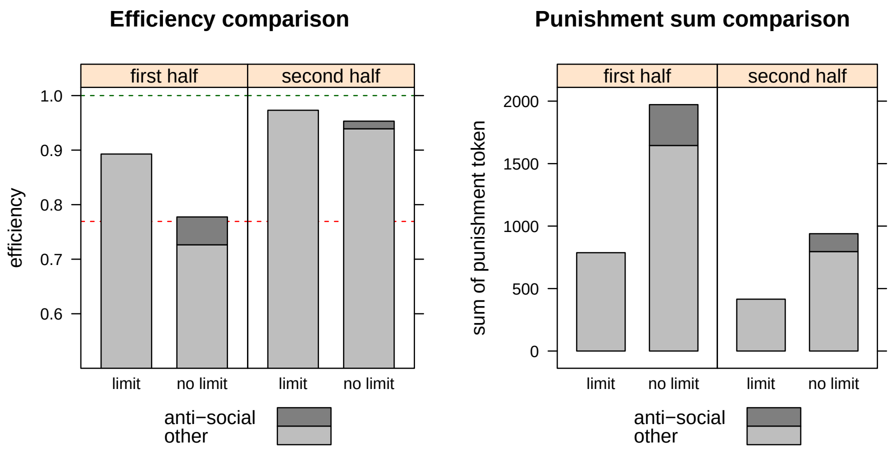

Figure 4.4: Efficiency and punishment revenue (sum of all punishment tokens) comparison of our data ("limit") with those of Gürerk et al. (2010a) ("no limit").

tests of Table 4.6 on page 100 (Appendix) show. Since contributions were not significantly different it is obvious that the difference is due to the punishment as Figure 4.4, right panel, confirms. It is also clear from the same figure that the sanction points distributed without limit are even lower compared to Gürerk et al. (2010a) if one deducts there the "anti-social" punishments. The dotted line in Figure 4.3 represents the profits if we simply subtract all "anti-social" punishment tokens. The difference between the treatments is still significant ($p = 0.0008$ for the first half of the experiment and $p = 0.0028$ for all periods, two-sided exact Wilcoxon–Mann–Whitney test).

Table 4.4: Random-effects logit estimates of the change in institution. The data of Gürerk et al. (2010a) are depicted as "no pun limit".

	no pun limit		**with pun limit**	
	Coef.	bootstr.s.e.	Coef.	bootstr.s.e.
$const.$	-1.0948^{**}	0.5061	0.1587	0.5045
t	0.0334^{*}	0.0472	-0.0030	0.0584
t^2	-0.0018	0.0015	-0.0008	0.0020
$\max(\pi_i - \overline{\pi}^{other}; 0)$	-0.0814^{***}	0.0254	-0.1946^{***}	0.0406
$\max(\overline{\pi}^{other} - \pi_i; 0)$	0.0388^{*}	0.0228	0.0478	0.0308
$\max(\pi_i - \overline{\pi}^{own}; 0)$	0.0376	0.0238	-0.1425^{***}	0.0711
$\max(\overline{\pi}^{own} - \pi_i; 0)$	0.0074	0.0264	0.0412	0.0610
$\max(c_i - \overline{c}^{other}; 0)$	-0.0924^{***}	0.0267	-0.0622^{***}	0.0263
$\max(\overline{c}^{other} - c_i; 0)$	0.0279	0.0254	0.0349	0.0360
$\max(c_i - \overline{c}^{own}; 0)$	0.0089	0.0477	-0.0387	0.0933
$\max(\overline{c}^{own} - c_i; 0)$	-0.0836	0.0641	0.0453	0.0912
$recpun_{t-1}$	0.0239	0.0736	0.1073	0.1722
n_{t-1}	-0.0011	0.0499	-0.0712^{*}	0.0387
Wald χ^2	169.20	$p < 0.0001$	188.83	$p < 0.0001$

Signif. codes: *** 1%, ** 5%, * 10%

Result 5. *Subjects in the experiment with the punishment limit had higher average profits in the punishment institution in the first periods compared to Gürerk et al. (2010a). This is due to a much lower sanction points revenue.*

To analyze in more detail what made the subjects change the institution we estimate random-effects logit models (Table 4.4) and use spine plots as an indication as well (Appendix, Figures 4.6, 4.7, 4.8). The regression and the spine plots support the hypothesis that the individuals look at the average profits but also at the other institution average contribution in relation to the own contribution.[7] If the individual's profit is higher than the average profit of the other group the subject is less likely to change the institution while it was more likely to change if the own profit was lower than the average. The same holds true for the contribution: Subjects who contribute more than the average of the other group have a significantly lower probability to move to the other institution. This is probably caused by the high contributors in *pun*. We furthermore observed that a punished subject is not more likely to go back to the *free* institution, they stay in *pun* and adjust their behavior.

4.5.3 Punishment Behavior

To analyze the punishment behavior more in detail we used a so-called hurdle regression model because we believe that a punishment decision can be split into two separate decisions: first whether the subject wants to

[7] They could see the summary statistics in advance of the institution decision.

punish at all and second at what size the sanction will be exercised. The first part of the regression is estimated with a logit model and the second with the negative binomial model. The latter is a so-called count model which can be used because the sanction points are discrete data and are approximately distributed like a count variable. We choose the negative binomial over the Poisson distribution by model selection criteria (Vuong test). Each treatment was estimated separately.

As explanatory variables we have taken into account: the received punishment points of the last period ($recpun_{t-1}$), the difference in the subjects' own and the other subjects' contribution (separated into a positive and negative distance), the period, whether the potentially punishing subject was a maximum- or minimum-contributor in the last period, a minimum contributor in the current period, whether the potentially punished "other" subject itself was a minimum and/or maximum contributor and finally the number of the group members n.

The estimates in Table 4.5 show that the punishment decision first of all appears to be a question about whether or not to punish and not so much at what extent. With the punishment limit introduced there is hardly anything that explains the size except for the weakly significant dummy variable of whether or not the punished subject was a minimum contributor. This is different without the punishment restriction (the data of Gürerk et al. (2010a)). Without the limit the sanction size is attributable mostly to the difference in contribution of the punished and the punishing subject ($\max(c_{i,t} - c_{j,t}; 0)$) as well as the dummy variable whether the punished subject was the subject that had the lowest contribution or not (*other min contr*$_t$).

There is evidence for revenge as a source of "anti-social" punishment since we observe in the data without limit (Gürerk et al., 2010a) that the probability that someone punishes is higher if someone became punished himself in the last period, as the positive and significant logit coefficient for $recpun_{t-1}$ shows. The same coefficient is not significant for the data with the punishment constraint. However, the received punishment with the limit imposed seems to have another effect as the minimum contributor of the last period (*i: min contr*$_{t-1}$), has a high probability to sanction herself, though there is only one direction she can sanction only subjects that contribute less. We already have stated with Result 3 that the punishment is more efficient with the constraint as it causes a faster adjustment.

