DFG Research Group 2104

Need-Based Justice and Distribution Procedures

Trigger-Happy or Precisionist? On Demand for Monitoring in a Noisy Social Dilemma Game

Andreas Nicklisch, Louis Putterman and Christian Thöni

Working Paper Nr. 2017-20

http://bedarfsgerechtigkeit.hsu-hh.de/dropbox/wp/2017-20.pdf

Date: 2019-05-20

DFG Research Group 2104 at Helmut Schmidt University Hamburg

http://needs-based-justice.hsu-hh.de

Trigger-Happy or Precisionist? On Demand for Monitoring in a Noisy Social Dilemma Game

Andreas Nicklisch[*], Louis Putterman[†], and Christian Thöni[‡]

May 20, 2019

Abstract

Recent experimental studies question whether societies can self-govern social dilemmas with the help of decentralized punishment opportunities. One important challenge for the mechanism is imperfect information about cooperative behavior. It has been shown that imperfect information increases misdirected punishment and thereby hampers the efficacy of the punishment mechanism. We study an environment with monitoring opportunities, in which subjects can improve the quality of their information at a cost. We find experimentally that the majority of subjects are willing to pay a modest cost to improve their information. The demand for monitoring is price sensitive, but does not systematically depend on whether other subjects are informed about the monitoring decision. Almost no subjects take up the chance to monitor partially at a lower price. Rather subjects choose to monitor either perfectly or not at all. Little punishment takes place with imperfect information. The large majority of those subjects who monitor subsequently punish non-cooperative behavior, leading to a substantial and significant improvement in efficiency.

Keywords: public goods, peer punishment, costly monitoring,
JEL-Classification: C92, D02, H41

[*]Nicklisch: HTW Chur and research group FOR 2104 "Need-based justice and distributive procedures", 7004 Chur, Switzerland, email: andreas.nicklisch@htwchur.ch.

[†]Corresponding author: Brown University, Providence, Rhode Island 02912, USA, email: louis_putterman@brown.edu, tel: (401) 863-3837, fax: (401) 863-1970.

[‡]University of Lausanne, 1015 Lausanne, Switzerland, email: christian.thoeni@unil.ch.

1 Introduction

As experimental and behavioral approaches have increased their contributions to economics, there has been both good news and bad news regarding the ability of human agents to cooperate in the face of social dilemmas. On the one hand, empirical studies have suggested to many that the traditional assumption of universal selfish maximization might yield less accurate predictions than models consistent with conditional cooperation and assuming heterogeneity of type and belief in place of traditional "common knowledge." These views result from the observation of considerable amounts of cooperation in one shot and finitely repeated games, and of relatively sustained cooperation when players are able to sort by type, to communicate, or to sanction each other. On the other hand, important questions have been raised regarding the realism of the environments in which what Ostrom, Walker, and Gardner (1992) called "self-governance" has been demonstrated. One line of questioning to which the present authors have contributed concerns the problem of imperfect and/or costly information. For example, Page, Putterman, and Unel (2005) find that cooperators endogenously sort into groups to sustain cooperation in public goods games, but Kamei and Putterman (2017) find less success in this regard the less perfect is players' information about one another's behaviors. In related examples, Fehr and Gächter (2000) and Gächter, Renner, and Sefton (2008) find that high contributors incur costs to punish free riders and thereby to promote cooperation, when the contribution of each group member is accurately shown to the others at no cost, but Grechenig, Nicklisch, and Thöni (2010) and Ambrus and Greiner (2012) find large increases in "misdirected" punishment that result in less or even negative effects of punishment opportunities on cooperation, when similar interactions occur under imperfect observability.

The present paper addresses the challenge that imperfect and costly information poses for cooperative self-governance by introducing a natural extension to the experimental literature: the incorporation of a costly monitoring option. Put differently, we endogenize the imperfectness of information by letting the agents concerned improve their information's accuracy, should they be willing to incur the requisite cost. The message of Grechenig et al. (2010) and Ambrus and Greiner (2012) might appear to be summarized by the phrase "punishment despite reasonable doubt," which suggests that agents in their social dilemma experiments who were offered the opportunity to engage in peer punishment based on information known to be frequently inaccurate, were not especially reticent about punishing one another although they might be punishing a fellow cooperator when attempting to enforce cooperation. However, a truly cavalier attitude towards punishing in such conditions would imply having little willingness to pay for better information, were it to be available. By offering opportunities to improve information at relatively low as well as higher cost, we investigate how far that characterization should be pushed.[1] We obtain, in the event, the good news that many experimental participants are in fact quite willing to pay a modest cost to improve their information, and that little punishment in fact takes place with imperfect information when a monitoring opportunity of modest cost is placed on offer. A particularly impressive aspect of our findings is that although we offer our subjects the chance to buy a partial improvement in information for only half the price of attaining complete accuracy, almost no subjects take up this offer. Subjects who choose to monitor almost always choose to do so to the highest available degree. And subjects who choose not to buy information also choose not to punish in the large majority of instances.

Despite the positive finding that subjects are willing to pay a modest cost to improve their information, this may not mean that the provision of costly monitoring combined with the peer punishment mechanism guarantees well-functioning self-governance and persistent voluntary cooperation. The reason is that monitoring is in itself a third-order public good. That is, one may choose to free-ride on the monitoring of others when individual monitoring decisions are public. Moreover, in cases in which monitoring is public knowledge, if one knew that others are not paying to improve the accuracy with which they observe one's behavior, there are good reasons to reduce one's contributions: first, if one anticipates that others who do not incur the cost to monitor accurately are reluctant to punish, free-riding is associated with a smaller risk of punishment the less monitoring is done. Second, if there is punishment based on noisy monitoring, there is a substantial chance that free-riding will go undetected and unpunished, and a chance that contributors will receive misdirected punishment, so the presence of peer punishment opportunities may do less to motivate contributing the less monitoring is done.

At the same time, monitoring may send an indirect public warning to those monitored that punishment could be forthcoming should they ignore norms of cooperation. Hence, monitoring might make punishment less necessary.[2] Anticipation of that effect might provide additional motivation for

[1] Related to our work is the paper by DeAngelo and Gee (2017): they analyze people's willingness to monitor others' contributions at all when monitoring is costly and a precondition for sanctioning subsequently. That is, the issue of their paper is not whether people want to improve their knowledge on which basis they may or may not execute sanctions, but whether people want to spend money for observing others and may or may not sanction them. Aoyagi, Bhaskar, and Fréchette (2016) have also a somewhat related paper in which they analyze the impact of public versus private knowledge about the noisiness of others' signal on voluntary cooperation in infinitely repeated prisoner's dilemmas. They observe quite different strategies depending on whether information quality is private vs. public but do not study information quality as a costly decision variable.

[2] In a similar vein, Ramalingam, Godoy, Morales, and Walker (2016) analyze the effect of individuals' announcements to acquire the right to punish other players in a voluntary contribution mechanism. Their results indicate that the unilateral announcement of individuals regarding their sanctioning right increase cooperation rates when the right to

incurring the cost to monitor. Since there are plausible reasons why publicness of monitoring may spur, but also why it may hinder, both contributions and monitoring investments, the question is best addressed by empirical means. For this reason, we investigate how (if at all) the observability of the monitoring decision affects both contribution and monitoring choices: we introduce "public" and "private" monitoring versions of our high and low monitoring cost treatments in a 2x2 factorial design, testing whether costly monitoring yields different results when the monitoring decision is public vs. private.

Overall, our results are consistent with monitoring having a "warning" effect: more monitoring is clearly associated with less free riding on contributions, when it is public. However, the anticipation of this effect appears to induce greater monitoring, if at all, only in the low monitoring cost condition. Perhaps a considerable number of subjects in the treatment having both high monitoring cost and publicness of the monitoring decision consider both monitoring and punishment cost before sending a signal, then, deciding that the combined cost is too high, demur from engaging in monitoring altogether. In our paper, we consider the inconsistent impact of monitoring's publicness on the decision to undertake it at some length, but we view our results in that dimension to be more provisional and of less fundamental importance than those about the willingness to incur a cost to monitor at a more general level, as discussed above.

There are already large empirical and theoretical literatures on costly monitoring, but little in the way of laboratory or other controlled experiments on the topic. One prominent area of application has been to the organization of work teams and their management. In a seminal paper, Alchian and Demsetz (1972) (which had over 16,000 citations according to Google scholar, when recently accessed) argued that workers' abilities to reduce effort without correspondingly sharp reductions in compensation, due to the imperfect observability of effort in teams, provides the core explanation of why economies of specialization and team production are not usually realized by worker partnerships. Those economies, the authors argued, are instead harnessed in asymmetric employment relationships where specialist monitors can claim the profit they obtain by carefully observing worker effort, paying proportionate rewards, and earning the profit as an incentive to accurately perform those roles. Note that for Alchian and Demsetz, good monitoring is something workers would view as being in their own interest, not a warning of potential punishment for choosing the wrong effort level. In efficiency wage models like that of Shapiro and Stiglitz (1984) and Bowles and Gintis (1990), in contrast, firms can induce higher or lower effort simply by announcing a higher or lower monitoring level—an anticipatory role of monitoring that, as mentioned above, requires that the occurrence of

sanction entails a monetary cost.

monitoring be publicly known.

Yet the claim that costly information in the workplace necessitates top down or specialized monitoring has been a controversial one at the empirical level. Labor economists who have investigated the matter, especially Pencavel (2013) and Craig and Pencavel (1995), find that in enterprises of modest size, at least, mutual monitoring is in fact a comparative advantage of profit-sharing enterprises, and their ability to save on the cost of hiring specialist monitors is a recognized cost advantage.[3] Arguably, the knowledge of being observable by fellow workers is one of the factors leading workers to exert more effort in such firms, but we are unaware of any empirical analysis of how (in this context) varying the observability of monitoring affects the extent to which it is undertaken.[4]

Problems of observation and monitoring likewise arise in other contexts. Ostrom's (1990) discussion of attempts to resolve commons problems among fishermen suggests that differences of observability may be key to the success or failure of alternative governance arrangements. The same applies in areas like the policing of restrictions on harvesting trees, where Ostrom found village self-policing could be superior to government supervision in part due to superior observability of one another by community members. Experimental public economics has focused on quality of information, with differences in likelihood of detection being one of the main variables manipulated in studies of tax compliance (Alm, 2012). Manipulation of public perception of the frequency of audits has also been discussed as a policy tool (see, for example, the survey article by Torgler, 2002).

More broadly, "transparency" is frequently mentioned as a key desideratum of effective governance, and it may well depend not only on rules and practices of officials, but also on the inclination of citizens to expend resources on monitoring their behavior. The free press itself, which is touted as a crucial underpinning of democracy, may be capable of fulfilling that function in a self-sustaining manner only to the extent that citizens show sufficient interest as consumers of its investigations and exposés. Govern-

[3]An experiment that finds some support for the Alchian and Demsetz argument is van der Heijden, Potters, and Sefton (2009), while one obtaining more mixed results on the question is Grosse, Putterman, and Rockenbach (2011). A theoretical appraisal of Alchian-Demsetz and related literature is provided by Dow (2017). Kremer (1997) argues that worker-owned enterprises may typically fail due to internal political pressures towards wage compression, a phenomenon indirectly supported by the experimental results of Balafoutas, Kocher, Putterman, and Sutter (2013), and by the empirical findings of Burdín (2016).

