Partial Coercion, Conditional Cooperation, and Self-Commitment in Voluntary Contributions to Public Goods

Elena Cettolin
Arno Riedl

Partial Coercion, Conditional Cooperation, and Self-Commitment in Voluntary Contributions to Public Goods

Abstract

In this paper we experimentally investigate whether partial coercion can in combination with conditional cooperation increase contributions to a public good. We are especially interested in the behavior of the non-coerced populations. The main finding is that in our setting conditional cooperation is not a strong enough force to increase contribution levels. Although, non-coerced subjects rationally adjust their beliefs about contribution behavior of coerced subjects they do not increase their own contributions to the public good accordingly. This points to the limits of the actual strength of conditional cooperation and puts some doubt on the idea that it is crucial in overcoming social dilemma problems.

Elena Cettolin
Maastricht University
Department of Economics (AE1)
P.O. Box 616
NL – 6200 MD Maastricht
The Netherlands
e.cettolin@maastrichtuniversity.nl

Arno Riedl
Maastricht University
Department of Economics (AE1)
P.O. Box 616
NL – 6200 MD Maastricht
The Netherlands
a.riedl@maastrichtuniversity.nl

1 Introduction

In its purest form the consumption of a public good is non-rival and non-excludable, leading to well-known undersupply results that are presented and discussed in all undergraduate and graduate text books in public economics (Rosen and Gayer, 2010; Laffont, 1988). The problem of insufficient contributions to public goods when organized in a decentralized and voluntary way is perhaps one of the best known examples where individual and collective interest are standing in stark contrast. When pondering about whether to voluntarily contribute to a public good, individual material self-interest dictates to free ride on others contributions while collective interest asks for high contributions of all involved. Traditionally it is assumed that in such a situation individual material self-interest prevails and that the economy indeed ends up with an inefficiently low supply of public goods. Therefore, and because a sufficient supply of public goods is perceived of utmost importance for any economy, the development of means and mechanisms to overcome undersupply has a long history in economics (e.g., Lindahl, 1919).

Following the pessimistic statement of Samuelson (1954) that in public goods economies no decentralized process can lead to efficient allocations of resources the traditional theoretical economic literature developed sophisticated mechanisms proving Samuelson wrong on theoretical grounds. The proposed mechanisms indeed ensure that people contribute to public goods when it is in their interest not to contribute in the absence of the mechanism. For instance, the Groves-Ledyard mechanism (Groves and Ledyard, 1977) is a decentralized mechanism where the government sets a taxation-allocation scheme such that it is in each individuals' interest to contribute the efficient amount. A similar but simpler mechanism is proposed by Falkinger (1996). Ledyard (2010) provides a recent critical survey of these and other mechanisms discussed in the literature. All proposed mechanisms, although allowing for decentralized individual decisions, rely on the assumption that some benevolent decision maker can coerce people to contribute their share, usually by sanctioning deviations from the target contribution. Thus, these mechanisms all rely on centralized means of coercion and a paternalistic government.

Laboratory experiments testing these mechanisms show that if the centralized sanctioning is severe enough the mechanisms indeed implement efficient outcomes (see, e.g., Chen and Plott, 1996; Falkinger et al., 2000). The experimental literature on public goods also suggests that there may be other, truly decentralized, ways to ensure efficient or close to efficient provisions of public goods. The research investigating the role of decentralized costly punishment and reward (see, e.g., Fehr and Gächter, 2000; Sefton et al., 2007) indeed shows that especially decentralized individual punishment can have positive efficiency effects in public goods problems, at least when

1

people interact repeatedly (Gächter et al., 2008).[1] These are informative and important results but it is unlikely that decentralized costly punishment alone is a sustainable solution for overcoming the underprovision problem to public goods in complex developed societies. For instance, recent experimental research has shown that heterogeneity among economic actors may undermine the effectiveness of this decentralized enforcement mechanism (Tan, 2008; Reuben and Riedl, 2009). In addition, decentralized punishment may not be feasible in many 'real life' situations. For instance, citizens of a society may not have sufficient information about individual contributions to fine tune punishment towards free riders, which has been shown to be essential for the success of punishment in increasing contributions and efficiency (see, Herrmann et al., 2008). Even if all citizens could know all others contributions, they may still lack appropriate punishment technologies for targeted sanctioning or the law may simply not allow citizens to privately enforce contributions (see also Kosfeld and Riedl, 2007). Another potential limitation of decentralized punishment is that it may be too costly for sustaining efficient contributions to the public good (Anderson and Putterman, 2006; Carpenter, 2007) or not effective enough (Egas and Riedl, 2008; Nikiforakis and Normann, 2008). Hence, it is likely that some coercion by a central planner will be unavoidable for sustaining efficient levels of public good provision.

An important behavioral regularity uncovered by experimental studies on public goods problems is that many people are conditional cooperative (Keser and Winden van, 2000). Being conditional cooperative means that one is ready to contribute one's share if one believes or, better, can be sure that others also contribute their share. Some studies indeed show that beliefs about others contribution behavior strongly influence a person's actual contributions to a public good (Fischbacher et al., 2001; Fischbacher and Gächter, 2010). What these studies also show is that voluntary contributions triggered by conditional cooperation are fragile because they depend on a person's expectation about others contributions. Such uncertainty is inherent in the simultaneous move nature of most public goods problems and, together with pessimistic expectations about others contributions, it may easily lead to a breakdown of voluntary contributions to the public good.[2] One way of taking out this uncertainty in simultaneous move public goods problems is to coerce a subset of players to contribute a minimum amount.

In this paper we combine the idea of conditional cooperation and coercion. In particular, we investigate if partial coercion can be sufficient for guaranteeing high contribution levels to the

---

[1]In experiments where people only interact once this positive efficiency effect is either very weak (Fehr and Gächter, 2002) or not existent (e.g. Egas and Riedl, 2008). For a recent overview of the effects of decentralized punishment on contributions and efficiency see Gächter and Herrmann (2009).