One remarkable fact is that the size of the punishment is highly significant if the other subject is a maximum contributor. Interestingly, in the data of Gürerk et al. (2010a) only a few subjects were responsible for most of the "anti-social" punishments: The top 5 anti-social punishers where responsible for 65.2% of the anti-social punishment tokens and 62.7% of the anti-social punishment actions.

Table 4.5: Hurdle estimate coefficients for punishment behavior. (Newey/West heteroscedasticy and autocorrelation consistent covariance matrix estimators) The estimates on the data of Gürerk et al. (2010a) are depicted as "no pun limit".

logit part	**no pun limit**		**with pun limit**	
(punish yes or no)	Coef.	Robust s.e.	Coef.	Robust s.e.
const	−0.2884	0.4053	−0.4864	0.3236
$recpun_{t-1}$	0.0543**	0.0258	−0.0563	0.0414
$\max(c_{j,t} - c_{i,t}; 0)$	0.1553***	0.0254		
$\max(c_{i,t} - c_{j,t}; 0)$	0.2557***	0.0285	0.3227***	0.0155
t	−0.1720***	0.0491	−0.2419***	0.0553
t^2	0.0046***	0.0014	0.0028*	0.0016
i: max $contr_{t-1}$	−0.3854	0.2785	−0.2332	0.2002
i: min $contr_{t-1}$	−0.0248	0.2144	0.9222***	0.3136
i: min $contr_t$	−0.3201	0.3023		
other min $contr_t$	0.4939*	0.2783	1.0645***	0.2231
other max $contr_t$	−1.8875***	0.2555		
n	−0.0664	0.0450	−0.1859***	0.0468
negative binomial part (if punish: size of punishment)				
const	0.5604**	0.2646	−12.0008	$5.36e+05$
$recpun_{t-1}$	−0.0020	0.0157	−0.0385	0.0620
$\max(c_{j,t} - c_{i,t}; 0)$	0.0240	0.0261		
$\max(c_{i,t} - c_{j,t}; 0)$	0.0915***	0.0137	0.1022	0.0981
t	−0.0354	0.0355	−0.1683	0.1737
t^2	0.0013	0.0011	0.0058	0.0066
i: max $contr_{t-1}$	−0.0209	0.1172	−0.4706	0.9888
i: min $contr_{t-1}$	−0.1540	0.1319	−0.6704	0.4840
i: min $contr_t$	−0.0428	0.2674		
other min $contr_t$	0.3119*	0.1498	0.6761*	0.4090
other max $contr_t$	0.2972***	0.2590		
n	−0.1104***	0.0387	−0.0332	0.1115

Signif. codes: *** 1%, ** 5%, * 10%

Result 6. *The limit on the sanctions prevents that punishment induces more punishment. It is also more focused on minimum contributors.*

The significant coefficients in the treatment without limit for sanctions towards higher contributing individuals and the size of punishment towards maximum contributors can be explained if one assumes the existence of the "do-gooders" derogation. There is some weak evidence for revenge in the treatment with the limit as well: a minimum contributor of the last period had a higher likelihood to punish others. However in that treatment neither the coefficient for the received punishment in the last period is significant nor is there a possibility to punish a higher contributor.

In the data with the punishment constraint the logit coefficient of the size of negative deviation – the punished subject contributing less than the punishing subject ($\max(c_{i,t} - c_{j,t}; 0)$), the logit and negative binomial coefficient for the minimum contributor dummy in t (*other min contr*$_t$) are larger compared to the estimates on the data of Gürerk et al. (2010a). Although the standard deviations indicate that the difference is not significant it gives an indication that someone contributing a unit less than others had with the limit a higher probability of being punished by others than without the limit.

4.6 Conclusion

Sanctions are a successful measure in order to achieve a high cooperation level in public-good experiments. But this comes at a cost since sanctions are expensive as they effectively burn welfare. In some cases the payoffs would be even higher with no cooperation at all due to excessive punishment. One phenomenon that makes sanctions more costly is that of "anti-social" punishments: low contributing subjects punishing high or maximal contributing subjects. We partially attribute those punishments to revenge and "do-gooders" derogation both being potentially triggered by emotions. The inequity aversion model of Fehr and Schmidt (1999) is not sufficient to explain this behavior.

As Gürerk et al. (2010a, 2006) have shown, one measure to improve the efficiency is to give the subjects the individual per-period choice whether they want to join a group with punishment option or another group without punishment option. Although this endogenous-institution-choice design leads to a success of the punishment institution at a much higher efficiency than comparable designs without institution-choice there still is an efficiency issue in the first periods with average payoffs far below the Nash equilibrium which represents the minimum payoff level when no one contributes. Thinking in evolutionary terms – a society could extinguish due to too heavy punishment before that punishment actually proves successful in establishing cooperation. Furthermore the punishment group was highly unpopular as less than a third of the subjects chose that group.

With our experiment we introduced a limit on punishment: it was only allowed to punish those subjects in the punishment group that contributed less than the sanctioning subject, to the endogenous-institution-choice design of Gürerk et al. (2010a). We had two aims: firstly, to check whether the low popularity of the sanctioning in related to the "anti-social" punishments, in other words that high contributors must fear to get punished, and secondly, to see whether such a rule improves the efficiency of the sanctioning institution. We compare our data with the imposed limit to the data of Gürerk et al. (2010a).

We found that the limit indeed induces higher efficiency: Average profits in the sanctioning institution never fall below the Nash equilibrium in contrast to Gürerk et al. (2010a). Although it does not increase the overall contribution level significantly it reduces the sum of the distributed punishment point to a great extent beyond the share of "anti-social" punishment that we observed in the data of Gürerk et al. (2010a). Thus a society with such a rule on punishment has to bear much lower costs to establish cooperation which in turn could mean a higher survival probability in evolutionary terms.

However, the punishment institution still remains highly unpopular. Exactly the same number of subjects joint the punishment institution as with Gürerk et al. (2010a). Thus it is certainly not the fear of the "anti-social" punishment that lets the subject abstain from joining the punishment institution. It could be the dislike of sanctions in general but also could be some optimizing behavior as in the first two periods the payoffs were significantly larger in the group without sanctions. But as soon the cooperation was established the "optimizers" could join the sanctioning institution which was earlier profitable compared with Gürerk et al. (2010a).

4.7 Appendix

4.7.1 Statistics

Table 4.6: Per-period exact Wilcoxon–Mann–Whitney tests, *p*-values.