[4]For public goods games where subjects have the choice between different punishment institutions, the results of Nicklisch, Grechenig, and Thöni (2016) suggest that under limited observability subjects tend to favor centralized structures of enforcement, while with perfect information decentralized punishment prevails. Markussen, Putterman, and Wang (2017) obtain a similar result under imperfect information, whereas centralized and decentralized punishment are approximately equally popular when neither suffers problems of observability.

mental monitoring of citizens' (e.g., of their tax compliance) and citizens' monitoring of governmental non-corruption are both examples of costly monitoring in the public sphere.

The rest of our paper proceeds as follows. In Section 2, we describe the design of our experiment, which uses the well-known experimental paradigm of a public goods game with punishment played a finitely repeated number of times in groups of fixed composition, extends it to the domain of imperfect information, and introduces the possibility of monitoring. In Section 3, we discuss theoretical considerations and offer behavioral conjectures on monitoring. Section 4 presents and analyzes the results of the experiment, and Section 5 concludes.

2 Design and procedures

Our experimental tool is the standard voluntary contribution mechanism (VCM) with decentralized punishment. We analyze behavior in a standard repeated VCM game with four players per group and twenty periods. The group composition remains constant over the periods (partner design). At the beginning of each period, each player receives an endowment of 22 ECU (experimental currency units). In a first stage, each player chooses whether or not to monitor. We denote monitoring as m_i, where $m_i \in \{0,1,2\}$. Monitoring is costly, the cost per unit of information being κ ECU.

Players simultaneously—in some treatments knowing, in others not knowing, the first stage's monitoring decisions—choose how many ECU from their endowment to contribute to the public good, g_i, with $g_i \in \{0,1,2,\dots,20\}$. Notice that we restrict contributions to at most 20 ECU, which (given the κ's specified below) leaves it possible to both maximally contribute and fully monitor while staying within the available budget. Each ECU contributed to the public good yields a benefit of 0.4 ECU (the marginal per capita return) to every player in the group.

After the contributions are made, each player receives a signal s_j $(j \neq i)$ about the contributions of each other player in the group, such that

$$s_j = \begin{cases} g_j & \text{with probability } 0.5 + 0.25 m_i \\ \tilde{g}_j & \text{with probability } 0.5 - 0.25 m_i, \end{cases} \tag{1}$$

where $\tilde{g}_j$ is an independent random draw from $\{0,1,2,\dots,20\} \setminus \{g_j\}$, all numbers with equal probability. Thus, for the contribution signals of the other three players, there is one independent random draw for each player within each group determining whether players with the same accuracy level of information receive the accurate signal, and if not, another independent draw which determines a random number to display. That is, all players in the group who choose $m_i = 0$ $(m_i = 1)$ see the same accurate or false signal.

Each number (except g_i) is equally likely to appear if the signal does not correspond to the true contribution. For example, suppose that player 1 contributes 10 ECU and player 2 chose $m_i = 1$ beforehand. There is a probability of 75 percent that player 2 sees the signal "10 ECU" for player 1's contribution, while with a probability of 25 percent player 2 sees a randomly picked number between 0 and 20, except 10 (for instance "3 ECU"). The labels "player 1", "player 2", etc. are randomly assigned anew to players at the beginning of each period, making the identification of other players across periods impossible.

Then players enter a third stage. Here they can punish the other players. Each punishment point assigned to another player leads to a deduction of three ECUs from the punished player's account, but also reduces the punisher's income by one ECU. Each player can spend up to 10 ECU to punish each other player in the group. Amounts spent on punishment are deducted from the player's earnings. Thus, player i's payoff in a given period is

$$\pi_i = \max(22 - g_i + 0.4 \sum_j g_j - 3 \sum_{j \neq i} p_{j \to i} - \kappa m_i, 0) - \sum_{j \neq i} p_{i \to j}, \qquad (2)$$

with $0 \leq g_i \leq 20$ and $0 \leq m_i \leq 2$. After each period, players learn their own payoff and the points they received (but get no detailed information on who distributed points). Players then proceed to the next period; payoffs accrue over periods. All parameters, the signal technology, and payoff functions are common knowledge.

We investigate two treatment dimensions: First, we vary the marginal monitoring cost κ. In treatment *Low* we choose $\kappa = 0.2$, for treatment *High* we set $\kappa = 1$.

Second, we vary whether players receive information about the monitoring decisions of other group members. In *Private* treatments players do not learn about the monitoring decisions of the other players. In *Public* treatments we inform all players in a group about the individual monitoring decisions of the other group members before each makes her choice of g_i.

We ran a total of twelve sessions with 60 groups (240 subjects) in a 2×2 factorial design. For each of the four treatment combinations (*PublicLow*, *PublicHigh*, *PrivateLow*, and *PrivateHigh*) we have 15 independent observations (i.e., groups). Each subject participated in only one treatment condition. The experiments were conducted at the Wiso-Lab of the University of Hamburg with mostly undergraduate students from various fields. Once all subjects were seated, the written instructions were handed to them before the experimenter read them out loud (see appendix A.2). Subjects were given the opportunity to ask questions (in private). Before the experiment started subjects had to solve a set of control questions. A session lasted for about 90 minutes. Payoffs were converted at an exchange rate of 3 Euro-Cent per ECU. Subjects earned on average 20.20 Euro[5] (standard

[5]Approximately $22.70 at the time of the experiments.

deviation 3.20 Euro), including a show-up fee of five Euro.

3 Behavioral Predictions

For our behavioral predictions we focus our attention mainly on the decision to monitor. While having more accurate information about the other player's contributions could have various benefits, we assume that the main motivation for monitoring is related to punishment. We investigate two motives: (i) monitoring reduces the punishment necessary to enforce cooperation, and (ii) monitoring reduces the risk of erroneous punishment.

The first motive, that monitoring lowers the punishment necessary to enforce cooperation, is relatively straightforward to quantify. For our analysis we make the simplifying assumption that there are two types of players, (i) players with purely selfish preferences, who neither contribute nor punish, and (ii) enforcers, who are willing to contribute $\bar{g}$ even in the absence of a punishment threat and who are willing to enforce contributions of other players with punishment.[6] For simplicity we will assume perfect information about the players' types, and we limit our attention to the stage game.

Players with selfish preferences contribute zero and do not punish and, consequently, have no willingness to pay for monitoring. Enforcers may monitor because this allows them to save money in terms of punishment expenditures.

To calculate the value of monitoring, consider an enforcer who seeks to mete out deterrent punishment, that is, punishment sufficiently strong in order to make the other players indifferent between contributing and free riding (assuming selfish and risk-neutral preferences for the other players). Nicklisch et al. (2016) derive an expression for minimal deterrent punishment in the light of imperfect information:

$$p(s, \bar{g}, \lambda) = \max\left\{\frac{4(\bar{g} - s)}{(21\lambda - 1)}, 0\right\}, \tag{3}$$

where s is the signal, and $\bar{g}$ is the enforced cooperation level, and λ is the probability of receiving accurate signals, which in turn is a function of monitoring. All signals below $\bar{g}$ receive positive punishment, linearly increasing in the difference. Signals weakly above $\bar{g}$ are not punished. Deterrent punishment is increasing in $\bar{g}$, and, most importantly for our purpose, decreasing

[6] We do not directly model the enforcer's preferences. The willingness to punish could be rationalized by assuming social preferences, like distributional preferences or reciprocal preferences, but one could alternatively model norm enforcement as an innate or evolved taste. For an example of literature in which punishment types are taken as primitives over which evolutionary selection operates to determine type prevalence, see Boyd, Gintis, and Bowles (2010).

in monitoring (which increases λ).[7] That is, the more accurately punishment can be targeted, the less punishment is required for deterrence, so monitoring and punishing are substitutes from an enforcer's standpoint.

Initially, assume that an enforcer has no preference for punishment accuracy as such, and only strives to achieve enforcement with the least cost combination of the two tools at hand, monitoring and punishing. We derive an enforcer's demand for monitoring by first determining how the required level of punishment varies depending on the contribution level she wishes to enforce and on the signal accuracy associated with each level of monitoring. How much punishment is necessary to deter free riding? Given the signals, we can derive an expression for the expected punishment expenditures necessary to enforce a contribution of $\bar{g}$ by a selfish player, assuming the player actually contributes $\bar{g}$:

$$P(\bar{g}, \lambda) = \frac{2(1-\lambda)(\bar{g}^2 + \bar{g})}{21(21\lambda - 1)}. \tag{4}$$

The RHS equals zero when the signals are fully accurate ($\lambda = 1$), that is, under full monitoring the threat of punishment is sufficient and no actual punishment is necessary. Noisy signals require punishment in case the signal is below $\bar{g}$.

Expression (4) shows the expected punishment expenditures for a single enforcer facing a single potential free rider. The overall value of monitoring for an enforcer depends on the number of selfish players and enforcers in the group. If, for example, a single enforcer faces three selfish players, then the expected punishment costs are $3P(.)$. On the other hand, if three enforcers face one selfish player, then they can split the punishment equally, such that individual punishment expenditures are $\frac{1}{3}P(.)$. We denote by $r \in [\frac{1}{3}, 3]$ the ratio of selfish players per enforcer.

Figure 1 shows the demand function for monitoring for possible values of r when $\bar{g} = 20$. The marginal value or benefit to the enforcer from a given unit of monitoring is simply the amount of punishment cost she can forgo while still achieving deterrence, when substituting punishment accuracy (via monitoring) for punishment amount.[8] For lower $\bar{g}$ demand for monitoring decreases. The dashed lines indicate the cost of a unit of monitoring in the *High* ($\kappa = 1$) and *Low* ($\kappa = 0.2$) cost treatment pairs. For $r = 1$ (e.g. if a group consists of two enforcers and two selfish subjects) the model predicts one unit of monitoring by each enforcer in the *High* treatments and two

[7] More precisely, this is true for $\lambda \in (\frac{1}{21}, 1]$. $\lambda = \frac{1}{21}$ refers to an uninformative signal, for which deterrent punishment is not possible. In our experiment there are—depending on monitoring—three possible values for λ: $.5, .75, 1$.

[8] Recall again that we define the enforcer type as one who undertakes to enforce a cooperation level $\bar{g}$ from others at lowest cost. Possible valuation of accuracy as desirable in its own right or for, say, inequity aversion reasons, is at the moment assumed absent, but taken up in the subsection which begins in the paragraph below.

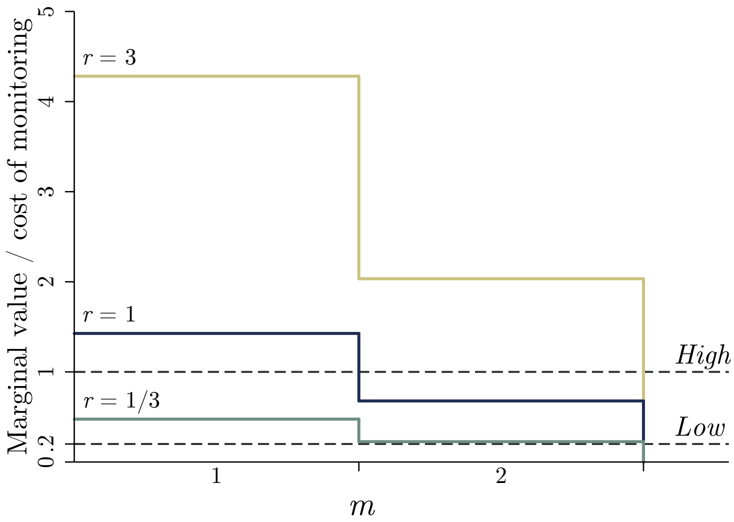


Figure 1: Demand for monitoring for $\bar{g} = 20$, and cost of monitoring in *Low* ($\kappa = 0.2$) and *High* ($\kappa = 1$).

units in the *Low* treatments. For $r = 3$ full monitoring is optimal for the enforcer in all treatments, while in case of $r = \frac{1}{3}$ the enforcers monitor only in the *Low* treatments. Importantly, the marginal value of the first unit of monitoring (from $\lambda = .5$ to $.75$) is always higher than the marginal value for the second unit, making intermediate monitoring a likely outcome for many parameter constellations given the simple enforcement motive assumed thus far.