[2]For a comparison of simultaneous and sequential move public goods problems see, e.g., Andreoni et al. (2003), Coats et al. (2009), and Gächter et al. (2010).

public good. The idea is that if conditional cooperation is indeed such a strong force as suggested in the literature than coercion of parts of the society could be used to coordinate beliefs of those not coerced. However, there could also be a downside. The knowledge that some people are forced to contribute their share may increase the temptation to free ride on them. In that case, partial coercion may lead to even worse results than no coercion at all.

In brief we conduct the following public goods experiment, which is played only once. Three subjects are randomly matched into a group where they receive 50 tokens (the experimental money unit) as endowment. They have to decide individually about their contribution to a linear public good. In a baseline treatment everybody is free to contribute whatever amount. In a low-coercion treatment one of the three subjects is randomly chosen and coerced to contribute at least 13 out of 50 tokens, and in a high-coercion treatment one subject has to contribute at least 38 tokens. In a final treatment, we explore the effectiveness of voluntary self-commit in comparison to exogenous coercion and allow subjects to self-commit to minimal contributions of either 0, 13, or 38 tokens.

The main result of the experiment is that partial coercion has no positive effect on contributions beyond the pure coercion effect. In particular, although the non-coerced subjects rationally adjust their beliefs about the contributions of coerced subjects they do *not* adjust their contributions. The picture is very similar in the self-commitment treatment where we also observe a self-selection effect. Those who do not self-commit are also those who free ride in the public good game.

The rest of the paper is organized as follows. Next we give a brief overview of the most closely related experimental literature. Thereafter, Section 2 introduces the experimental design and reports on the procedures, Section 3 presents and discusses the results regarding contributions to the public good, the role of beliefs, and the effect of the possibility of self-commitment. Section 4 provides a brief discussion and concludes.

Related literature.   To the best of our knowledge, there are no experimental papers directly investigating how cooperation in public good games is affected by the presence of players who are coerced to contribute a certain amount. There are a few studies on the related question of how (perceived) obligations affect voluntary contributions to public goods.

In the theoretical law and economics literature obligations are thought to facilitate coordination and cooperation. Besides introducing incentives for compliance, obligations may make the prescribed outcome salient, or focal (McAdams, 2000). The positive effects of obligations on cooperation levels are indeed supported by some experimental studies. Tyran and Feld (2006) experimentally analyze the effects of a law that makes full contribution to a public good game

obligatory while mildly sanctioning free riding. The effects of such mild, non-deterrent law are compared to those of a severe law and to the absence of any legal obligation to contribute. The authors find that an obligation backed by mild incentives does not significantly increase contributions when it is exogenously imposed. However, contributions increase significantly if the mild law is approved in a referendum; the fact that the majority votes in favor of the law seems to induce self-fulfilling expectations of cooperation. The authors also find that without law contributions are very low and that with strong law – basically full coercion – contributions are close to the societal efficient level.

The influence of obligations on beliefs about others' behavior is supported by two studies of Galbiati and Vertova (2007, 2008). In the public good game studied by the authors in the 2008 paper, individuals face an exogenously fixed obligation to contribute a minimum amount; individuals contributing less (more) than the minimum are subject to a probabilistic punishment (reward). The minimum level of contribution varies across treatments, while marginal incentives stay the same: the classical prediction for all treatments is that the obligation to contribute does not affect individual behavior. The authors find that obligations initially significantly affect contributions in a positive way. However, over time they fail to sustain cooperation. In Galbiati and Vertova (2007) the authors investigate whether obligations affect cooperation by coordinating individuals' beliefs about others' contributions to a focal point. The study shows that obligations have indeed a positive effect on individuals' expectations about others' contributions, which is consistent with conditional cooperative behavior, as discussed above. The authors also suggest that a minimum obligatory contribution can have direct effects on preferences if it urges people to update their contribution norms. The authors conclude that the presence of the obligation positively influences preferences for cooperation.

Kroll et al. (2007) investigate cooperation in public good games when collective agreements on obligatory contribution levels are possible. Each participant proposes a desired contribution level and votes for a proposal in her group. In the binding vote treatment, the winning proposal is imposed on all members in a group. In the non-binding treatment, voting is cheap talk and participants can deviate from the chosen contribution level. In the non-binding treatment with punishment participants can sanction group members who deviate from the voting outcome. The authors observe that voting always increases contributions to the public good, especially when the outcome of the vote is imposed on all members. However, contributions in the non-binding treatment only increase marginally and temporarily, suggesting that the contribution norm established with the vote does not survive if violators cannot be punished.

Finally, cooperation is enhanced also by mechanisms other than obligations which seemingly coordinate beliefs around a focal point. Croson and Marks (2001) study the impact of non-

4

binding recommendations on contributions to a threshold public good. The authors find that when valuations for the public good are heterogeneous, recommended contributions significantly increase the likelihood of efficient provision.

2 Experimental design and procedures

The basic game in our experiment is a linear public goods game, also known as voluntary contribution mechanism (VCM). The game is played after all subjects have participated in another, unrelated experiment. The preceding experiment is separated from the reported public goods game by means of a filler task that lasts approximately 7 minutes. Each participant takes part in only one of four experimental treatments. The treatments differ in the extent to which players are coerced to contribute to the public good. The game is played only once, in groups of three participants. We employ a procedure guaranteeing that participants who interacted in the previous experiment are never in the same group. All participants are informed about this.

At the beginning of the experiment each group member $i \in \{1, 2, 3\}$ receives detailed instructions on the computer screen. Participants are informed that everyone in the group receives an endowment of 50 tokens.[3] The marginal monetary benefit of keeping a token is equal to 1, whereas the marginal monetary benefit of contributing a token to the public good is given by $\alpha = 0.5$. Therefore, from a self-interested monetary perspective, the dominant strategy of each member i is not contributing any token to the public good.