Period	Contribution	Profit	People in PunWorld
1	0.69	0.02**	0.77
2	0.49	0.00***	0.73
3	0.82	0.49	0.95
4	0.49	0.01***	0.75
5	0.70	0.31	0.32
6	0.96	1.00	0.56
7	0.44	0.00***	0.18
8	0.24	0.05**	0.05**
9	0.91	0.57	0.01***
10	0.14	0.01***	0.16
11	0.59	0.37	0.01***
12	0.80	1.00	0.04**
13	0.18	0.09*	0.09*
14	0.77	0.14	0.01***
15	0.79	0.22	0.14
16	0.57	0.45	0.57
17	1.00	0.86	0.22
18	0.71	0.48	0.46
19	1.00	0.77	0.27
20	1.00	0.58	0.63
21	0.86	0.41	0.35
23	0.48	0.08*	0.16
24	0.47	0.86	0.01***
25	0.43	1.00	0.30
26	0.78	0.43	0.50
27	1.00	0.61	0.71
28	1.00	0.31	1.00
29	0.93	0.41	0.77
30	0.08	0.18	0.84

Sign.levels: *** 1%, ** 5%, * 10%

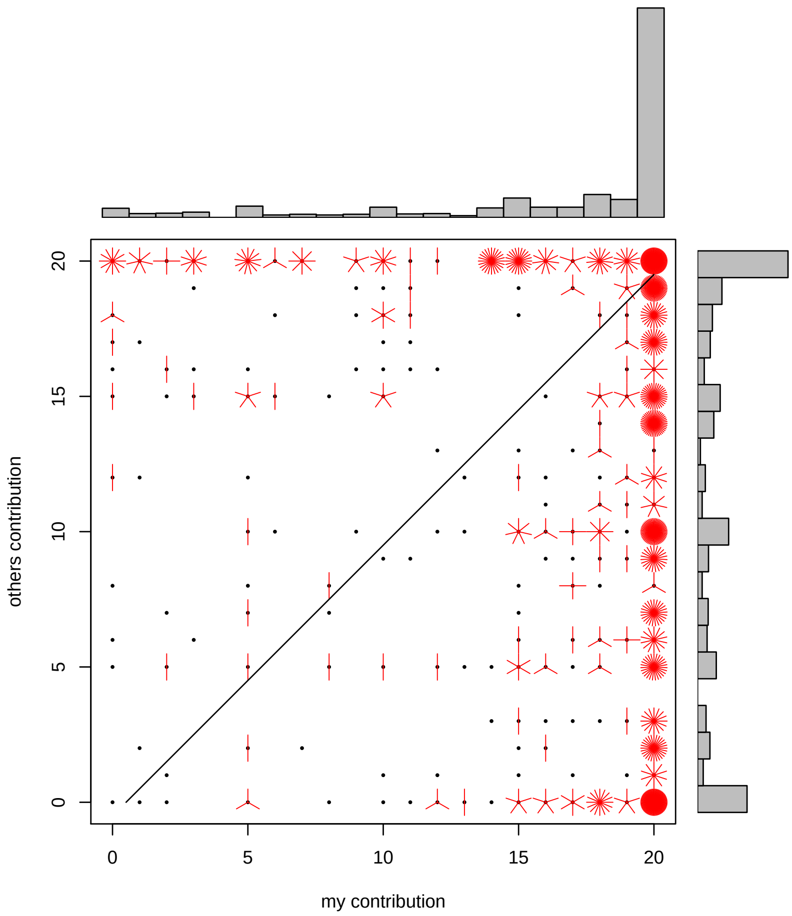

Figure 4.5: Sunflower plot of all punishment actions in Gürerk et al. (2010a), including a histogram for each axis. The x-axis is the perspective of the punishing (i.e., my contribution) and the y-axis the perspective of the punished subject (i.e., others contribution). The line represents the theoretical punishment rule. All sunflowers above that line represent "anti-social" punishments.

Table 4.7: Mixed-effects estimates of contribution changes.

increased contribution ($c_{i,t} \geq c_{i,t-1}$)	**no pun limit** Coef.	s.e.	**with pun limit** Coef.	s.e.
const	1.5155***	0.1726	2.6935***	0.1673
t	−0.1320***	0.0183	−0.2543***	0.0187
t^2	0.0028***	0.0005	0.0057***	0.0005
$\bar{c}^{pun}_{t-1} - c_{i,t-1}$ if positive	0.8143***	0.0355	0.4236***	0.0291
$c_{i,t-1} - \bar{c}_{t-1}$ if positive	−0.1012***	0.0268	−0.1976***	0.0412
$recpun_{t-1}$	0.0570***	0.0170	0.0526**	0.0218
decreased contribution ($c_{i,t} \leq c_{i,t-1}$)	**no pun limit** Coef.	s.e.	**with pun limit** Coef.	s.e.
const	−0.3305	0.2078	−0.9301***	0.2253
t	0.0265	0.0211	0.0990***	0.0253
t^2	−0.0006	0.0006	−0.0026***	0.0007
$\bar{c}^{pun}_{t-1} - c_{i,t-1}$ if positive	0.1161*	0.0665	0.1491***	0.0514
$c_{i,t-1} - \bar{c}^{pun}_{t-1}$ if positive	−0.1898***	0.0271	−0.1832***	0.0446
$recpun_{t-1}$	−0.1370***	0.0229	−0.1104***	0.0402