Aside from enforcing cooperative outcomes, an alternative motive for monitoring might be that players have a genuine desire to punish free riders, be it due to reciprocal motives or an aversion against earning less than the free rider (inequality aversion, see Fehr and Schmidt, 1999). This motive is likely to be accompanied by desires to avoid two types of errors: a type I error if the enforcer punishes a player with $s_j < \bar{g}$ but $g_j \geq \bar{g}$, and a type II error if the enforcer fails to punish a player with $g_j < \bar{g}$ due to $s_j \geq \bar{g}$. Independent of the exact specification of the psychological costs associated with an erroneous punishment act, it seems plausible that the value of monitoring is linear. Let us illustrate this for inequality aversion. Under perfect information, Thöni (2014) shows that inequality averse enforcers punish free riders such that all players in the group end up with the same monetary payoff. For very weak signals the enforcer might prefer not to punish at all, but for some intermediate range she would punish the other players based on the signals and accept the fact that, with probability $1-\lambda$, she might be creating instead of reducing inequality. As an example, assume the signals a player i receives indicate that all others players in the group cooperate fully. An inequality averse player i would not punish any other player. If the signal of a player j is correct, then there is no disutility from inequality in comparison with j. However, with probability $1-\lambda$ the signal is incorrect and player i suffers a disutility of Δu from inequality towards the other player.[9] The expected disutility at the punishment stage has the form $(1-\lambda)\Delta u$, which means that it is linear in λ.[10] Consequently, the demand for monitoring should produce corner solutions where players either monitor two units or not at all (with the exception of the case where, by accident, the expected disutility per unit of monitoring equals κ). To conclude, we derived two theoretical accounts for the demand for monitoring. If the primary goal of the enforcers is to ensure a certain level of contributions via deterrent punishment, then we should observe interior solutions ($m = 1$) relatively frequently. If, on the other hand, enforcers monitor to avoid errors in punishment, then we should mostly observe either no monitoring ($m = 0$) or full monitoring ($m = 2$).

Although a preference for avoiding punishing errors is thus expected to be associated with a greater bifurcation between full and no monitoring than is a simple goal of enforcing cooperation, alone, we see no reason why the demand for monitoring would not be declining in monitoring's cost, despite preference for error avoidance. For the treatment differences between *Low* and *High*, both motives predict that monitoring should be decreasing in the monitoring cost κ.

The treatment variation along the *Private* versus *Public* dimension may yield either of two possible effects: monitoring can help to signal enforcer presence and help to coordinate punishment among enforcers in *Public* treatments (i.e., the more enforcers are known to be present, the lower the required punishment and monitoring expenditures by a given enforcer according to expression (4)). At the same time, as monitoring might be taken as a signal of a willingness to punish, monitoring may be a more effective substitute for punishing when it is public. For these reasons, publicness is likely to increase the attractiveness of monitoring. On the other hand, if publicness of monitoring is likely to create some commitment to punish noncontributors if necessary subsequently, player may decide that doing both is simply too expensive, and so their best option is to do neither one. That is, publicness may decrease monitoring as players eschew such a commitment. In addition, publicness allows players to free-ride on others' public

[9]To be sure, in some settings including our experimental design, the prospective punisher might never learn the actual inequality outcome. Our discussion assumes that the disutilities associated with probability-weighted (expected) outcomes affect decisions whether or not there will be ex post knowledge of what inequality was realized.

[10]In the Fehr and Schmidt (1999) framework we could write this anticipated disutility as $(1-\lambda)\alpha(\frac{1}{20}\sum_{g_j=0}^{19}(\pi_j(g_j)-\pi_i(g_j)))$, where $\pi_i(g_j)$ is i's payoff in case j's contribution is g_j, and $\pi_j(g_j)$ is the same for player j, and α is the disutility from disadvantageous inequality.

monitoring: that is, the publicness includes sharing such signals, entailing a collective action problem in its own right.

Notice that it seems plausible for us to assume that the relative importance of both eschewing commitments and the collective action problem interact with the cost dimension of our treatment variations. That is, with high costs of monitoring, avoiding commitments as well as the collective action problem may be more important than the signaling effect of monitoring, while with low costs of monitoring, the greater efficacy of monitoring as a substitute for punishing leads more monitoring to be done when it is public. Therefore, we hypothesize that low monitoring costs lead to more monitoring in a *Public* than in a *Private* setting, while high monitoring costs may lead to less monitoring in *Public* than in *Private* settings. Consequently, signalling and precise punishment may lead to higher contribution levels in treatment *PrivateHigh* than in treatment *PublicHigh*, whereas they may lead to lower contribution levels in treatment *PrivateLow* than in treatment *PublicLow*.

4 Results

We provide a brief overview of the results of our four experimental treatments before beginning a more detailed analysis in which our initial focus is on our main interest, the monitoring decisions. Figure 2 shows the averages of the three main dependent variables across treatment. Spikes indicate clustered standard errors. In addition, Table 1 shows the significance levels of the differences for all bilateral treatment comparisons based on Wilcoxon rank-sum tests.

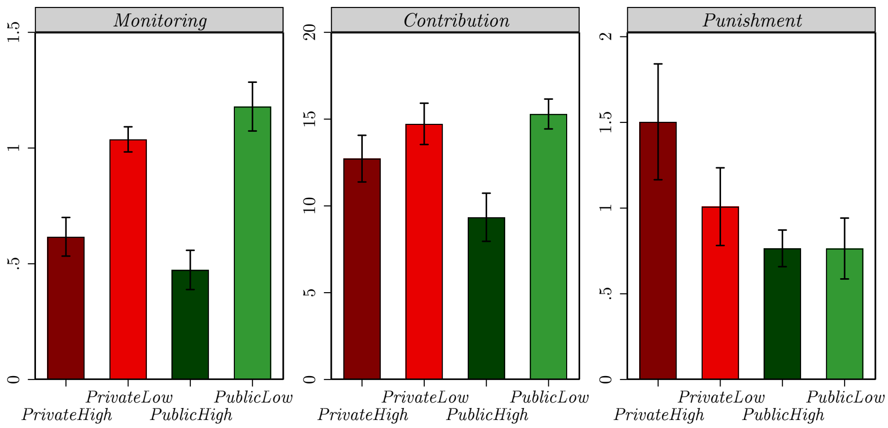


Figure 2: Averages of the main dependent variables over the 20 periods and by treatment. Spikes indicate standard errors, clustered on group.

Table 1: Bilateral treatment comparisons.

			p-values of bilateral comparison		
Variable	Treatment	Mean	*PrivateHigh*	*PrivateLow*	*PublicHigh*
Monitoring	*PrivateHigh*	0.62			
	PrivateLow	1.04	.000		
	PublicHigh	0.47	.177	.000	
	PublicLow	1.18	.000	.254	.000
Contribution	*PrivateHigh*	12.72			
	PrivateLow	14.73	.330		
	PublicHigh	9.35	.071	.007	
	PublicLow	15.30	.191	.983	.003
Punishment	*PrivateHigh*	1.50			
	PrivateLow	1.01	.290		
	PublicHigh	0.76	.191	.803	
	PublicLow	0.76	.089	.395	.604

Notes. Mean of monitoring, contributions, and punishment across the 20 periods and p-values of Wilkoxon ranksum tests for bilateral treatment comparisons. All tests based on independent group averages.

The left panel in Figure 2 shows that monitoring is strongly price sensitive: in the two treatments with low cost (*PrivateLow* and *PublicLow*) subjects monitor on average around one unit, but only around half a unit in the two treatments with high costs (Wilcoxon rank-sum test with pooled data, *High* vs. *Low*: $p = 0.000$). On the other hand, whether the monitoring decisions are made public or not does not seem to have a uniform effect. While monitoring is slightly higher in *PublicLow* than in *PrivateLow*, the reverse is the case when comparing *PublicHigh* to *PrivateHigh*. When pooled, the two treatments in which monitoring is public are not different from the two in which it is private, with respect to amount of information purchased (*Private* vs. *Public*: $p = 0.848$).

The middle panel of Figure 2 shows average contribution to the public good. In all treatments but *PublicHigh*, average contributions are well above 50% of the endowment. The stylized fact from past VCM experiments is that contributions begin in the neighborhood of 50% of the endowment, and decline with repetition in the absence of well-targeted peer punishment. That contributions average near 75% of endowment in the two treatments with low cost of information, and 64% of endowment in *PrivateHigh*, suggests that the threat and use of punishment increased and sustained cooperation in our environment with costly monitoring.

The right panel of Figure 2 shows that punishment is lower in the two public treatments compared to their counterpart with private monitoring decisions. This is consistent with the idea that a public monitoring signal serves as a warning; but the story turns out to be different as between the two public treatments, and we reserve it for later discussion. First, we will focus on the monitoring decisions and show how they function as precursors to punishment. We show that the warning effect of monitoring can affect observed punishment by rendering it unnecessary, at least in the *PublicLow* treatment.

4.1 Monitoring

We start with the structure of the demand for monitoring. Recall that our two theoretical conjectures suggested a downwards sloping individual demand function if monitoring is primarily motivated by economizing on enforcement costs, whereas if the motive is to avoid misguided punishment, then we conjectured that subjects monitor either fully or not at all. Figure 3 shows that the latter is the case: in all treatments subjects either monitor fully or not at all, while intermediate monitoring is very rare.[11] This result suggests that the value subjects attach to additional information accuracy on the margin is non-decreasing—i.e., avoiding of mis-targeted punishing plays a substantial part in the desire to monitor.

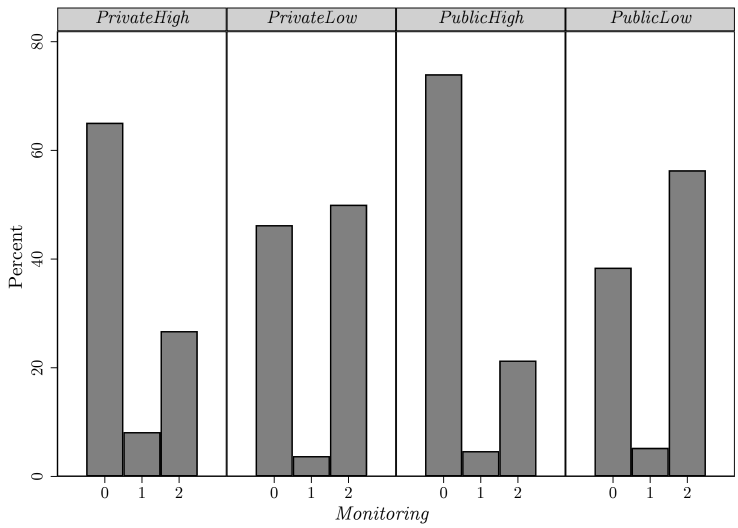


Figure 3: Histograms of the individual monitoring decisions in the four treatments.

Figure 4 shows the demand for monitoring across the 20 periods. Throughout the game the difference between high and low monitoring cost remains highly significant (we will provide some statistics on this later). On the

[11]The intermediate level of monitoring is significantly less frequent than the two other levels in all treatments ($p < .001$).