In treatment 1 (baseline), after reading the instructions, all participants in a group independently and simultaneously decide how many tokens, if any, they want to contribute to the public good. In treatment 2 (low coercion) after reading the instructions and before the contribution stage, one member per group is randomly selected. The selected group member is coerced to contribute at least 13 tokens, approximately 25% of the endowment, to the public good. In treatment 3 (high coercion) a group member is also randomly selected after the instruction phase and is then coerced to contribute at least 38 tokens, approximately 75% of the endowment, to the public good. In both treatments, the fact that a member will be coerced to contribute the mentioned minimum amount to the public good is common knowledge in the group. In treatment 4 (self-commitment) after reading the instructions and before the contribution stage, subjects (independently and simultaneously) decide voluntarily whether they commit to contribute a minimum amount to the public good. The level of commitment can be low, 13 tokens, or high, 38 tokens. Commitment levels are always binding and revealed to all group members before the contribution stage.

---

[3]For the payment of participants tokens were converted into cash at the rate of 1 token = € 0.025.

The final earnings of a participant i from the public good are given by:

$$\pi_i = 50 - c_i + \alpha \sum_{j=1}^{3} c_j,$$

where $c_i \geq 13$ for one randomly selected i in the low-coercion treatment, $c_i \geq 38$ for one randomly selected i in the high coercion treatment, and $c_i \geq 13$ or $c_i \geq 38$ for any i who chooses to commit to the respective minimum contribution in the self-commitment treatment.

Apart from contributions, we are interested in the effect of coercion on players' anticipation of what the other group members will contribute. This is of importance because other studies (Fischbacher et al., 2001; Fischbacher and Gächter, 2010) have shown that many people condition their behavior on their beliefs about others' contributions. Specifically, if they expect others to contribute much they contribute more than if they expect others to contribute little. Therefore, in all treatments, right before deciding how many tokens they want to contribute to the public good, participants are asked to provide their best estimate on how much the other group members will contribute to the public good. Rather than eliciting a point belief, we decided to elicit an interval, which also gives us as an idea about the guesser's confidence in his/her estimate. Specifically, each participant is asked to indicate what s/he thinks will be the minimum and maximum contribution of each group member. The belief elicitation is incentivized using the interval scoring rule (Schlag and van der Weele, 2009). The rule works as follows: if the true contribution of another member lies in the indicated interval the guessing subject earns an amount that is inversely related to the length of the indicated interval. If the true contribution of a member lies outside the indicated interval the subject earns nothing. More precisely, let c_j be the actual contribution of j, and $\underline{c}_{ij}$ and $\overline{c}_{ij}$ be the minimum and maximum, respectively, of the interval indicated by i regarding j's contribution, then the earnings of i from guessing j's contribution is determined as follows:

$$\pi_{ij} = \begin{cases} 0.12 \times (50 - (\overline{c}_{ij} - \underline{c}_{ij})), \text{ if } c_j \in [\underline{c}_{ij}, \overline{c}_{ij}] \\ 0, \text{ if } c_j \notin [\underline{c}_{ij}, \overline{c}_{ij}], \end{cases}$$

with $c_j, \underline{c}_{ij}, \overline{c}_{ij} \in [0, 50]$.

Eliciting an interval has the advantage that it gives information not only about the location of the belief distribution, but also about its dispersion. The width of the interval determines an upper bound on the variance of the belief distribution, which makes the width of the interval a proxy for how confident the decision maker is about his/her guess. This holds for any risk neutral or risk adverse decision maker.

In the self-commitment treatment subjects may also form beliefs about the other group members' level of commitment to the public good. Therefore, after the commitment phase and

before eliciting beliefs on the contribution levels, we ask subjects to indicate the likelihood in percentage points of the six possible commitment scenarios. In the first scenario none of the other two group members commits and in the last scenario both commit to high contributions. This belief elicitation is incentivized with the quadratic scoring rule, which rewards subjects for the accuracy of their prediction according to the following formula:

$$\pi_i \;=\; 3 - 0.00015 * \left[(p_0 - r_0)^2 + (p_{13} - r_{13})^2 + (p_{26} - r_{26})^2 + (p_{38} - r_{38})^2 + \right.$$
$$\left. + (p_{51} - r_{51})^2 + (p_{76} - r_{76})^2 \right]$$

Where p_n indicates participant's i likelihood estimate, expressed in percentage points, that the sum of commitments equals n. As an example, if i beliefs that with 50% chance both other members commit to contribute 13 tokens then $p_{26} = 50$. The realized commitments are indicated by r_n. If, for instance, i's other group members commit to contribute respectively 13 and 38 tokens, then $r_{51} = 100$ and all other r_n equal zero.

The experiment ends after the contribution phase where subjects are informed about the total amount contributed to the public good, their earnings from the public good game and from the beliefs elicitations. Thereafter they answer a short questionnaire, are paid in cash and dismissed from the laboratory. The experiment lasted approximately 25 minutes inclusive the instructions. In total 276 subjects participated in the experiment: 72 in each of the baseline, low coercion, and self-commitment treatments and 60 in the high coercion treatment.

3 Results

In the following we first report on group level contributions to the public good for all treatments. We then proceed with a discussion of contributions of non-coerced (non-committed) group members, where we also zoom into the role of beliefs in contribution behavior. Finally, we report the results of the self-commitment treatment in more detail.

3.1 Contributions to the public good

Table 1 reports first and second order descriptive statistics of the contributions to the public good, separate for treatments.