Signif. codes: *** 1%, ** 5%, * 10%

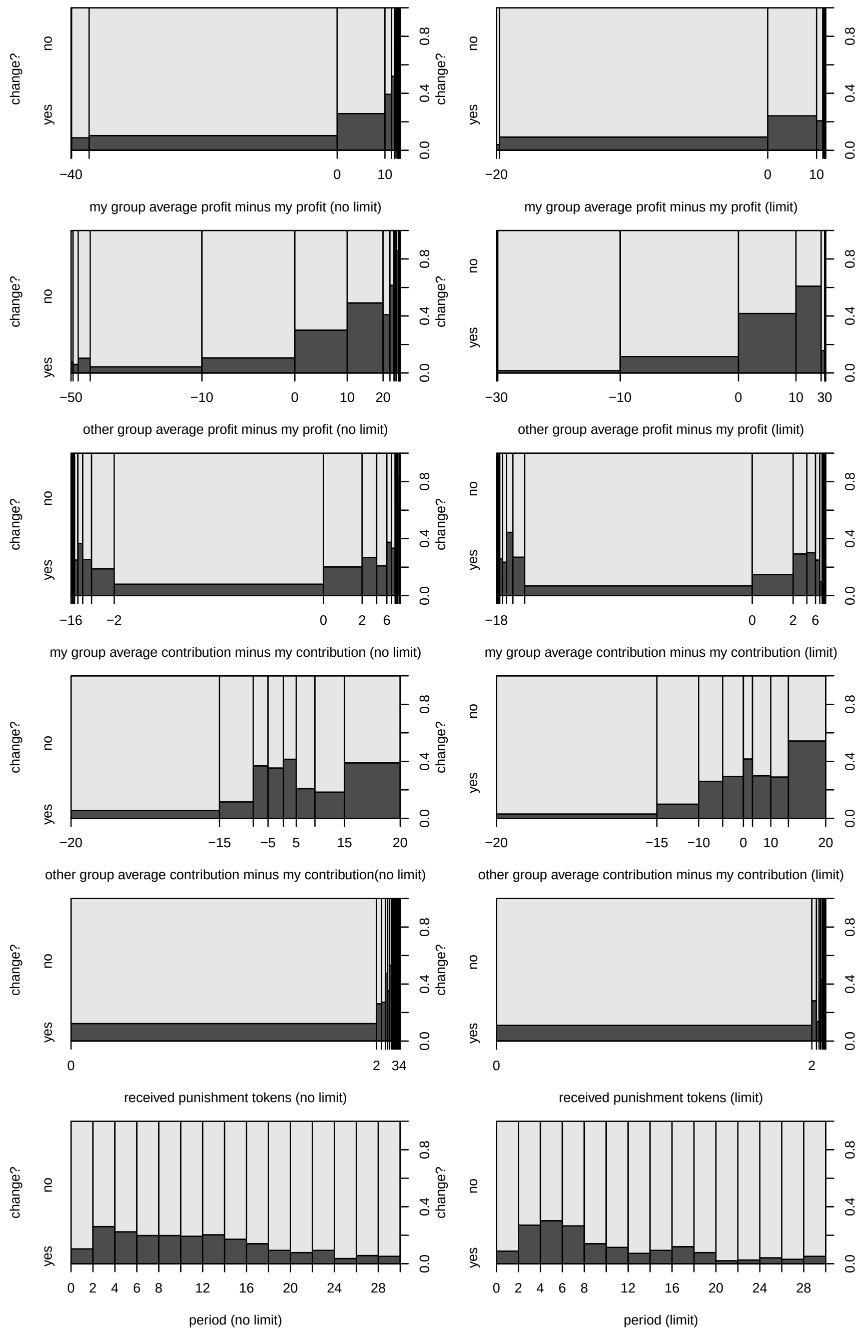

Figure 4.6: Spine-plots of the **general** institution change behavior. On the x-axis the dependent variable is depicted. The width of the bars represent the number of cases and the height of the dark shaded are the number of institution changes as percentage. The data of Gürerk et al. (2010a) are marked as "no limit".

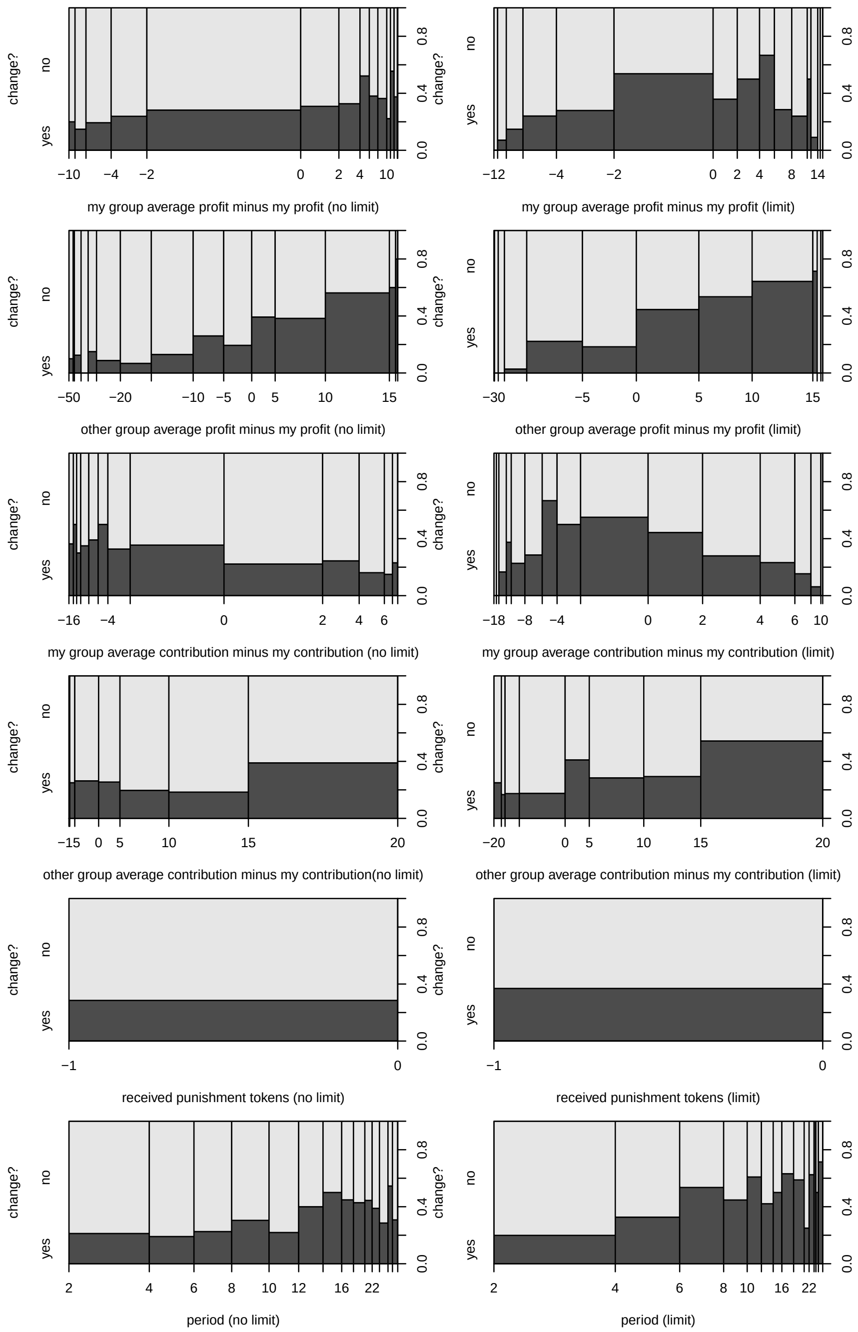

Figure 4.7: Spine-plots of the institution change behavior **towards pun**. On the x-axis the dependent variable is depicted. The width of the bars represent the number of cases and the height of the dark shaded are the number of institution changes as percentage. The data of Gürerk et al. (2010a) are marked as "no limit".