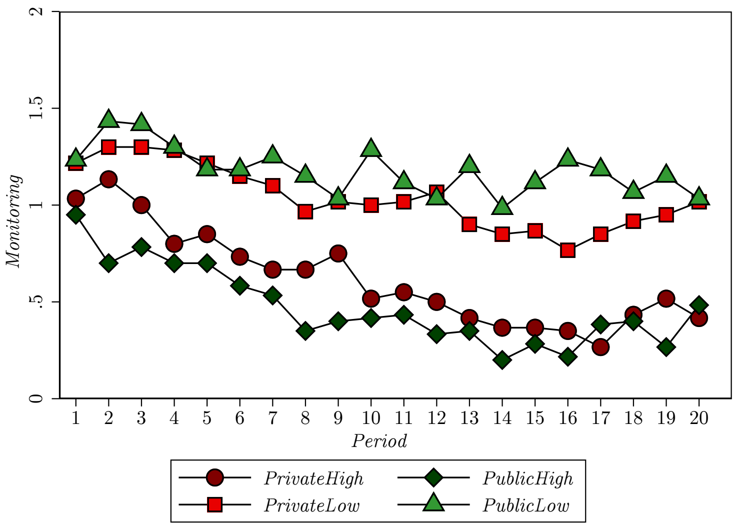


Figure 4: Monitoring over the course of the 20 periods.

other hand, the graphs for the *Private* and *Public* member of each treatment pair are never far apart and often overlap. In the last third of the game it seems like public results in more monitoring, but the differences do not reach significance. There is a negative trend in monitoring, especially in the two treatments with high cost. This might indicate that—similarly to what is often observed for punishment—monitoring is initially necessary to establish the credibility of punishment, but no longer later in the game. The negative trend is, however, weaker for monitoring than for punishment, especially for low cost of information. This indicates that continued monitoring, serving as a warning of potential punishment, may be necessary to ensure self governance.

Table 2 shows the results of OLS estimates explaining individual monitoring decisions over time. We report robust standard errors with clustering on group. Model (1) explains monitoring by period, a dummy for the final period, and two dummies for the treatment variations. We confirm the overall negative time trend, but observe a significant increase in the final period, presumably indicating that subjects anticipate the danger of end-game effects in contributions. The cost level is highly significant, while *Public* does not seem to matter. In Model (2) we add an interaction of the two treatment variables. Our point estimate for the dummy *Public* is positive for low monitoring cost, albeit insignificant. In accordance with our hypothesis, the publicness of the monitoring decision has a negative effect under high monitoring costs (weakly significant). In Model (3) we investigate whether contributions from the previous period affect monitoring. We add the subject's own contribution (g_i^{t-1}) and the average contribution of the other three group members ($\bar{g}_{-i}^{t-1}$) in the previous period as explanatory variables. The average contribution of the others does not seem to affect monitoring, while there is a strong correlation between a subject's contribution in the previous round and her monitoring, perhaps because those who contribute more also monitor more. When we replace the others' average contribution by the standard deviation of the four contributions in the previous period, in Model (4), the effect is significant, that is, subjects are more likely to monitor if the contributions were more heterogenous in the previous period. A plausible interpretation is that the presence of low contributors in the previous period encourages cooperators to monitor in preparation for potential punishing.

Table 2: OLS estimates for monitoring

	Dependent variable: *Monitoring*			
	(1)	(2)	(3)	(4)
Period	-0.028^{***}	-0.028^{***}	-0.028^{***}	-0.025^{***}
	(0.004)	(0.004)	(0.004)	(0.004)
Final period	0.186^{***}	0.186^{***}	0.193^{***}	0.186^{***}
	(0.054)	(0.054)	(0.051)	(0.050)
High	-0.563^{***}	-0.421^{***}	-0.410^{***}	-0.421^{***}
	(0.084)	(0.097)	(0.102)	(0.101)
Public	-0.001	0.142	0.142	0.136
	(0.084)	(0.116)	(0.117)	(0.121)
High × *Public*		-0.285^{*}	-0.249	-0.213
		(0.164)	(0.165)	(0.168)
g_i^{t-1}			0.018^{***}	0.018^{***}
			(0.005)	(0.005)
$\bar{g}_{-i}^{t-1}$			-0.007	
			(0.007)	
sd(g^{t-1})				0.022^{**}
				(0.011)
Constant	1.394^{***}	1.322^{***}	1.158^{***}	0.961^{***}
	(0.064)	(0.062)	(0.122)	(0.121)
F-test	28.4	23.1	25.4	24.7
Prob > F	0.000	0.000	0.000	0.000
R^2	0.112	0.117	0.130	0.133
N	4800	4800	4560	4560

Notes: OLS estimates. *High* and *Public* indicate dimensions of treatment variation; g_i^{t-1} ($\bar{g}_{-i}^{t-1}$) indicates a subject's (the others' average) contribution in the previous period; sd(g^{t-1}) denotes the standard deviation of the contributions in the previous period. Robust standard errors, clustered on group, in parentheses. * $p < .1$, ** $p < .05$, *** $p < .01$.

Result 1 *Demand for monitoring is bimodal: subjects either do not monitor or invest two units to receive fully accurate signals about the contributions. Demand for information is price sensitive, whereas* Private *and* Public *monitoring condition result in similar demand for monitoring.*

4.2 Contributions

In all four treatments contributions are, after some initial increase, fairly stable over time. Average contribution in *PublicHigh* is consistently lower than in *PrivateHigh*, while contributions in the two treatments with low monitoring cost are higher than in their *High* counterparts and show patterns very similar to one another (see Figure A1 in the appendix).

Table 3: OLS estimates for contribution

	Dependent variable: *Contribution*				
	(1) All obs.	(2) All obs.	(3) *Public*	(4) *Public*	(5) *Private*
Period	0.059	-0.031^{***}	-0.029^{**}	-0.006	-0.032^{**}
	(0.037)	(0.009)	(0.013)	(0.013)	(0.013)
Final period	-1.675^{***}	-0.720^{*}	-0.069	-0.228	-1.383^{**}
	(0.415)	(0.375)	(0.487)	(0.444)	(0.549)
High	-2.003	-0.281	-0.885^{**}	0.157	-0.213
	(1.750)	(0.219)	(0.383)	(0.525)	(0.171)
Public	0.568	0.169			
	(1.431)	(0.172)			
High × *Public*	-3.947	-0.416			
	(2.365)	(0.310)			
g_i^{t-1}		0.567^{***}	0.532^{***}	0.522^{***}	0.607^{***}
		(0.031)	(0.038)	(0.040)	(0.048)
$\bar{g}_{-i}^{t-1}$		0.358^{***}	0.361^{***}	0.350^{***}	0.350^{***}
		(0.032)	(0.050)	(0.049)	(0.046)
$p_{\to i}^{t-1}$		0.234^{***}	0.214^{*}	0.149	0.278^{**}
		(0.082)	(0.126)	(0.128)	(0.110)
$p_{\to i}^{t-1} \times (g_i^{t-1} \geq \bar{g}^{t-1})$		-0.227^{*}	-0.192	-0.141	-0.289^{*}
		(0.117)	(0.171)	(0.176)	(0.153)
$\bar{m}_{-i}$				1.394^{***}	
				(0.397)	
High × $\bar{m}_{-i}$				-0.327	
				(0.508)	
Constant	14.192^{***}	1.426^{***}	2.036^{***}	0.511	0.968^{**}
	(1.189)	(0.323)	(0.516)	(0.526)	(0.376)
F-test	7.2	689.8	437.0	482.9	729.9
Prob > F	0.000	0.000	0.000	0.000	0.000
R^2	0.117	0.636	0.638	0.648	0.628
N	4800	4560	2280	2280	2280

Notes: OLS estimates. *High* and *Public* indicate dimensions of treatment variation; g_i^{t-1} ($\bar{g}_{-i}^{t-1}$) indicates a subject's (the others' average) contribution in the previous period; $p_{\to i}^{t-1}$ indicates punishment i received by others in the previous period (with $p_{\to i} = \sum_{j \neq i} p_{j \to i}$); $\bar{m}_{-i}$ denotes the average monitoring of the other subjects. In Model (1) and (2) we use all observations, in Models (3) and (4) we use only the observations from the *Public* treatments, in Model(5) only the observations from the *Private* treatments, respectively. Robust standard errors, clustered on group, in parentheses. * $p < .1$, ** $p < .05$, *** $p < .01$.

Table 3 reports OLS regression estimates for the contribution decision. Model (1) shows that high monitoring costs do not significantly affect contributions in the *Private* treatments, whereas in the *Public* treatments the joint effect is significantly negative (Wald test for joint significance of *High* and *High* $\times$ *Public*: $p = 0.000$). The overall time trend is insignificant, but there is a significant end-game effect in contributions. In Model (2) we add lagged explanatory variables from the previous round. We control for a subject's contribution (g_i^{t-1}), the other group members' average contribution ($\bar{g}_{-i}^{t-1}$), and the number of punishment points received from other subjects ($p_{\rightarrow i}^{t-1}$).[12] In addition, we interact the variable for received punishment with a dummy for whether the subject's contribution was higher than the average contribution in the group. Lagged contributions as well as received punishment significantly increase contributions, unless punishment is received in combination with above average contributions, in which case the joint coefficient is insignificant (Wald test for joint significance of $p_{\rightarrow i}^{t-1}$ and the interaction: $p = 0.806$). In Model (3) we look at the results of the *Public* treatments only. We confirm that high monitoring costs significantly reduce contributions, while the reaction to punishment and contributions in the previous periods is qualitatively similar to the estimates on the whole sample. In Model (4) we add the average monitoring by the other subjects in the group ($\bar{m}_{-i}$) and its interaction with high monitoring costs. The coefficient on $\bar{m}_{-i}$ is highly significant and positive, indicating that monitoring works as a signal that improves others' cooperation, when public. The interaction with the dummy for high monitoring costs is small and insignificant, suggesting that this effect is present in both treatments with public monitoring. The estimated coefficients on the contribution variables remain about as large and significant as in Model (3), but the punishment measures lose significance, suggesting that the warning effect of monitoring delivers in advance much of the news that then comes with punishment. In Model (5) we show the results for the *Private* treatments. The results are very similar to those of Model (2).

Result 2 *In* Private *treatments monitoring costs do not seem to influence contributions importantly. On the other hand, when subjects learn about others' monitoring decisions (i.e., in* Public *treatments), contributions significantly increase in monitoring such that low monitoring costs increase contributions.*

4.3 Punishment

As usual in public goods experiments with punishment, we observe the use of punishment to decline over time. This is true for all treatments, and the ordering of the treatments shown in Figure 2 remains relatively stable throughout the 20 periods (see Figure A2 in the appendix). This is partly

[12] Model (2) shows a significant negative coefficient for period. This should not be interpreted as pure time trend, as we control for lagged contributions. In particular, the coefficient loses significance if we remove $\bar{g}_{-i}^{t-1}$.