Result 1. *On the group level, contributions are highest in the high-coercion treatment, followed by the low-coercion, the self-commitment and the baseline treatment. The difference between high coercion and the other treatments is statistically significant. All other differences are insignificant.*

Support for this result can be found in Figure 1 and Table 1. Average contributions to the public good vary roughly between 15 tokens (30 percent) in the baseline treatment and 25 tokens (50 percent) in the high coercion treatment. The medians of contributions are very close to the averages. When comparing the group average contributions across all four treatments a Kruskal-Wallis equality-of-population rank (henceforth, KW) test detects a significant difference ($\chi_3^2 = 11.831$, $p = 0.0080$, two-sided). Pair-wise comparisons between treatments with Mann-Whitney rank sum (henceforth, MW) tests reveal that contributions in the high coercion treatment are significantly higher than in all other treatments ($p \leq 0.0328$, two-sided). All other pair-wise comparisons of group average contributions return insignificant results ($p > 0.1766$, two-sided).

Table 1: Group level contributions

treatment	average	st.dev.	median	no. of obs.
baseline	15.61	11.18	16.67	24
low coercion	18.75	8.76	20.17	24
high coercion	25.52	8.64	25.17	20
self-commitment	16.61	9.51	16.67	24

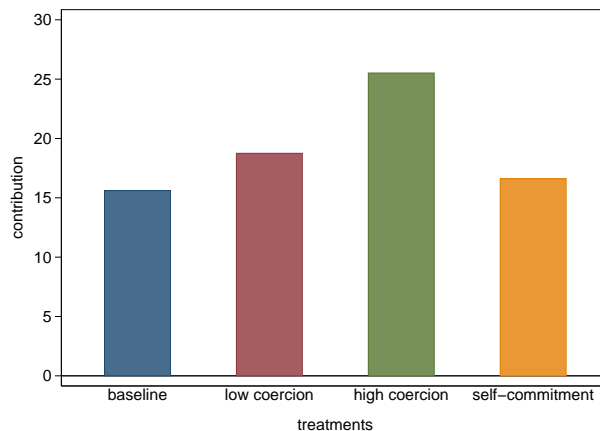Note: unit of observation is group average contribution.

Figure 1: Group level average contributions

We are especially interested in the contribution behavior of non-coerced (non-committed) subjects. If it holds that conditional cooperation is an important behavioral mechanism for achieving high voluntary contribution rates then the knowledge that some subjects are coerced into (committed to) relatively high contributions should have a positive effect on the contributions of non-coerced (non-committed) subjects. Specifically, we expect that non-coerced subjects in the low coercion treatment contribute more than subjects in the baseline treatment and, in turn, that subjects in the high coercion treatment contribute more than subjects in the low coercion

treatment. Our next result shows that these hypotheses are not supported by the data.

Result 2. *In the low and high coercion treatments those subjects who are not coerced contribute the same amount in both treatments and not more than the average amount observed in the baseline treatment. In the self-commitment treatment non-committed subjects in groups with at least one self-committed subject contribute significantly less than non-coerced subjects in the other treatments.*

Figure 2 and Table 2 show data in support of the result. When comparing the group average contributions of non-coerced (non-committed) subjects across all four treatments a KW test detects a significant difference ($\chi_3^2 = 26.221$, $p = 0.0001$, two-sided). Pair-wise comparisons between treatments with MW tests reveal that in the self-commitment treatment contributions of not self-committed subjects in groups with at least one committed subject are significantly lower than contributions of non-coerced subjects in either of the other three treatments ($p < 0.0001$, two-sided). All other pair-wise comparisons of average contributions return insignificant results ($p \geq 0.6540$, two-sided).[4]

Table 2: Average contributions of non-coerced (non-committed) subjects

treatment	average	st.dev.	median	no. of obs.
baseline	15.61	11.18	16.67	24
low coercion	15.71	11.80	12.75	24
high coercion	17.48	13.28	18.25	20
self-commitment	2.06	4.13	0.0	18

Note: in the baseline treatment the unit of observation is the group average contribution; in the low- and high coercion treatments it is the average contribution of non-coerced subjects in a group, and in the self-commitment treatment it is the average contribution of not self-committed subjects in groups with at least one self-committed subject.

This result seems surprising. First, the fact that knowing that others are coerced to contribute relatively high amounts to the public good does not have any positive effect on contributions of those who are not coerced contradicts the conditional cooperation hypothesis. Second, the fact that not self-committed subjects contribute even less when they know that others self-commit to relatively high contributions than when there is no self-commitment possible reinforces this interpretation. It also suggests that self-commitment works as a selection device, in the sense that (conditional) cooperators commit themselves whereas selfish people do not commit. However,

---

[4]The above result does only change marginally when also taking those groups into account where no subject self-committed in the self-commitment treatment (see Appendix A).
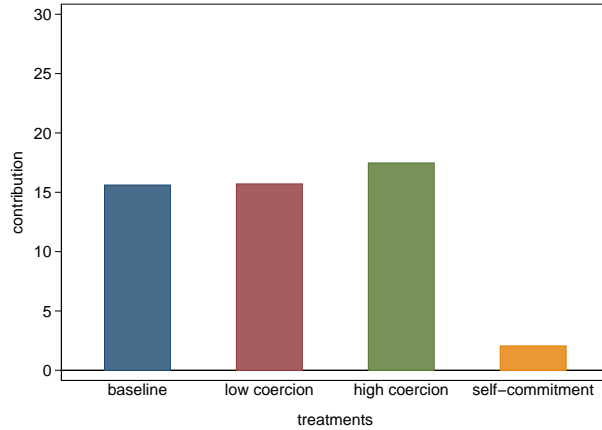
Figure 2: Average contributions of non-coerced (non-committed) subjects

the studies on conditional cooperation also point out that conditional cooperators contribute only when they expect others to contribute. Therefore, we next investigate the role of beliefs in contribution behavior.

3.2 The role of beliefs

Conditional cooperation can increase contributions in our coercion treatments only if non-coerced subjects adjust their beliefs about coerced subjects accordingly. The next result shows that this is the case for non-coerced subjects in the high coercion treatment, but that otherwise coercion has no effect on beliefs about others contributions.