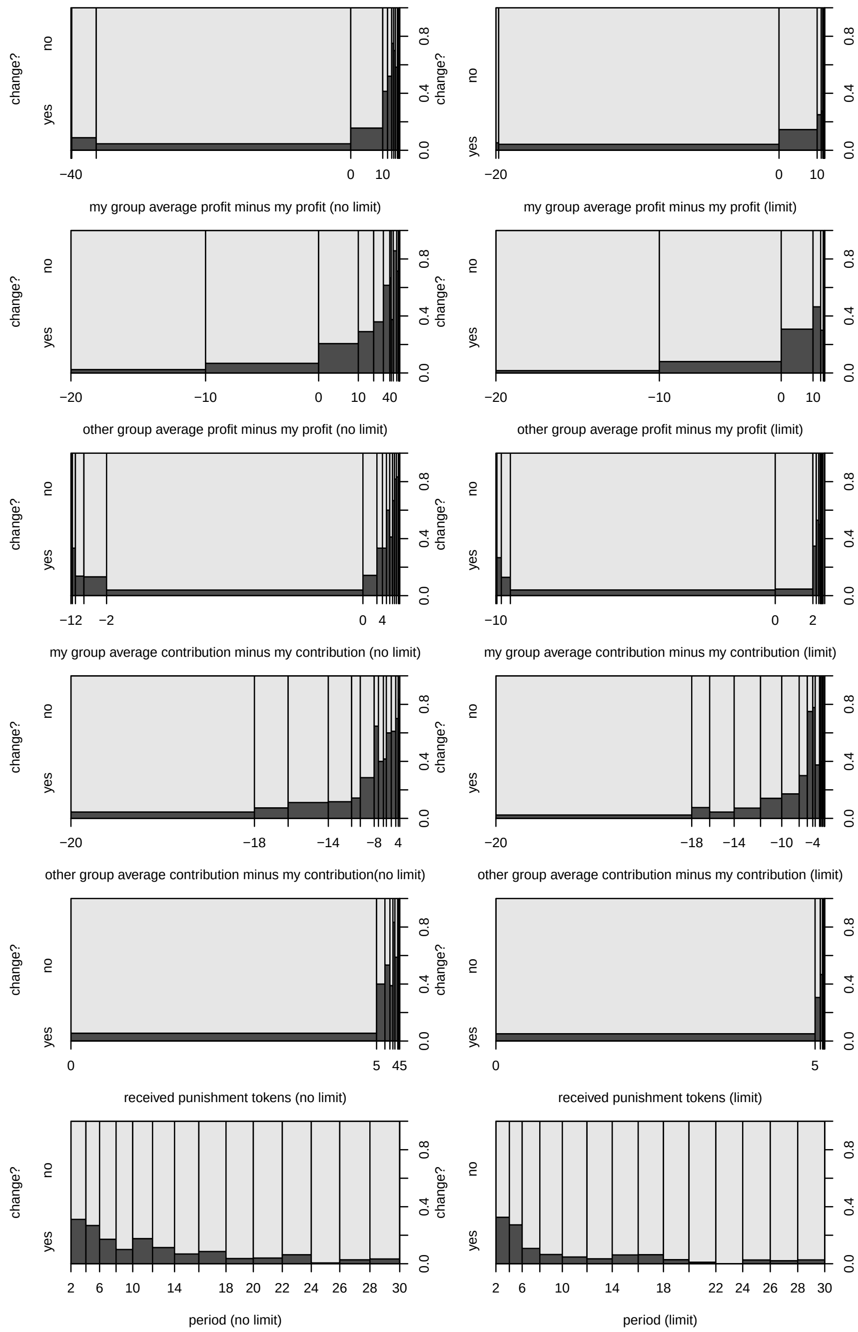

Figure 4.8: Spine-plots of the institution change **towards free** behavior. On the x-axis the dependent variable is depicted. The width of the bars represent the number of cases and the height of the dark shaded are the number of institution changes as percentage. The data of Gürerk et al. (2010a) are marked as "no limit".

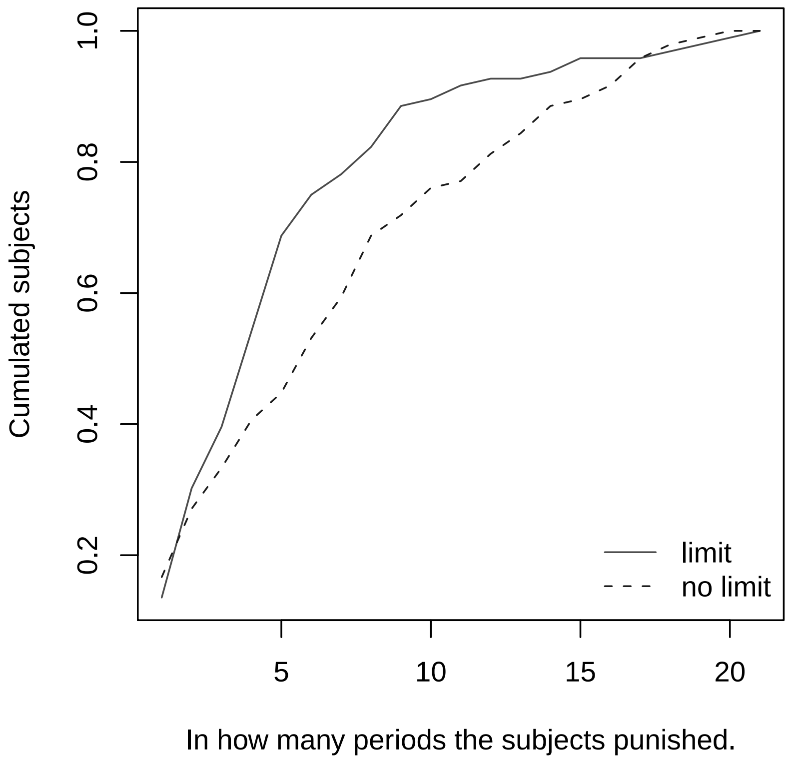

Figure 4.9: Distribution of punishment acts in the sense of how many subjects punished at least once x-times (on the x-axis). The data of Gürerk et al. (2010a) are marked as "no limit".

4.7.2 Translation of the Instructions

Common Information

At the beginning of the experiment you will be randomly assigned to *two* populations of 12 members each. During the whole experiment you will interact only with the members of your population. At the beginning of the experiment you will receive a *start capital of* 1000 *tokens* on your account.

Course

The experiment consists of 30 periods. Every period has two phases. In Phase 1 you choose your group and your contribution to the group project. In Phase 2 you can potentially influence the income of other group members.

Phase 1

(i) Choice of Group In Phase 1 every member decides which group to choose. There are two groups:

	Influence on the others income possible?
Group	**A** No
	B Yes, by distribution of negative points

(ii) The Contribution to the Group Project Every group member will receive in every period an **endowment** of **20 tokens**. You must decide how many of the 20 tokens you will **contribute** to the group project. You will **keep** the remaining tokens.

Calculation of Your Income in Phase 1 Your income in Phase 1 consists of two parts:

- The **tokens that you kept** = Endowment - contribution to the group project
- The **income out of the group project** = 1.6× sum of the contributions of all group members to the project / number of group members

Thus your income in Phase 1 calculates as follows: $20 -$ your contribution to the group project $+1.6\times$ sum of the contributions of all group members to the project / number of group members.

This formula is the same for all group members. **Please note** that every team member receives the same payout of the group project. Thus every group member benefits from all contributions towards the group project.

Phase 2

Distribution of Tokens In Phase 2 it will be revealed how much every group member contributed to the group project. **Note** that in every period the order of the list is reshuffled. Thus it is not possible to identify a group member just by the position in the list.

By **your token distribution** you can **reduce** the income of each other group members or **leave it unchanged**. You are only allowed to reduce the income of those group members that contribute **less** than you.

In every period of Phase 2 each member of both groups receives **20 additional tokens**. In group B you may decide how to **distribute** the tokens among the other team members. The remaining tokens will be added to your payoff. You can validate the costs of your token distribution with the button *Token calculation*. This button is not available if you are not able to distribute any tokens. In this case all the 20 additional tokens will be added to the period's payout. This is the case for all group A members.