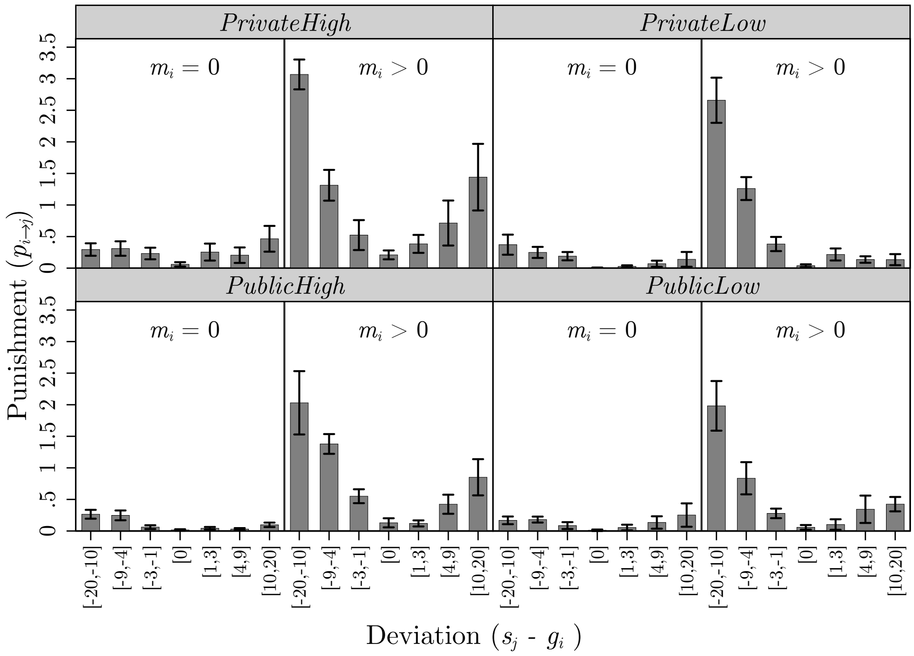


Figure 5: Assigned punishment in relation to the difference between the signal and the punisher's contribution ($s_j - g_i$). Averages over all cases (including zero punishment). Separate panels show the punishments in case the punisher did not monitor ($m_i = 0$), or monitored one or two units ($m_i > 0$).

due to some groups reaching full contribution after some initial rounds of punishment. It is also noteworthy that the expenses for punishment drop more sharply than the expenses for monitoring. If we make a simple comparison of the first and second half of the periods, then punishment drops from 1.35 to 0.67 (reduction by 50%), while the reduction in monitoring is from 0.96 to 0.70 (reduction by 27%). Both reductions are highly significant ($p < .001$, Wilcoxon matched-pairs signed-ranks test).

There is a strong correlation between monitoring and the use of punishment. Subjects who monitored punish more often than uninformed subjects. Among the subjects who did not monitor in a given period, 14.3 percent punish another subject in that period. The corresponding percentage is 44.7 percent among the subjects who monitored.[13] Figure 5 shows how the strength of punishment varies with the difference between the signal and the punisher's contribution ($d_{ij} = s_j - g_i$), by treatment. We distinguish between situations in which the punisher monitored ($m_i > 0$), and ones in

[13]Correlating individual frequency of punishment to the individual frequency of monitoring yields $\rho = .424$ with $p = .000$

Table 4: OLS estimates for assigned punishment ($p_{i \to j}$)

	Dependent variable: Punishment			
	(1) All obs.	(2) All obs.	(3) *High*	(4) *Low*
Period	-0.011^{***}	-0.004	-0.009^{*}	0.000
	(0.003)	(0.003)	(0.004)	(0.003)
Final period	0.035	-0.020	-0.032	-0.003
	(0.038)	(0.033)	(0.051)	(0.050)
High	0.228^{*}	0.193^{*}		
	(0.119)	(0.111)		
Public	-0.094	-0.044	-0.170^{*}	-0.044
	(0.084)	(0.074)	(0.095)	(0.075)
Public $\times$ *High*	-0.071	-0.137		
	(0.135)	(0.123)		
d_{ij}^{-}	0.073^{***}	0.014^{***}	0.013^{**}	0.014^{**}
	(0.008)	(0.004)	(0.005)	(0.006)
d_{ij}^{+}	0.023^{***}	0.009^{**}	0.010^{*}	0.008
	(0.005)	(0.004)	(0.005)	(0.007)
m_i	0.264^{***}	0.003	-0.009	-0.009
	(0.028)	(0.019)	(0.042)	(0.010)
$m_i \times d_{ij}^{-}$		0.090^{***}	0.100^{***}	0.083^{***}
		(0.007)	(0.010)	(0.009)
$m_i \times d_{ij}^{+}$		0.013^{***}	0.030^{***}	0.005
		(0.005)	(0.010)	(0.005)
Constant	-0.076	0.103^{*}	0.308^{**}	0.092
	(0.060)	(0.059)	(0.119)	(0.056)
F-test	22.0	36.2	37.6	17.3
Prob > F	0.000	0.000	0.000	0.000
R^2	0.143	0.264	0.258	0.283
N	14400	14400	7200	7200

Notes: OLS estimates. *High*, *Low*, and *Public* indicate dimensions of treatment variation; d_{ij} indicates the deviation between a signal and a punisher's contribution. It is calculated as $d_{ij} = s_j - g_i$, and $d_{ij}^{+} = \max\{d_{ij}, 0\}$, $d_{ij}^{-} = |\min\{d_{ij}, 0\}|$; m_i indicates monitoring. Robust standard errors, clustered on group, in parentheses. * $p < .1$, ** $p < .05$, *** $p < .01$.

which she did not monitor ($m_i = 0$). (We pool all non-zero investments in information, because there are only a few observations for $m = 1$.) Punishment is much stronger among informed subjects, and in both panels we observe non-negligible amounts of antisocial punishment.[14] Interestingly, it seems that the treatments with higher monitoring cost provoke stronger punishment of high contributors, much of it by informed subjects.

Table 4 shows OLS estimates for the punishment decision. We control for the two treatment dimensions and for the deviation between the signal and a punisher's contribution. We estimate separate slopes for positive (d^+_{ij}) and negative deviations (d^-_{ij}).[15] In addition, we control for monitoring and periods. Punishment drops over time, but there is no additional end-game effect in the final period. We observe marginally significantly higher punishments when monitoring costs are high. The slopes are highly significant both for signals below the punisher's contribution and for signals above the punisher's contribution. In Model (2) we interact the deviation with monitoring and find, in accordance to Figure 5, the reaction to the signal is a lot stronger among subjects who monitored, and this holds equally for ostensibly antisocial punishment (i.e., monitoring does not, as some might conjecture, cause less such punishing among those informed that the signal is accurate). In Model (3) and (4) we re-estimate the model for the *High* and *Low* treatments separately. We find that the coefficient for antisocial punishment (d^+) reaches weak significance only when monitoring costs are high. Likewise, the interaction effect with monitoring is highly significant when monitoring costs are high, but not when they are low. The coefficients for the interaction effects with monitoring confirm the results from Figure 5 that punishment of high contributors is particularly strong in the treatments with high monitoring costs and for punishers who monitor. High monitoring costs may either create some sort of commitment to punish, or screen out those less determined to punish strongly. Alternatively it might be that in the treatments with high monitoring costs the groups no longer share a common cooperative norm. With increased free riding low contributors might not be willing to accept punishment as they might feel that they have a moral right to free ride (they are not violating a norm). They may feel that high contributors violate a norm by punishing low contributors, therefore such a low contributor has a right to punish a high contributor in the hope of punishing back ('blind revenge').

[14] Here, we define antisocial punishment as punishment targeted at subjects with a weakly higher contribution than the punisher (Herrmann, Thöni, & Gächter, 2008). However, we substitute signal of potential target s_j for j's actual contribution g_j, and in case of imperfect information ($m_i < 2$) we cannot be certain that the punishment of a high signal member is due to antisocial motives, since we cannot rule out belief that the signal is not accurate.

[15] Note that both d^+ and d^- terms are increasing in the size of the difference, assured by the absolute value operator (see Table 4 table note). d^- (d^+) > 0 when prospective punishment receiver j is reported to contribute less (more) than punisher i.

Table 5: OLS estimates for received punishment ($p_{\to i}$)

	Dependent variable: Received punishment			
	(1)	(2)	(3)	(4)
	All obs.	All obs.	*Public*	*Public*
Period	-0.070^{***}	-0.050^{***}	-0.059^{***}	-0.061^{***}
	(0.011)	(0.011)	(0.013)	(0.013)
Final period	0.429^{***}	0.223	0.266^{*}	0.270^{*}
	(0.130)	(0.133)	(0.147)	(0.146)
High	0.495	0.639^{*}	0.050	0.055
	(0.397)	(0.363)	(0.227)	(0.230)
Public	-0.244	-0.299		
	(0.281)	(0.268)		
High × *Public*	-0.494	-0.541		
	(0.445)	(0.431)		
$\bar{g}_{-i}$		0.107^{***}	0.104^{***}	0.105^{***}
		(0.017)	(0.023)	(0.023)
g_i		-0.163^{***}	-0.133^{***}	-0.143^{***}
		(0.018)	(0.022)	(0.024)
$\bar{m}_{-i}$		0.607^{***}	0.453^{***}	0.466^{***}
		(0.112)	(0.152)	(0.147)
m_i			-0.143^{***}	-0.369^{**}
			(0.051)	(0.178)
$m_i \times g_i$				0.016
				(0.011)
Constant	1.725^{***}	1.713^{***}	1.443^{***}	1.543^{***}
	(0.251)	(0.393)	(0.356)	(0.372)
F-test	10.0	21.6	9.3	7.8
Prob > F	0.000	0.000	0.000	0.000
R^2	0.054	0.235	0.205	0.208
N	4800	4800	2400	2400

Notes: OLS estimates. Dependent variable is the sum of received punishment points, i.e., $p_{\to i} = \sum_{j \neq i} p_{j \to i}$; *High* and *Public* indicate dimensions of treatment variation; $\bar{m}_{-i}$ denotes the average monitoring of the other subjects; g_i ($\bar{g}_{-i}$) indicates a subject's (the others' average) contribution; m_i denotes the subject's monitoring. Robust standard errors, clustered on group, in parentheses. * $p < .1$, ** $p < .05$, *** $p < .01$.

In Table 5 we show the results of OLS estimates for punishment received, at receiver and period level. Each observation combines the punishment a recipient j receives from up to three group members (i, k, l), yielding one third the observation numbers of Table 4. Model (1) uses the dimensions of treatment variation and period as explanatory variables. Apart from a negative trend over time and the final period term, none of the variables are significant.[16]

In Model (2) we add measures for contribution and the monitoring of the other subjects in the group. All coefficients are highly significant and have the expected sign: received punishment is significantly increasing in others' monitoring and contributions, and received punishment is strongly decreasing in the subject's contribution. In Model (3) we restrict the sample to the *Public* treatments, and we add the subject's own monitoring (m_i) to the model, in order to test whether monitoring might be seen by the group as a punishment threat and motivate retaliatory punishment acts. The results suggest that this is not the case. It is in fact the opposite: the negative coefficient suggests, remarkably, that the group on balance punishes a player for not monitoring (i.e. it punishes less those who monitor). Finally, in Model (4) we test for interaction effects between the recipient's contribution and monitoring. The coefficient is close to zero and insignificant.

Result 3 *Punishment is predominantly assigned by informed subjects, while the likelihood of receiving punishment increases when other subjects monitor more. High costs for information increase the likelihood of antisocial punishment. When monitoring is public we observe that those who monitor are punished less by the others.*

4.4 Earnings

Finally, we compare the overall efficiency of the treatments in Figure 6. The vertical axis shows the percentage of the potential efficiency gains from cooperation that subjects realize in the four treatments.[17] We find *PublicLow* to be the most efficient treatment condition and *PrivateHigh* the least efficient. The intermediate treatments suggest that the lion's share of the variation

[16] The control for final period checks for an end-game decline, which might reveal that punishment is mainly strategically motivated (Falk, Fehr, & Fischbacher, 2005). The significant increase in punishment in the final period is likely to be caused by subjects free riding towards the end of the game. Model (2) shows that the coefficient for the final period loses significance once we control for contributions. However, the results still favor the view that punishment is mainly non-strategic, as we do not observe a decline in punishment in the final period.