Result 3. *Neither in low- nor in the high-coercion treatment does coercion affect the coerced subject's beliefs about the contributions of the non-coerced subjects. The non-coerced subjects adjust their beliefs only in the high-coercion treatment.*

Figure 3 shows the average belief intervals regarding the contributions of the other group members in the baseline treatment and the average belief intervals of coerced (non-coerced) group members regarding there non-coerced (coerced) counterparts in the low- and high-coercion treatments. The figure indicates that beliefs of coerced subjects about the contributions of non-coerced subjects do not differ between the low- and high coercion treatments and are also not different from beliefs in the baseline treatment. This visual impression is corroborated by a Kruskal-Wallis test. When testing the equality of beliefs of minimum contributions between the baseline and the two coercion treatments the test does clearly not reject the null hypothesis ($\chi^2_2 = 1.284$, $p = 0.5263$, two-sided). The same holds for the beliefs of maximum contribution ($\chi^2_2 = 0.150$, $p = 0.9279$, two-sided) and the belief interval length ($\chi^2_2 = 1.108$, $p = 0.5745$, two-sided). Hence, coerced subjects do neither believe that the non-coerced subjects will conditionally contribute

10

nor do they belief that they will exploit the situation and more strongly free ride than in the baseline treatment. As one would expect, non-coerced subjects beliefs about coerced subjects are affected by coercion. When testing equality of minimum, maximum, and interval length of beliefs about coerced subjects' contributions the null hypothesis of equality of distributions is clearly rejected ($\chi_2^2 = 41.957$, $p = 0.0001$; $\chi_2^2 = 34.055$, $p = 0.0001$, $\chi_2^2 = 15.848$, $p = 0.0004$, respectively; KW tests, all two-sided). Pairwise comparisons with Mann-Whitney tests show that these difference are solely driven by beliefs in the high coercion treatment where coerced subjects had to contribute at least 38 tokens. In particular, there is no belief shift in the low coercion treatment in comparison to the baseline treatment ($|z| \leq 0.712$, $p \geq 0.4764$, for all three measures; MW tests, two-sided). In contrast, when comparing the minimum, maximum, and interval length of beliefs of non-coerced subjects in the baseline and low-coercion treatment with the high-coercion treatment, MW tests reject equality of distributions for all pair-wise comparisons and all three measures ($|z| \geq 3.258$, $p \leq 0.0011$, two-sided).
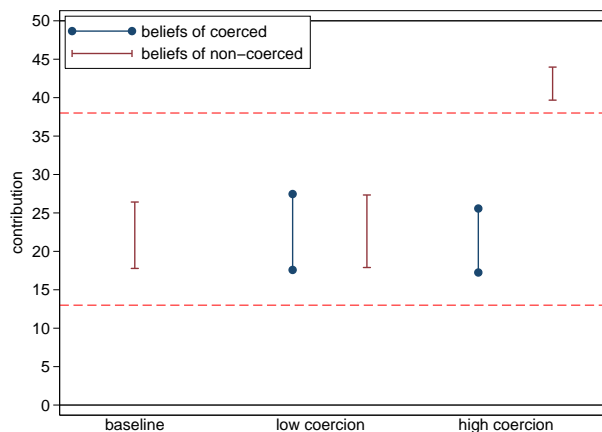

Figure 3: Average belief intervals of coerced (non-coerced) subjects about non-coerced (coerced) contributions

In the previous section we have seen that neither low nor high coercion changes the contribution behavior of non-coerced subjects. The above results on beliefs of coerced subjects show that they correctly anticipate this contribution behavior of their counterparts.

A precondition for the existence of conditional cooperation is that contribution behavior is positively correlated with the expectations about others contributions. Figure 4 shows the relation between the elicited (average) belief intervals and the (average) contributions in the baseline treatment and for both coercion treatments; for the latter treatments separate for coerced and non-coerced subjects. For the baseline and the low coercion treatment, the plots indicate a positive correlation between belief intervals and own contributions. Hence, there is evidence for conditional cooperative behavior in both treatments and in case of low coercion, for both,
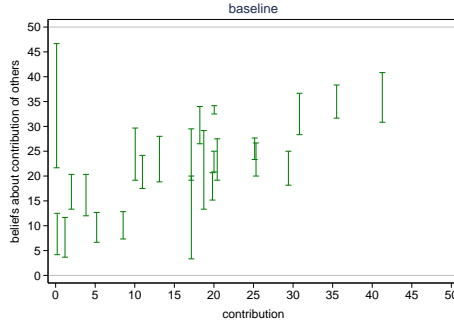
(a) baseline
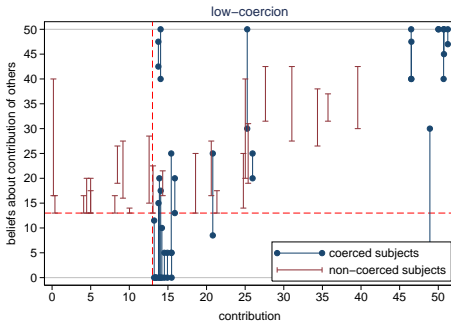
(b) low-coercion
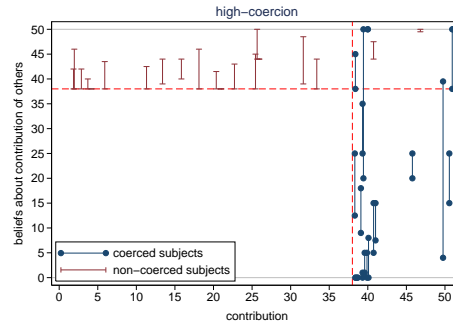
(c) high-coercion

Note: in baseline treatment group level data, in coercion treatments type level data; small random noise added to contributions in order to avoid overlap in graph.