- **Every token** that you distribute to a group member **reduces her income** by **3** tokens.
- If you do **not distribute** any token to another group member **her income will not change**.

Calculation of the Period Income in Phase 2 Your income in Phase 2 consists of 2 parts:

- The additional **tokens that you keep** = 20 − sum of tokens that you distributed towards other group members.
- **Three times the tokens** that **you received** by other group members.

Thus your income in Phase 2 calculates as follows: 20− sum of tokens that you distributed towards other group members − 3× received tokens.

Calculation of the Period Income

To sum up the overall income for each period calculates as follows:

Your income from Phase 1	= 20	- your contribution to the group + 1.6× (sum of all contributions) / number of group members
Your income from Phase 2	= 20	- sum of tokens you distributed - 3× tokens you received by other group members
your overall income		

Exception: Only one Group Member

If you are the only group member in your group you keep your 20 tokens endowment from Phase 1 as well as the additional tokens from Phase 2. Thus the period income in this case is 40. You have no possibility to interact.

Information at the End of the Period

At the end of the period you receive a detailed overview over the results: each member's contribution, income from Phase 1, distributed, received tokens as well as the income from Phase 2 and the overall income. The order again is randomized to prevent identification.

History

From period 2 on you will see an overview over the average results (as above) of the previous periods during the group choice.

Your Payout at the End of the Experiment

Your payout at the end is the starting capital of 1,000 tokens plus the sum of the period incomes of all 30 periods. The tokens will be converted into real money with an exchange rate of €1 per 100 tokens.

Note

During the whole experiment communication between the participants is not allowed. If you have a question you may raise the hand out of the cabin. All decisions are made anonymously. This means no participant will know who the other group members are and who made what decision. Furthermore the payment will be anonymous thus no one will know how much money the others received.

Good luck!

Bibliography

Alchian, A. A. and Demsetz, H. (1972). Production, information costs, and economic organization. *The American Economic Review*, 62(4):777–795.

Anderson, C. M. and Putterman, L. (2006). Do non-strategic sanctions obey the law of demand? The demand for punishment in the voluntary contribution mechanism. *Games and Economic Behavior*, 54(1):1–24.

Anderson, L. R., Mellor, J. M., and Milyo, J. (2008). Inequality and public good provision: An experimental analysis. *Journal of Socio-Economics, forthcoming*, 37(3):1010–1028.

Aquino, K., Steisel, V., and Kay, A. (1992). The effects of resource distribution, voice, and decision framing on the provision of public goods. *The Journal of Conflict Resolution*, 36(4):665–687.

Bagnoli, M. and McKee, M. (1991). Voluntary contribution games: Efficient private provision of public goods. *Economic Inquiry*, 29(2):351–366.

Baker, G. P., Jensen, M. C., and Murphy, K. J. (1988). Compensation and incentives: Practice vs. theory. *The Journal of Finance*, 43(3):593–616.

Banerjee, A. V. and Newman, A. F. (1993). Occupational choice and the process of development. *Journal of Political Economy*, 101(2):274–298.

Battaglini, M., Nunnari, S., and Palfrey, T. R. (2010). Political institutions and the dynamics of public investment. Technical report, Caltech.

Bochet, O., Page, T., and Putterman, L. (2005). Communication and punishment in voluntary contribution experiments. *Journal of Economic Behavior & Organization*, 60(1):11–26.

Bolton, G. E. and Ockenfels, A. (2000). ERC: A theory of equity, reciprocity, and competition. *American Economic Review*, 90(1):166–193.

Bowles, S. and Gintis, H. (1990). Contested exchange: New microfoundations of the political economy of capitalism. *Politics and Society*, 18(2):165–222.

Brosig, J., Weimann, J., and Ockenfels, A. (2003). The effect of communication media on cooperation. *German Economic Review*, 4(2):217–243.

Carpenter, J., Bowles, S., Gintis, H., and Hwang, S.-H. (2009). Strong reciprocity and team production: Theory and evidence. *Journal of Behavior & Organization*, 71(2):221–232.

Carpenter, J. P. (2007). The demand for punishment. *Journal of Economic Behavior & Organization*, 62(4):522–542.

Carpenter, J. P. and Matthews, Peter, H. (2005). Norm enforcement: Anger, indignation or reciprocity? Working Paper 1583, IZA.

Carpenter, J. P., Matthews, P. H., and Ong'ong'a, O. (2004). Why punish? Social reciprocity and the enforcement of prosocial norms. *Journal of Evolutionary Economics*, 14(4):407–429.

Chan, K. S., Mestelman, S., Moir, R., and Muller, A. R. (1999). Heterogeneity and the voluntary provision of public goods. *Experimental Economics*, 2(1):5–30.

Chaudhuri, A. (2011). Sustaining cooperation in laboratory public goods experiments: a selective survey of the literature. *Experimental Economics*, forthcoming.

Cherry, T. L., Frykblom, P., and Shogren, J. F. (2002). Hardnose the dictator. *The American Economic Review*, 92(4):1218–1221.

Cherry, T. L., Kroll, S., and Shogran, J. F. (2005). The impact of endowment heterogeneity and origin on public good contributions: Evidence from the lab. *Journal of Economic Behavior & Organization*, 57(3):357–365.

Cinyabuguma, M., Page, T., and Putterman, L. (2006). On perverse and second-order punishment in public goods experiments with decentralized sanctioning. *Experimental Economics*, 9(3):265–279.

Clark, J. E. (2002). House money effects in public good experiments. *Experimental Economics*, 5(3):223–231.

Craig, B. R. and Pencavel, J. (1995). Participation and productivity – a comparison of worker cooperatives and conventional firms in the plywood industry. *Brookings Papers on Economic Activity, Microeconomics Annual*, 1995:121–160.

Croson, R. T. and Marks, M. B. (2000). Step returns in threshold public goods: A meta- and experimental analysis. *Experimental Economics*, 2:239–259.

Davis, D. D. and Holt, C. A. (1993). *Experimental Economics*. Princeton University Press.

de Quervain, D. J. F., Fischbacher, U., Treyer, V., Schellhammer, M., Schnyder, U., Buck, A., and Fehr, E. (2004). The neural basis of altruistic punishment. *Science*, 305(5688):1254–1258.

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

Denant-Boemont, L., Masclet, D., and Noussair, C. (2007). Punishment, counterpunishment and sanction enforcement in a social dilemma experiment. *Economic theory*, 33(1):145–167.

Dow, G. K. (2003). *Governing the Firm: Workers Control in Theory and Practice*. Cambridge University Press, New York.

Dow, G. K. and Putterman, L. (2000). Why capital suppliers (usually) hire workers: what we know and what we need to know. *Journal of Economic Behavior & Organization*, 43(3):319–336.

Duffy, J., Ochs, J., and Vesterlund, L. (2007). Giving little by little: Dynamic voluntary contribution games. *Journal of Public Economics*, 91(9):1708–1730.