[17] That is, the proportion of the potential gains from moving from the predicted selfish Nash equilibrium under standard assumptions, where average earnings are 22 (no monitoring, zero contribution and punishment) to the social optimum, where average earnings are 34 (with full contribution, no monitoring and no punishment). Note that values below zero are possible due to punishment.

in efficiency gains is attributable to the variation in the cost of monitoring, while the publicness of monitoring seems of minor importance.

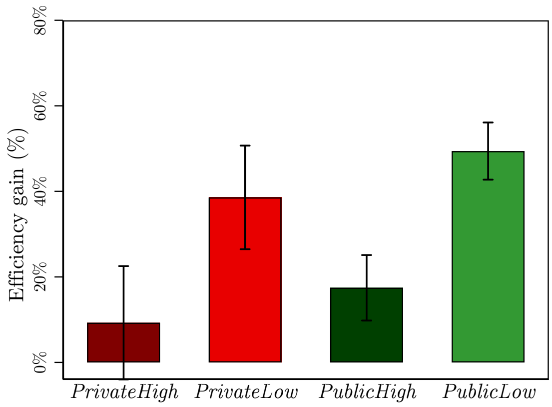


Figure 6: Average efficiency gain in the four treatments, all periods. Efficiency gain is calculated as the percentage of distance between Nash equilibrium earnings under standard assumptions (22) and maximum possible average earnings (34).

Table 6 corroborates these findings in a regression analysis based on independent group averages across the 20 periods. Dependent variable is the efficiency gain as defined above. Model (1) shows that efficiency gains are significantly higher in the treatments with low monitoring costs, whereas the publicness of monitoring remains insignificant. In Model (2) we add the interaction term of the treatment dummies. The coefficient for *High* is roughly the same as in Model (1) but loses significance.

Is monitoring efficient? Within groups, it is likely that those who monitor end up with relatively low monetary payoffs, both because of the direct monitoring cost and because, as we have shown earlier, punishment is predominantly meted out by subjects who monitor.[18]

However, on the group level, it might be that groups with higher levels of monitoring achieve more efficient outcomes than those with less monitoring. Model (3) in Table 6 shows that this is the case: fully monitoring groups ($\bar{m} = 2$) earn 44 percentage points more than groups with no monitoring, when measuring earnings relative to the two theoretical benchmarks, i.e., the Nash equilibrium and the efficient outcome. In Model (4) we control for average punishment. The effect of monitoring remains strong, while punishment has a negative impact on earnings. Apart from the direct costs of punishment, this is because highly cooperative groups often need hardly

[18]We find a significant correlation between a subject's monitoring and rank in the payoff distribution within the group ($\rho = -0.269$, $p = .000$)

any punishment, while uncooperative groups need a lot of punishment to establish cooperation.[19] Finally, in Model (5) we show that the effects of monitoring and punishment remain significant when we add dummies for the two treatment dimensions and their interaction. Notice that, judging by the R^2, the percentage of variance these regressions explain jumps—when punishment is added—from under 20% in columns (1) to (3) to over 70% in columns (4) and (5), with the treatment dimension dummies adding about 3% to the already high R^2 of column (4).

Table 6: Group estimates: Efficiency and monitoring

	Dependent variable: Efficiency gain				
	(1)	(2)	(3)	(4)	(5)
High	-0.306^{***}	-0.293			-0.018
	(0.102)	(0.179)			(0.089)
Public	0.095	0.108			-0.014
	(0.102)	(0.138)			(0.069)
High × *Public*		-0.027			-0.162
		(0.206)			(0.124)
Monitoring			0.223^{**}	0.288^{***}	0.197^{**}
			(0.097)	(0.081)	(0.097)
Punishment				-0.379^{***}	-0.388^{***}
				(0.032)	(0.037)
Constant	0.393^{***}	0.386^{***}	0.102	0.432^{***}	0.573^{***}
	(0.099)	(0.121)	(0.096)	(0.085)	(0.117)
F-test	4.8	4.5	5.3	75.4	36.2
Prob > F	0.012	0.007	0.025	0.000	0.000
R^2	0.147	0.148	0.052	0.722	0.750
N	60	60	60	60	60

Notes: OLS estimates. Dependent variable is the efficiency gain, ranging from zero (earnings in the Nash equilibrium of 22) to one (maximum average group earnings of 34). Independent variables are treatment dummies, monitoring and punishment. All estimates based on independent group averages. Robust standard errors in parentheses. * $p < .1$, ** $p < .05$, *** $p < .01$.

Result 4 *The cost of monitoring significantly affects efficiency, while the public vs. private variation does not. Average earnings in groups with high levels of monitoring are substantially higher than in groups with low levels of monitoring.*

[19] A negative relation between the average use of punishment in a group and earnings is very common in public goods games with punishment. For example, in similar regressions using the data of Herrmann et al. (2008) one finds a significantly negative effect in all of the 16 subject pools.

5 Conclusion

Recent experimental studies have demonstrated a number of important challenges for the "self-governance" ability of societies. One such challenge is that there may well be imperfect information regarding others' behaviors, and subjects' willingness to impose sanctions despite grounds for doubt may not auger well for societal self-organization (Grechenig et al., 2010 and Ambrus and Greiner, 2012). In the absence of affordable remedies, noise sharply increases "misdirected" punishment and eliminates or even reverses the effects on cooperation of peer punishment opportunities that, with perfect information, often successfully promote it.

The good news of our study is that the majority of our experimental participants are willing to pay a modest cost to monitor, while little punishment takes place with imperfect information when perfect information is affordably available. Moreover, although we offer the chance to buy a partial improvement in information at lower price, almost no subjects take up this offer. Rather subjects choose to monitor fully or not to monitor at all. Those who choose not to monitor also choose not to punish in the large majority of instances. In other words, not to monitor serves as some kind of commitment device not to execute punishment. In turn, we observe significantly more antisocial punishment when monitoring costs are high: spending more on improved signals may serve as a commitment device, also for antisocial punishment. While the condition with public monitoring decisions offers the opportunity to signal strategically one's willingness to punish noncooperators, participants do not monitor significantly more than when the monitoring decisions are kept private. A reason for the absence of a treatment effect could be that public information may motivate some subjects to free ride on other subjects expenses in monitoring.

Overall, there is a substantial and statistically significant improvement in terms of efficiency when participants monitor. Our subjects do not punish despite reasonable doubts, but try to resolve those doubts. However, the demand for monitoring is price-sensitive. This implies that it is in the best interest of a group to make relevant information about the contribution of each member be available to the others at moderate costs, if possible. However, the costs of monitoring or verification are to some degree exogenously given, and there are in some cases economies of scale in observation or advantages of having an entity with the power to insist on information disclosure. The relative ease of decentralized versus centralized monitoring may therefore be a major determinant of which social dilemmas are resolved via decentralized and which via centralized mechanisms.

Acknowledgements: Financial support for Nicklisch by the German Research Foundation (Ni 1610/1-1) and the Swiss National Science Foundation (100019E_178317/1) is gratefully acknowledged.

References

Alchian, A. A. & Demsetz, H. (1972). Production, information costs, and economic organization. *American Economic Review*, *62*(5), 777–795.

Alm, J. (2012). Measuring, explaining, and controlling tax evasion: Lessons from theory, experiments, and field studies. *International Tax and Public Finance*, *19*(1), 54–77. doi:10.1007/s10797-011-9171-2

Ambrus, A. & Greiner, B. (2012). Imperfect public monitoring with costly punishment: An experimental study. *American Economic Review*, *102*(7), 3317–3332. doi:10.1257/aer.102.7.3317

Aoyagi, M., Bhaskar, V., & Fréchette, G. R. (2016). *The impact of monitoring in infinitely repeated games: Perfect, public, and private.*

Balafoutas, L., Kocher, M. G., Putterman, L., & Sutter, M. (2013). Equality, equity and incentives: An experiment. *European Economic Review*, *60*, 32–51. doi:10.1016/j.euroecorev.2013.01.005

Bowles, S. & Gintis, H. (1990). Contested exchange: New microfoundations for the political economy of capitalism. *Politics & Society*, *18*(2), 165–222. doi:10.1177/003232929001800202

Boyd, R., Gintis, H., & Bowles, S. (2010). Coordinated punishment of defectors sustains cooperation and can proliferate when rare. *Science*, *328*(5978), 617–620. doi:10.1126/science.1183665

Burdín, G. (2016). Equality under threat by the talented: Evidence from worker-managed firms. *Economic Journal*, *126*(594), 1372–1403. doi:10.1111/ecoj.12272

Craig, B. & Pencavel, J. (1995). Participation and productivity: A comparison of worker cooperatives and conventional firms in the plywood industry. *Brookings Papers on Economic Activity. Microeconomics*, *1995*, 121–174. doi:10.2307/2534773

DeAngelo, G. & Gee, L. K. (2017). *Peers or police? detection and sanctions in the provision of public goods.* Mimeo.

Dow, G. K. (2017). *The labor-managed firm.* Cambridge University Press.

Falk, A., Fehr, E., & Fischbacher, U. (2005). Driving forces behind informal sanctions. *Econometrica*, *73*(6), 2017–2030. doi:10.1111/j.1468-0262.2005.00644.x

Fehr, E. & Gächter, S. (2000). Cooperation and punishment in public goods experiments. *American Economic Review*, *90*(4), 980–994. doi:10.1257/aer.90.4.980

Fehr, E. & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics*, *114*(3), 817–868. doi:10.1162/003355399556151

Gächter, S., Renner, E., & Sefton, M. (2008). The long-run benefits of punishment. *Science*, *322*(5907), 1510–1510. doi:10.1126/science.1164744

Grechenig, K., Nicklisch, A., & Thöni, C. (2010). Punishment despite reasonable doubt—a public goods experiment with sanctions under un-

certainty. *Journal of Empirical Legal Studies*, *7*(4), 847–867. doi:10.1111/j.1740-1461.2010.01197.x
Grosse, S., Putterman, L., & Rockenbach, B. (2011). Monitoring in teams: Using laboratory experiments to study a theory of the firm. *Journal of the European Economic Association*, *9*(4), 785–816. doi:10.1111/j.1542-4774.2011.01026.x
Herrmann, B., Thöni, C., & Gächter, S. (2008). Antisocial punishment across societies. *Science*, *319*(5868), 1362–1367. doi:10.1126/science.1153808
Kamei, K. & Putterman, L. (2017). Play it again: Partner choice, reputation building and learning from finitely repeated dilemma games. *Economic Journal*, *127*(602), 1069–1095. doi:10.1111/ecoj.12320
Kremer, M. (1997). *Why are worker cooperatives so rare?* NBER Working Paper Series. doi:10.3386/w6118
Markussen, T., Putterman, L., & Wang, L. (2017). Governing collective action in the face of observational error. *Brown University Department of Economics Working Paper 2017-2.*
Nicklisch, A., Grechenig, K., & Thöni, C. (2016). Information-sensitive Leviathans. *Journal of Public Economics*, *144*, 1–13. doi:10.1016/j.jpubeco.2016.09.008
Ostrom, E. (1990). *Governing the commons: The evolution of institutions for collective action.* Cambridge University Press.
Ostrom, E., Walker, J., & Gardner, R. (1992). Covenants with and without a sword: Self-governance is possible. *American Political Science Review*, *86*(2), 404–417. doi:10.2307/1964229
Page, T., Putterman, L., & Unel, B. (2005). Voluntary association in public goods experiments: Reciprocity, mimicry and efficiency. *Economic Journal*, *115*(506), 1032–1053. doi:10.1111/j.1468-0297.2005.01031.x
Pencavel, J. (2013). Worker cooperatives and democratic governance. In A. Grandori (Ed.), *Handbook of economic organization: Integrating economic and organization theory.* Cheltenham,UK: Edward Elgar.
Ramalingam, A., Godoy, S., Morales, A. J., & Walker, J. M. (2016). An individualistic approach to institution formation in public good games. *Journal of Economic Behavior & Organization*, *129*, 18–36.
Shapiro, C. & Stiglitz, J. E. (1984). Equilibrium unemployment as a worker descipline device. *American Economic Review*, *74*(3), 433–444.
Thöni, C. (2014). Inequality aversion and antisocial punishment. *Theory and Decision*, *76*(4), 529–545. doi:10.1007/s11238-013-9382-3
Torgler, B. (2002). Speaking to theorists and searching for facts: Tax morale and tax compliance in experiments. *Journal of Economic Surveys*, *16*(5), 657–683. doi:10.1111/1467-6419.00185
van der Heijden, E., Potters, J., & Sefton, M. (2009). Hierarchy and opportunism in teams. *Journal of Economic Behavior & Organization*, *69*(1), 39–50. doi:10.1016/j.jebo.2008.09.007