Figure 4: Relation between own contribution and belief (interval) about others contribution in the baseline and the two coercion treatments

coerced and non-coerced subjects (cf. Figures 4a,b). Indeed, Spearman rank order correlations between beliefs about minimum contributions of others and own contributions are positive and statistically significant in the baseline treatment and for coerced as well as non-coerced subjects in the low coercion treatment (baseline: $\rho = 0.6565, p = 0.0005$; low coercion - coerced subjects: $\rho = 0.7221, p = 0.0001$; low coercion - non-coerced subjects: $\rho = 0.7188, p = 0.0001$; all two-sided).

In the high-coercion treatment (cf. Figure 4c), there is a significantly positive correlation between non-coerced subjects own contributions and their beliefs on coerced subjects' minimum contribution ($\rho = 0.5902, p = 0.0062$, two-sided). For coerced subjects there is also a positive but statistically insignificant correlation between their contributions and their beliefs ($\rho = 0.3079, p = 0.1867$, two-sided). We attribute this insignificance mainly to the fact that there is much less room left for coerced subjects to act according to their beliefs.[5] Tobit regressions estimating

---

[5]The correlation statistics do not change significantly when looking at the believed maximum of the other subject type's contribution.

contributions as a function of believed minimum contributions and believed contribution interval length largely corroborate the non-parametric tests and also indicate a statistically significant positive correlation between coerced subjects' beliefs on the minimum contributions of their non-coerced counterparts and their own contributions. Table 3 reports these estimates. The regressions are run separately for each treatment and within the coercion treatments separately for coerced and non-coerced subjects.

Table 3: Tobit regressions of contributions as function of believed contribution by others

| | Dependent variable: own contribution | | | | |
| | baseline | low-coercion | | high-coercion | |
		coerced	non-coerced	coerced	non-coerced
$belief\_minimum$	0.893^{***}	0.692^{***}	1.474^{***}	0.178^{*}	3.242^{***}
	(0.207)	(0.144)	(0.254)	(0.071)	(0.726)
$belief\_interval$	-0.378	0.358	0.032	0.115	1.122
	(0.385)	(0.364)	(0.311)	(0.130)	(0.831)
const	2.619	10.454	-11.242^{*}	38.168^{***}	-116.181^{***}
	(5.643)	(5.483)	(4.829)	(2.060)	(30.315)
LR χ^2_2	15.71	17.55	22.07	6.21	13.59
Prob $> \chi^2$	0.0004	0.0002	0.0000	0.0448	0.0011
No. of obs.	24	24	24	20	20

Note: *** $^{(**)}$ $^{[*]}$ indicate two-sided significance levels at 0.1 (1) [5] percent; standard errors between parentheses.

We summarize the evidence discussed above in our next result.

Result 4. *In the baseline and both coercion treatments coerced and non-coerced subjects' contributions are positively correlated with their beliefs about their counterparts contributions.*

Table 4 reports average contributions for non-coerced subjects split according to whether their beliefs about minimum contributions of coerced subjects where below or equal to the coercion level or strictly above it. For the coercion treatments this returns the contributions of those who believe that the coerced subjects contribute minimally exactly the coerced level and of those who believe that coerced subjects contribute strictly more than the coerced level, respectively. For the baseline treatment it returns the contributions of those who believe that their other group members (which are not coerced by definition) will contribute equal to or less than the coercion levels of the coercion treatments and of those who believe that their counterparts contribute strictly more than those coercion levels, respectively. The question we want to answer here is whether a 'forced' lift – due to coercion – in expectations regarding others' contributions

makes those who are otherwise pessimistic contribute more. If conditional cooperation is an important motivational force in contribution behavior, even if contributions are exogenously enforced, we should find that contributions are higher in case of lifted expectations. At first sight, the results reported in Table 4 suggest that this could be the case in the low coercion treatment. Comparing contributions in that treatment with those in the baseline treatment indicates that lifted expectations increase contributions from 5.44 tokens to 8.00 tokens. However, a Mann-Whitney test rejects the hypothesis that this difference is statistically significant ($z = -0.707, p = 0.4798$, two-sided). Interestingly, in comparison to the baseline treatment, in the low coercion treatment the belief that the minimum contribution of coerced subjects will be strictly larger than the coercion level is less frequent (79 percent in baseline vs. 58 percent in low coercion). This is different in the high coercion treatment where the frequency of beliefs that contributions of coerced subjects will be strictly above the coercion level of 38 is 29 percent, whereas it is 0 percent in the baseline treatment. However, in the high coercion treatment exogenously lifted beliefs have a negative albeit statistically insignificant effect on contributions. Comparing contributions in that treatment with those in the baseline treatment indicates that lifted expectations decrease contributions from 15.61 tokens to 12.12 tokens ($z = 0.764, p = 0.4448$). This following result summarizes.

Table 4: Non-coerced subjects' average contributions for beliefs below and above coercion levels

	average contributions for minimum beliefs			
	equal or below 13[†]	strictly above 13	equal or below 38[†]	strictly above 38
baseline	5.44	19.00	15.61	–
	(6.24)	(10.43)	(11.18)	–
	[n=6]	[n=19]	[n=24]	[n=0]
low-coercion	8.00	21.21		
	(6.86)	(11.63)		
	[n=10]	[n=14]		
high-coercion			12.12	27.43
			(10.83)	(12.10)
			[n=13]	[n=7]

Note: [†] for the low- and high coercion treatments these are the average contributions of non-coerced subjects who believe that the minimum of contributions of their coerced counterparts is exactly 13 and 38, respectively.