Dufwenberg, M. and Kirchsteiger, G. (2000). Reciprocity and wage undercutting. *European Economic Review*, 44(4-6):1069–1078.

Egas, M. and Riedl, A. (2008). The economics of altruistic punishment and the maintenance of cooperation. *Proceedings of the Royal Society B: Biological Sciences*, 275(1637):871–878.

Ertan, A., Page, T., and Putterman, L. (2009). Who to punish? Individual decisions and majority rule in the solution of free rider problems. *European Economic Review*, 53(5):495–511.

Estrin, S., Jones, D. C., and Svejnar, J. (1987). The productivity effects of worker participation – producer cooperatives in western economies. *Journal of Comparative Economics*, 11(1):40–61.

Eswaran, M. and Kotwal, A. (1984). The moral hazard of budget-breaking. *Rand Journal of Economics*, 15(4):578–581.

Eswaran, M. and Kotwal, A. (1989). Why are capitalists the bosses. *Economic Journal*, 99(394):162–176.

Falk, A. and Fischbacher, U. (2006). A theory of reciprocity. *Games and Economic Behavior*, 54(2):293 – 315.

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

Fehr, E. and Gächter, S. (2000b). Fairness and retaliation: The economics of reciprocity. *Journal of Economic Perspectives*, 14(3):159–181.

Fehr, E. and Gächter, S. (2002). Altruistic punishment in humans. *Nature*, 415(6868):137–150.

Fehr, E., Hoff, K., and Kshetramade, M. (2008). Spite and development. *American Economic Review*, 98(2):494–499.

Fehr, E. and Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics*, 114(3):817–868.

Fehr, E. and Schmidt, K. M. (2000). Fairness, incentives, and contractual choices. *European Economic Review*, 44(4-6):1057–1068.

Fershtman, C. and Nitzan, S. (1991). Dynamic voluntary provision of public goods. *European Economic Review*, 35(5):1057 – 1067.

Fischbacher, U. (2007). z-tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10(2):171–178.

Fischbacher, U. and Gächter, S. (2010). Social preferences, beliefs and the dynamics of free riding in public good experiments. *American Economic Review*, 100(1):541–56.

Fischbacher, U., Gächter, S., and Fehr, E. (2001). Are people conditionally cooperative? Evidence from a public goods experiment. *Economic Letters*, 71(3):397–404.

Frijda, N. H. (1994). The lex talionis: On vengeance. In van Goozen, S. H., van de Poll, N. E., and Sergeant, J. A., editors, *Emotions: Essays on emotion theory*, chapter 10, pages 263–289. Lawrence Erlbaum Associates, Hillsdale, New Jersey Hove.

Frohlich, N., Godard, J., Oppenheimer, J. A., and Starke, F. A. (1998). Employee versus conventionally-owned and controlled firms: An experimental analysis. *Managerial and Decision Economics*, 19(4-5):311–326.

Gächter, S. (2007). Conditional cooperation: behavioral regularities from the lab and the field and their policy implications. In Frey, B. S. and Stutzer, A., editors, *Psychology and economics*, Cambridge, MA. MIT Press.

Gächter, S. and Herrmann, B. (2009). Reciprocity, culture and human cooperation: Previous insights and a new cross-cultural experiment. *Philosophical Transactions of the Royal Society B - Biological Sciences*, 364(1518):791–806.

Gächter, S., Renner, E., and Sefton, M. (2008). The long-run benefits of punishment. *Science*, 322(5907):1510.

Gollwitzer, M. (2004). *Eine Analyse von Racheaktionen und rachebezogenen Reaktionen unter gerechtigkeitspsychologischen Aspekten*. PhD thesis, Universität Trier Fachbereich Psychologie.

Gradstein, M. and Nitzan, S. (1990). Binary participation and incremental provision of public goods. *Social Choice and Welfare*, 7(2):171–192.

Greiner, B. (2004). An online recruitment system for economic experiments. In Kremer, K. and Macho, V., editors, *Forschung und wissenschaftliches Rechnen 2003*, volume 63, pages 79–93. Ges. für Wiss. Datenverarbeitung, Göttingen.

Grosse, S., Putterman, L., and Rockenbach, B. (2011). Monitoring in teams: Using laboratory experiments to study a theory of the firm. *Journal of the European Economic Association*, 9(4).

Gürerk, Ö., Irlenbusch, B., and Rockenbach, B. (2006). The competitive advantage of sanctioning institutions. *Science*, 312(5770):108–111.

Gürerk, Ö., Irlenbusch, B., and Rockenbach, B. (2010a). Voting with feet – community choice in social dilemmas. Technical report, University of Erfurt.

Gürerk, Ö., Rockenbach, B., and Wolff, I. (2010b). The effects of punishment in dynamic public-good games. Technical report, University of Erfurt / CEREB.

Güth, W., Schmittberger, R., and Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior and Organization*, 3(4):367–388.

Hardin, G. (1968). The tragedy of the commons. *Science*, 162(3859):1243–1248.

Harrison, G. (2007). House money effects in public good experiments: Comment. *Experimental Economics*, 10(4):429–437.

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., McElreath, R., Alvard, M., Barr, A., Ensminger, J., Henrich, N. S., Hill, K., Gil-White, F., Gurven, M., Marlowe, F. W., Patton, J. Q., and Tracer, D. (2005). 'Economic Man' in cross-cultural perspective: Behavioral experiments in 15 small-scale societies. *Behavioral and Brain Sciences*, 28(6):795–855.

Henrich, J., McElreath, R., Barr, A., Ensminger, J., Barrett, C., Bolyanatz, A., Cardenas, J. C., Gurven, M., Gwako, E., Henrich, N., Lesorogol, C., Marlowe, F., Tracer, D., and Ziker, J. (2006). Costly punishment across human societies. *Science*, 312(5781):1767–1770.

Herrmann, B., Thöni, C., and Gächter, S. (2008). Antisocial punishment across societies. *Science*, 319(5868):1362–1367.

Holmström, B. (1982). Moral hazard in teams. *Bell Journal of Economics*, 13(2):324–340.

Hopfensitz, A. and Reuben, E. (2009). The importance of emotions for the effectiveness of social punishment. *Economic Journal*, 119(540):1534–1559.

Irlenbusch, B. and Ruchala, G. K. (2008). Relative rewards within team-based compensation. *Labour Economics*, 15(2):141–167.

Isaac, M. A. and Walker, J. M. (1988a). Group-size effects in public-goods provision - the voluntary contributions mechanism. *Quarterly Journal of Economics*, 103(1):179–199.

Isaac, M. R., Walker, J., and Thomas, S. (1984). Divergent evidence on free riding: An experimental examination of some possible explanations. *Public Choice*, 63(2):113–149.

Isaac, M. R. and Walker, J. M. (1988b). Communication and free-riding behavior: The voluntary contributions mechanism. *Economic Inquiry*, 26(4):585–608.

Kandel, E. and Lazear, E. P. (1992). Peer pressure and partnerships. *The Journal of Political Economy*, 100(4):801–817.