A Online appendix

A.1 Additional analysis

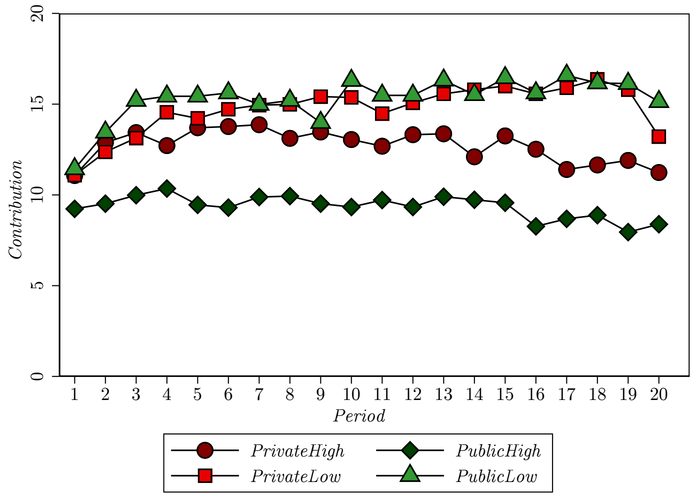


Figure A1: Monitoring over the course of the 20 periods.

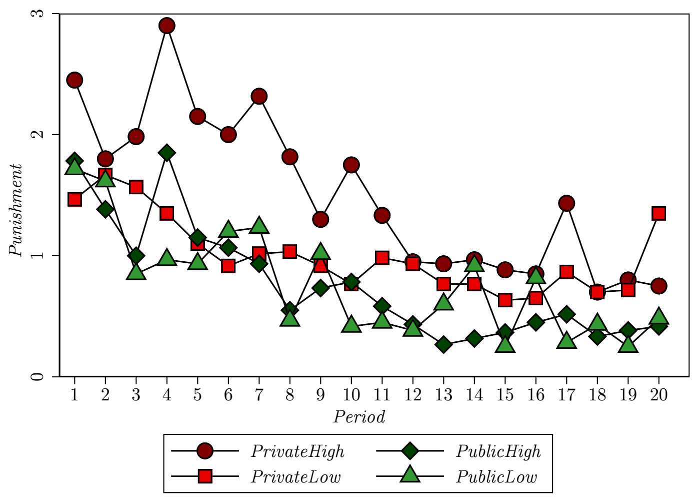


Figure A2: Punishment over the course of the 20 periods.

Table A1: OLS estimates for assigned punishment ($p_{i \to j}$)

	Dependent variable: Punishment			
	(1) All obs.	(2) All obs.	(3) *High*	(4) *Low*
Period	-0.011^{***}	-0.011^{***}	-0.004	-0.004
	(0.003)	(0.003)	(0.003)	(0.003)
Final period	0.035	0.039	-0.021	-0.017
	(0.038)	(0.039)	(0.033)	(0.036)
High	0.193^{***}	0.205^{***}	0.127^{*}	0.053
	(0.068)	(0.062)	(0.065)	(0.053)
Private	0.130^{*}	0.133^{*}	0.113^{*}	0.107^{*}
	(0.066)	(0.067)	(0.061)	(0.061)
d_{ij}^{-}	-0.073^{***}	-0.078^{***}	-0.015^{***}	-0.015^{**}
	(0.007)	(0.011)	(0.004)	(0.006)
d_{ij}^{+}	0.023^{***}	0.016^{***}	0.008^{**}	0.007
	(0.005)	(0.005)	(0.004)	(0.007)
m_i	0.266^{***}	0.269^{***}	0.006	-0.006
	(0.028)	(0.029)	(0.019)	(0.018)
$d_{ij}^{-} \times$ *High*		0.011		0.001
		(0.015)		(0.007)
$d_{ij}^{+} \times$ *High*		0.010		0.003
		(0.010)		(0.009)
$m_i \times d_{ij}^{-}$			-0.089^{***}	-0.082^{***}
			(0.007)	(0.009)
$m_i \times d_{ij}^{+}$			0.014^{***}	0.005
			(0.005)	(0.005)
$m_i \times d_{ij}^{-} \times$ *High*				-0.019
				(0.012)
$m_i \times d_{ij}^{+} \times$ *High*				0.026^{**}
				(0.012)
Constant	-0.190^{***}	-0.195^{***}	0.019	0.064
	(0.067)	(0.068)	(0.059)	(0.055)
F-test	25.1	20.3	39.6	37.6
Prob $> F$	0.000	0.000	0.000	0.000
R^2	0.143	0.144	0.263	0.268
N	14400	14400	14400	14400

Notes: OLS estimates. *High*, *Low*, and *Private* indicate treatment variations; d_{ij} indicates the deviation between a signal and a punisher's contribution. It is calculated as $d_{ij} = s_j - g_i$, and $d_{ij}^{+} = \max\{d_{ij}, 0\}$, $d_{ij}^{-} = \min\{d_{ij}, 0\}$; m_i indicates monitoring. Robust standard errors, clustered on group, in parentheses. * $p < .1$, ** $p < .05$, *** $p < .01$.

A.2 Instructions

Instructions[1]

General explanations for participants

You are taking part in an economic science experiment. If you read the following explanations carefully, you can earn a rather significant sum of money, depending on the decisions you make. It is therefore very important that you pay attention to the following points.

The instructions you have received from us are intended solely for your information. *During the experiment, you will not be allowed to communicate with anyone***.** Should you have any questions, please direct them directly to us. Not abiding by this rule will lead to exclusion from the experiment and from any payments.

In this experiment, we calculate in Taler, rather than in Euro. Your entire income will therefore initially be calculated in Taler. The total sum of Taler will later be calculated in Euro as follows:

10 Taler = 25 Euro cents

The euro you will have accrued plus 5.00 Euro for your participation will be paid to you privately *in cash* at the end of the experiment.

The experiment is divided into separate periods. It consists of a total of 20 periods. Participants are randomly assigned to groups of four. You will thus be in a group that has three other members, apart from you. During these 20 periods, the composition of your group will remain unaltered. *That is, you will be in the same group for 20 periods.* Please note that the identification number assigned to you and the other members of the group changes randomly in each period. Therefore, given group members cannot be identified from one period to the next.

The following pages outline the exact procedure of the experiment.

[1]English translation of the German original for High. Numbers change in Low accordingly. Treatment differences between Private and Public are indicated by {...} and […].

Information on the exact procedure of the experiment

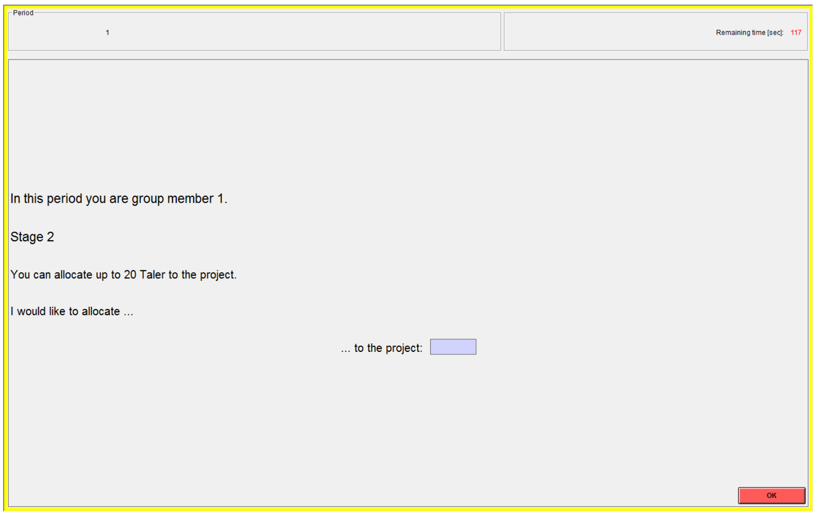
Each of the twenty periods contains three stages. In the first stage, you decide whether to invest in information about others' behaviors, information that will be made available to you in the third stage. [Prior to the second stage, you will be informed about what others in your group decided, with respect to information acquisition.] In the second stage, you will make a decision on how much to allocate to a group project. In the third stage, you will make decisions regarding possibly reducing others' incomes by assigning reduction points to them. When making your third stage decisions, you are given information about other group members' allocations from the second stage. That information may or may not be accurate. The accuracy of the information increases the more you invest in the first stage.

Stage 1

In each period, each participant is allotted *22 Taler,* which we shall henceforth refer to as his *endowment.* From this endowment, each participant can invest 0, 1, or 2 Taler for information acquisition. Taler spent on information acquisition are deduced from participant's endowment. [Prior to Stage 2, each participant is informed about the investments for information acquisition of the other group members.] {Group members are not informed about one another's investments for information acquisition.}

Stage 2

In Stage 2, each individual has to make a decision with regard to using parts of his or her endowment. You have to decide how many of 20 Taler of your endowment you wish to allocate to a *project* and how many you wish to keep for yourself. You will see the following input screen of Stage 2:



The input screen of stage 2

In the left upper corner of the screen you will find the *period number*. In the right upper corner you will find the *remaining time* for your decision *in seconds*.

You may allocate at most 20 Taler to the project. You make a decision on your project allocation by typing any whole number between 0 and 20 into the appropriate field on your screen. Together your decision in stage one and stage two determine how many Taler from your endowment you keep for yourself, i.e., *22 – Your investment for information acquisition – Your allocation to the project*. Once you have done this, your decision for this period is irreversible.