Result 5. *For low coercion exogenously lifted expectations insignificantly increase contributions but make less non-coerced people than in baseline believe that the coerced ones will contribute strictly more than the coercion level. For high coercion more non-coerced people than in baseline believe that the coerced ones will contribute strictly more than the coercion level but exogenously lifted expectations insignificantly decrease contributions of non-coerced subjects.*

This and the preceding results in this section indicate that, on the one hand, subjects are conditionally cooperative in the sense that they contribute more the more they expect others to contribute, but that, on the other hand, coercion fails to create an upward shift in contributions of non-coerced subjects. The latter holds, despite the fact that non-coerced subjects consistently adjust their expectations regarding the contributions of coerced subjects. In addition, in both coercion treatments there are countervailing forces at work regarding the effect of lifted expectations. Together these results explain why the contributions of non-coerced subjects in the coercion treatments do not significantly differ from subjects in the baseline treatment where nobody is coerced by definition (cf. Result 2).

An open question is whether non-coerced subjects' beliefs regarding the other non-coerced subjects' contributions is negatively (or positively) influenced by the fact that there is one subject coerced to contribute a minimum amount. Even conditional cooperative people may refrain from contributing more in the coercion treatments when they believe that the other non-coerced subject will free ride on the coerced subject. The data show that there is neither a positively nor a negatively significant effect of coercion on these beliefs. In the baseline treatment subjects' beliefs about the other members minimum and maximum contributions are 17.8 and 26.4 tokens, respectively. In the low- and high coercion treatments these beliefs are 15.7 and 15.2, respectively, for the minimum contributions, and 25.7 and 26.1, respectively, for the maximum contributions. The belief interval lengths in the baseline, low-, and high coercion treatment are 8.6, 10.0, and 10.8, respectively. None of these differences is statistically significant (KW tests: $\chi_3^2 < 2.246$, $p \geq 0.3254$, two-sided).

3.3 Self-comittment and contributions

In this section we focus on the results in the treatment where subjects could freely choose to commit themselves (i.e., self-coerce) to either a low (13 tokens) or a high (38 tokens) minimum contribution level. We first report whether subjects self-commit and, if so, on which levels. Thereafter, we examine how this affects contributions behavior of non-committed subjects. (Recall, that it is public knowledge to all group members how many in the group have committed themselves and at which levels.) The first result for this treatment shows that many subjects are indeed ready to self-commit to minimum contribution levels.

Result 6. *In the self-commitment treatment a slight majority of individuals self-commits to a minimum contribution level. About one-third commit to contribute at least 13 tokens and almost 20 percent to contribute at least 38 tokens.*

Support for this result can be found in Figure 5. It shows that in total about 53 percent of all subjects commit to either a low (33.3 percent) or a high (19.4 percent) minimum contribution
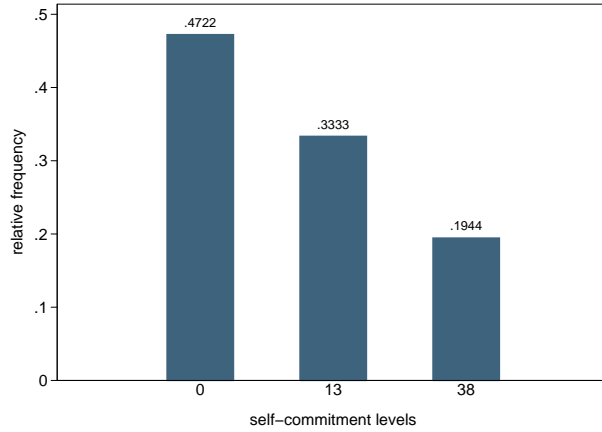
Figure 5: Frequencies of individual self-commitments

level. In addition, Figure 6 shows the frequency of all possible combinations of self-commitment levels across groups. It can be seen that all possible combinations occur at least once, except for the combinations where two group members choose 13 and one chooses 38 and where all three group members choose the high self-commitment level of 38 tokens. The most frequent combination (25 percent) is where each group member chooses a different self-commitment level. Perhaps not surprisingly, the self-commitment combinations involving a majority of group members choosing a high commitment level are rather scarce (8.34 percent in total). Interestingly, no self-commitment at all is also rather infrequent. It happens only in 12.5 percent of the cases that no member in a group commits.
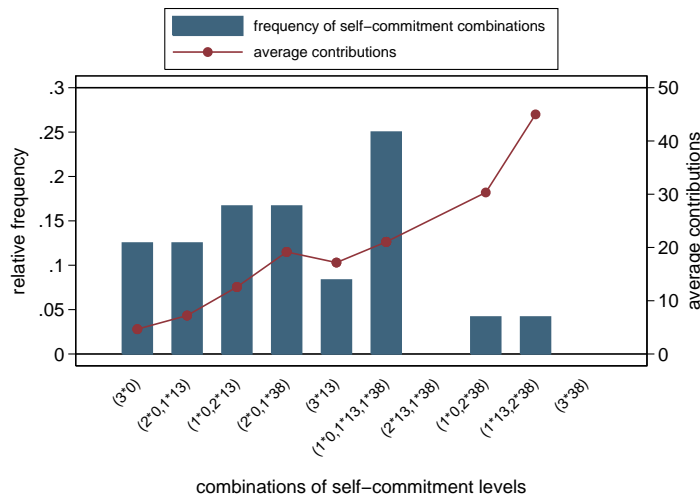

Figure 6: Frequencies of self-commitment combinations and associated average contributions

Figure 6 also shows the average contributions in groups given the indicated combinations of commitment levels. Not surprisingly the average contributions increase with the strength of

16

the commitments. Groups that do not commit at all contribute on average close to zero (4.67 tokens, $n = 3$). The one group where two members commit to at least 38 tokens and one member commits to at least 13 tokens contribute a group average of 45 tokens, which is clearly above their average commitment level of 29.67 tokens. Similarly, the groups where all three members committed to at least 13 tokens contribute on average 17.17 tokens. Hence, these examples suggest that the possibility to commit increases overall contributions.