Keser, C. (1996). Voluntary contributions to a public good when partial contribution is a dominant strategy. *Economics Letters*, 50(3):359–366.

Keser, C. (2000). Strategically planned behavior in public good experiments. Technical Report 2000s-35, CIRANO, Montreal.

Kihlstrom, R. E. and Laffont, J.-J. (1979). A general equilibrium entrepreneurial theory of firm formation based on risk aversion. *Journal of Political Economy*, 87(4):719–747.

Knight, F. H. (1921). *Risk, Uncertainty and Profit*. Hart, Schaffner & Marx; Houghton Mifflin Company, Boston.

Kremer, M. (1997). Why are worker cooperatives so rare? *NBER Working Paper*, 6118.

Kruse, D. L. (1993). *Profit Sharing: Does it Make a Difference?* Kalamazoo, MI: Upjohn Institute for Employment Research.

Ledyard, J. O. (1995). Public goods: A survey of experimental research. In Kagel, J. and Roth, A., editors, *Handbook of Experimental Economics*, pages 111–194. Princeton University Press.

Levati, V. M., Sutter, M., and van der Heijden, E. (2007). Leading by example in a public goods experiment with heterogeneity and incomplete information. *Journal of Conflict Resolution*, 51(5):793–818.

MacLeod, W. B. (1988). *Equity, Efficiency and Incentives in Cooperative Teams*, volume 3. JAI Press, Greenwich.

Marglin, S. (1974). What do bosses do? The origins and functions of hierarchy in capitalist production. *Review of Radical Political Economics*, 6(2):60–112.

Marwell, G. and Ames, R. E. (1979). Experiments on the provision of public goods. I. Resources, interest, group size and the free-rider problem. *The American Journal of Sociology*, 84(6):1335–1360.

Marwell, G. and Ames, R. E. (1980). Experiments on the provision of public goods. II. Provision points, stakes, experience, and the free-rider problem. *The American Journal of Sociology*, 85(4):926–937.

Marx, L. M. and Matthews, S. A. (2000). Dynamic voluntary contribution to a public project. *Review of Economic Studies*, 67(2):327–358.

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

Masclet, D. and Villeval, M. C. (2008). Punishment, inequality and welfare: A public good experiment. *Social Choice and Welfare*, 31(3):475–502.

Milinski, M., Semmann, D., Krambeck, H.-J., and Marotzke, J. (2006). Stabilizing the earth's climate is not a losing game: Supporting evidence from public goods experiments. *Proceedings of the National Academy of Sciences*, 103(11):3994–3998.

Monin, B. (2007). Holier than me? threatening social comparison in the moral domain. *International Review of Social Psychology*, 20(1):53–68.

Nalbantian, H. R. and Schotter, A. (1997). Productivity under group incentives: An experimental study. *American Economic Review*, 87(3):314–341.

Nicklisch, A. and Wolff, I. (2011). Cooperation norms in multiple-stage punishment. *Journal of Public Economic Theory*. forthcoming.

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

Nikiforakis, N. and Engelmann, D. (2010). Altruistic punishment and the threat of feuds. Working Paper 1709204, SSRN.

Nikiforakis, N. and Normann, H.-T. (2008). A comparative statics analysis of punishment in public-good experiments. *Experimental Economics*, 11(4):358–369.

Noussair, C. and Soo, C. (2008). Voluntary contributions to a dynamic public good: Experimental evidence. *Economics Letters*, 98(1):71–77.

Noussair, C. and Tucker, S. (2005). Combining monetary and social sanctions to promote cooperation. *Economic Inquiry*, 43(3):649–660.

Offerman, T., Schram, A., and Sonnemans, J. (1998). Quantal response models in step-level public good games. *European Journal of Political Economy*, 14(1):89 – 100.

Offerman, T., Sonnemans, J., and Schram, A. (1996). Value orientations, expectations and voluntary contributions in public goods. *The Economic Journal*, 106(437):817–845.

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

Page, T., Putterman, L., and Unel, B. (2005). Voluntary association in public goods experiments: Reciprocity, mimicry and efficiency. *The Economic Journal*, 115(506):1032–1053.

Potters, J., Sefton, M., and van der Heijden, E. (2009). Hierarchy and opportunism in teams. *Journal of Economic Behavior and Organization*, 69(1):39–50.

Putterman, L., Tyran, J.-R., and Kamei, K. (2010). Public goods and voting on formal sanction schemes: An experiment. Discussion papers 10-02, University of Copenhagen, Department of Economics.

Rapoport, A. and Suleiman, R. (1993). Incremental contribution in step-level public-goods games with asymmetric players. *Organizational Behavior and Human Decision Processes*, 55(2):171–194. English.

Reuben, E. and Riedl, A. (2009). Enforcement of contribution norms in public good games with heterogeneous populations. *IZA Discussion Paper*, 4303.

Rockenbach, B. and Milinski, M. (2006). The efficient interaction of indirect reciprocity and costly punishment. *Nature*, 444:718–723.

Sadrieh, A. and Verbon, H. A. (2006). Inequality, cooperation, and growth: An experimental study. *European Economic Review*, 50:1197–1222.

Saijo, T. and Nakamura, H. (1995). The "spite" dilemma in voluntary contribution mechanism experiments. *Journal of Conflict Resolution*, 39(3):535–560.

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

Sefton, M. and Steinberg, R. (1996). Reward structures in public good experiments. *Journal of Public Economics*, 61:263–287.

Shapiro, C. and Stiglitz, J. E. (1984). Equilibrium unemployment as a worker discipline device. *American Economic Review*, 74(3):433–444.

Smith, V. L. (1979). An experimental comparison of three public good decision mechanisms. *Scandinavian Journal of Economics*, 81(2):198–215.

Sutter, M., Haigner, S., and Kocher, M. G. (2008). Choosing the stick or the carrot? endogenous institutional choice in social dilemma situations. Working Paper 2008-07, University of Innsbruck.

Sutter, M. and Weck-Hannemann (2003). On the effects of asymmetric and endogenous taxation in experimental public goods games. *Economics Letters*, 79(1):59–67.

Theil, H. (1967). *Economics and Information Theory*. North-Holland Publishing, Amsterdam.

Trivers, R. L. (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology*, 46:35–57. article grosse.

van de Kragt, A. J., Orbell, J. M., and Dawes, R. M. (1983). The minimal contributing set as a solution to public goods problems. *The American Political Science Review*, 77(1):112–122.

van Dijk, E. and Grodzka, M. (1992). The influence of endowments asymmetry and information level on the contribution to a public step good. *Journal of Economic Psychology*, 13(2):329–342.

van Dijk, E. and Wilke, H. (1995). Coordination rules in asymmetric social dilemmas: A comparison between public good dilemmas and resource dilemmas. *Journal of Experimental Social Psychology*, 31(1):1–27.

Weitzman, M. and Kruse, D. (1990). *Profit Sharing and Productivity*. Brookings Institution, Washington, DC.