Your income at the end of stage two consists of two parts, namely:

(1) the Taler you have kept for yourself

(2) the *income gained from the project".* Your income from the project is calculated as follows: *Income from the project = .4 * sum of all contributions to the project*

Your *income in Stage 2* of each period equals:

22 (endowment)

– Your investment for information acquisition

– Your allocation to the project

+.4(sum of allocations to the project)*

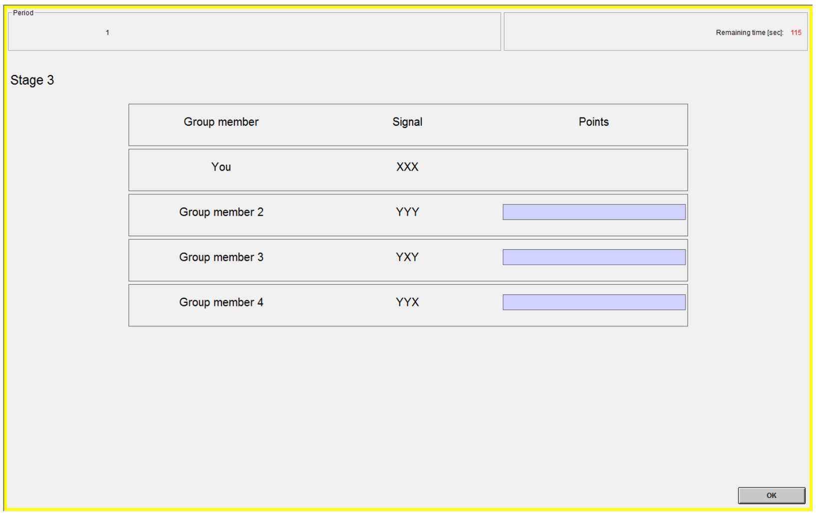
The allocations to the project are summed over the four group members, including yourself, and the total income in Stage 2, in Taler, is calculated using the same formula for each member of the group. If, for example, the sum of the allocations from all group members adds up to 60 Taler, you and all other members each receive a project income of .4 x 60 = 24 Taler. If the group members have allocated a total of 9 Taler to the project, you and all other members each receive an income of .4 x 9 = 3.6 Taler from the project.

For each Taler you keep for yourself, you earn an income of 1 Taler. If, on the other hand, you allocate one Taler from your endowment to your group's project, the sum of the allocations to the project increases by one Taler and your income from the project increases by .4 x 1 = .4 Taler. However, the income of each individual group member also increases by .4 Taler, so that the group's total income increases by .4 x 4 = 1.6 Taler. The other group members thereby also profit from your allocation to the project. In turn, you profit from other members' allocations to the project. For each Taler allocated to the project by another group member, you earn .4 x 1 = .4 Taler.

Stage 3

In Stage 3, you can decrease each individual group member's income by giving points, or leave it as it is. All other group members are allowed to decrease your income, too, if they so wish. You may assign points in the input screen of Stage 3 which shows, along with the period number and the remaining time, for each group member an indication or "signal" about their allocation to the project. Your allocation will be shown in the row "You".

Please notice that the signal for each of the three other group members only has a 50% probability of equaling his or her actual allocation to the project, unless you paid for information in Stage 1. This means that the signaled allocation to the project for each of the other group members is *accurate* (equals their actual allocation) in 5 out of 10 cases, on average. There is a 50% probability that you will see the *inaccurate signal*, which is a random number which does not correspond to the particular group member's allocation. In this case, there is an equal probability that any integer between 0 and 20 other than the actual allocation will appear.



The input screen of Stage 3

If you invested one Taler for information acquisition in Stage 1, there is a 75% probability that you receive the accurate signal. If you invested two Taler for information acquisition in Stage 1, there is a 100% probability that you receive the accurate signal. That is, in this case you will see the actual allocations of each of the others for sure.

Note that whether a given group member's allocation is signaled accurately is determined by one random draw for all group members who see the signal. That is, either all group members who did not pay for information see the accurate signal (50% chance) or they all see the inaccurate signal (50% chance). Likewise, either all group members who paid 1 Taler for information see the accurate signal (75% chance) or all see the inaccurate signal (25% chance). If the group members who did not pay for information see the accurate signal, then those who paid 1 Taler also see the accurate signal (but not necessarily vice versa). Those who paid 2 Taler always see the accurate signal. If group members see inaccurate signals, they see all the same inaccurate signal for a given group member's allocation.

Finally, the draw determining the signal for one group member's allocation is independent of the draws determining the signals for each other group member. This means there is a 50% chance that the signal for member 2 is accurate and a 50% chance that it is not accurate, for those not investing in information, and likewise there is a 50% chance of accuracy of signal in the case of member 3, and so on. Whether the random draw leads to an accurate or an inaccurate signal for member 2 has no impact on the random draw for member 3 or that for member 4.

Once you view the information on the Stage 3 screen, you have to decide how many points you wish to assign to each group member. You must enter a number at this stage. If you do not wish to alter a certain group member's income, please enter 0. You may still change your decision as long as you have not yet clicked on *OK*.

When distributing points, you incur costs in Taler: each distributed point costs you 1 Taler. The more points you assign, the higher your costs are:

Your cost for assigned points = the sum of points you assign (in Taler)

For example, if you have assigned 2 points to one group member, your cost is 2 Taler; if, in addition, you assign 7 points to another group member, your cost for that is 7 Taler; if you give the final group member 0 points, you have no cost for that member. The *total cost* to you is therefore 9 Taler (2+7+0) which are deducted from the income you had accrued as of the end of Stage 2.

If you choose 0 points for a certain group member, you do not alter this group member's income. If you allocate 1 point (choosing 1) to a group member, you *decrease* this particular group member's Taler income by *3 Taler*. If you allocate 2 points to a group member (choosing 2), you decrease his income by 6 Taler etc. *Each point allocated by you to a particular group member reduces the group member's Taler income by 3 Taler.*

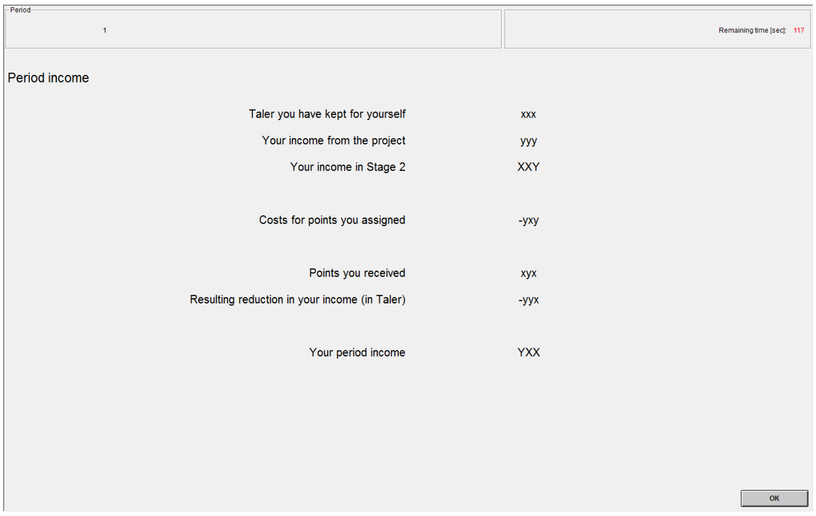
The *overall* reduction in a group member's income from Stages 1 and 2 depends on the total number of points received. If, for instance, one member receives *a total of 3 points* from all other members, the income as of the end of Stage 2 is reduced by *9 Taler*. If a member receives a total of *4 points*, the income is reduced by *12 Taler.* A person who receives points will be informed about the income reduction at the end of each period, without detailed information on the group member (or members) who distributed the point or points.

For your total income at the end of Stage 3, it follows that:

Total Taler income at the end of Stage 3 =
Income after Stage 2
– 3(sum of [effective] points others assign to you)*
– cost of [effective] points you assign to others

Please notice received points cannot decrease your income by more than the income after Stage 2. That is, if the expression [Income after Step 2 – 3*(sum of received points)] yields a negative number, we will reset it to zero. However, your total Taler income at the end of

Stage 3 can be negative if the costs for points that you assign exceed the income after Stage 2 minus the reduction of income due to received points. In other words, there is a limit on the cost others' reductions can impose on your earnings for a period insofar as these alone cannot drive your earnings to below zero, but you always incur the full cost of imposing reductions on others, even if they cause your period income to become negative.



The income screen at the end of Stage 2

Once all members of the group have made their decisions, you will be informed about your period income in the income screen at the end of Stage 3. Here, you see how many Taler you kept for yourself, your income from the project, and the resulting income in Stage 2. In addition, you are informed about the costs for points you assigned, the number of points you received, as well as the resulting reduction in income. Finally, you will see your period income. By pressing the OK button you will proceed to the next period where you receive a new endowment and face all three stages again. There are in total 20 periods and the group composition remains the same.

Your total income at the end of the experiment equals the sum of all period incomes:

Total income (in Taler) = Total sum of period incomes

(If the sum of period incomes is negative, your income is 0 Taler.)

Do you have any further questions?

DFG Research Group 2104

– Latest Contributions

2018:

Bauer, Alexander Max: Sated but Thirsty. Towards a Multidimensional Measure of Need-Based Justice. Working Paper Nr. 2018-03. http://bedarfsgerechtigkeit.hsu-hh.de/dropbox/wp/2018-03.pdf

Khadjavi, Menusch and Nicklisch, Andreas: Parent's Ambitions and Children's Competitiveness. Working Paper Nr. 2018-02. http://bedarfsgerechtigkeit.hsu-hh.de/dropbox/wp/2018-02.pdf

Bauer, Alexander Max: Monotonie und Monotoniesensitivität als Desiderata für Maße der Bedarfs-gerechtigkeit. Working Paper Nr. 2018-01. http://bedarfsgerechtigkeit.hsu-hh.de/dropbox/wp/2018-01.pdf

2017:

Schramme, Thomas: Mill and Miller: Some thoughts on the methodology of political theory. Working Paper Nr. 2017-25. http://bedarfsgerechtigkeit.hsu-hh.de/dropbox/wp/2017-25.pdf

Kittel, Bernhard, Tepe, Markus and Lutz, Maximilian: Expert Advice in Need-based Allocations. Working Paper Nr. 2017-24. http://bedarfsgerechtigkeit.hsu-hh.de/dropbox/wp/2017-24.pdf

Tepe, Markus and Lutz, Maximilian: The Effect of Voting Procedures on the Acceptance of Redistributive Taxation. Evidence from a Two-Stage Real-Effort Laboratory Experiment. Working Paper Nr. 2017-23. http://bedarfsgerechtigkeit.hsu-hh.de/dropbox/wp/2017-23.pdf

Tepe, Markus and Lutz, Maximilian: Compensation via Redistributive Taxation. Evidence from a Real-Effort Laboratory Experiment with Endogenous Productivities. Working Paper Nr. 2017-22. http://bedarfsgerechtigkeit.hsu-hh.de/dropbox/wp/2017-22.pdf

Kittel, Bernhard, Neuhofer, Sabine, Schwaninger, Manuel and Yang, Guanzhong: Solidarity with Third Players in Exchange Networks: An Intercultural Comparison. Working Paper Nr. 2017-21. http://bedarfsgerechtigkeit.hsu-hh.de/dropbox/wp/2017-21.pdf

Chugunova, Marina, Nicklisch, Andreas and Schnapp, Kai-Uwe: On the effects of transparency and reciprocity on labor supply in the redistribution systems. Working Paper Nr. 2017-19. http://bedarfsgerechtigkeit.hsu-hh.de/dropbox/wp/2017-19.pdf

Chugunova, Marina, Nicklisch, Andreas and Schnapp, Kai-Uwe Redistribution and Production with the Subsistence Income Constraint: a Real-Effort Experiment. Working Paper Nr. 2017-18. http://bedarfsgerechtigkeit.hsu-hh.de/dropbox/wp/2017-18.pdf

Nullmeier, Frank: Perspektiven auf eine Theorie der Bedarfsgerechtigkeit in zehn Thesen. Working Paper Nr. 2017-17. http://bedarfsgerechtigkeit.hsu-hh.de/dropbox/wp/2017-17.pdf

Springhorn, Nils: Comparative and Noncomparative Measurement of Need-based Justice. Working Paper Nr. 2017-15. http://bedarfsgerechtigkeit.hsu-hh.de/dropbox/wp/2017-15.pdf

FOR 2104

DFG Research Group 2104 at Helmut Schmidt University Hamburg
http://needs-based-justice.hsu-hh.de