(a) Average contributions of (not) self-committed subjects as a function of average commitment levels

(b) Self-commitment combinations and associated average contributions of (not self-committed subjects
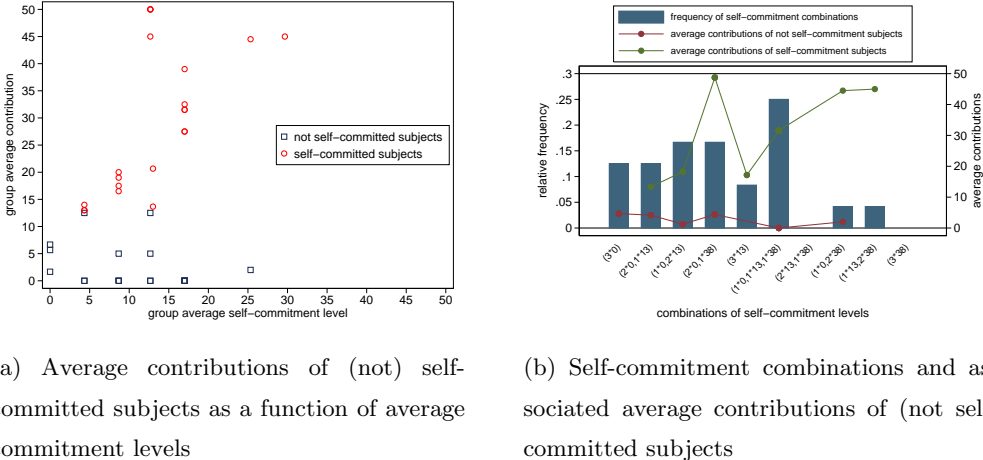
Figure 7: Relation between commitment levels and contributions of (not) self-committed subjects

However, recall from Section 3.1 (Results 1 and 2) that taken over all groups contributions in the self-commitment treatment are not higher than in the other treatments and that subjects who do not self-committed are contributing significantly less than non-coerced subjects in the other treatments. This suggests that in groups with self-commitment of some but not all members the not self committed members strongly free ride. Figure 7 shows that this is indeed the case. Figure 7a depicts a scatter plot of average contributions of committed and non-committed subjects on the average self-commitment level in the group. It clearly shows that not self-committed subjects contribute very little and that they do not increase their contributions with higher commitment levels of the other group member(s). The visual impression is corroborated by a correlation analysis. Spearman rank order correlation coefficients of contributions on commitment levels is even marginally significantly negative ($\rho = -0.3953$, $p = 0.0761$, two-sided, $n = 21$) for not committed subjects. Clearly, for self-committed subjects there is a strong and significant positive correlation between contributions and commitment levels ($\rho = 0.5675$, $p = 0.0073$, two-sided, $n = 21$). Figure 7b shows alternatively that for each chosen commitment combination the contributions of not committed subjects are very low. In fact, their contributions are highest when no one in the group self-committed, but even in that case it only reaches a level of 4.67 tokens on average. Hence, basically all positive contributions are made by those

17

who do self-commit to some minimum level. We summarize in our last result.

Result 7. *In the self-commitment treatment almost all positive contributions are made by self-committed subjects, while not committed subjects almost fully free ride.*

An interpretation of this rather somber result is that self-commitment works as a selection device. It is those subjects who have an inclination to conditionally or unconditionally contribute who commit to minimum levels. Those who do not commit could in principle be pessimistic conditional cooperators or free riders. The data strongly suggest that it is rather the latter type of subjects that does not self-commit, because even in groups with high commitment levels not self-committed subjects do contribute close to zero.

4 Discussion and Conclusion

In this paper we experimentally investigate whether partial coercion or self-commitment can increase voluntary contributions to a public good. The hypothesis that coercing a subset of people positively affects contributions of those not coerced is based on the well documented empirical observation that many people are conditional cooperative (Keser and Winden van, 2000; Fischbacher et al., 2001; Fischbacher and Gächter, 2010). Conditional cooperators contribute more to a public good the more others contribute or the more they expect others to contribute. If, as in our experiment, some of these others are coerced to contribute a minimum amount there is no uncertainty about these contributions and conditional cooperators should in principle be willing to contribute, too. Our results do not support this hypothesis. Although, non-coerced subjects rationally adjust their beliefs about contribution behavior of coerced subjects they do not increase their own contributions to the public good accordingly. In consequence, the overall contribution levels in the two environments with high and low partial coercion do not differ from an environment where all subjects are free to choose any contribution to the public good.

Notably, we observe this result although our subjects are clearly conditional cooperative minded. That is, most of our subjects behave conditionally cooperative in the sense that they contribute more the more they believe others to contribute. This, however, is only a second order effect. What partial coercion fails to create is a first order effect: it neither creates an upward shift in beliefs of coerced subjects about non-coerced contributions nor does it – thereby fulfilling these pessimistic beliefs – lead to an upward shift in contributions of non-coerced subjects. It seems that conditional cooperation only 'works' when positive contributions of others to the public good are perceived as voluntary and not if they are enforced by some third party. This strongly suggests that intentions (e.g., Falk et al., 2008) and accountability (e.g., Konow, 2000) matter. Hence, for conditional cooperators it is not enough that others contribute large positive amounts,

they also must 'mean' it. That is, conditional cooperators respond with higher contributions to high contributions of others if they believe that others had the intention to do so voluntarily and if they can be made responsible for these contributions.

Our finding may also be viewed as a special new case of motivation crowding-out (e.g., Frey, 1997; Frey and Jegen, 2001). In the motivation crowding-out literature it is shown that external intervention can undermine intrinsic motivation. For instance, Gneezy and Rustichini (2000) show that introducing a fee for parents coming late to pick up kids at a day care center has adverse effects on pick up time and, more recently, Holmås et al. (2010) found evidence that monetary incentives intended to decrease hospital length of stay increases it instead. These and other examples of motivation crowding-out refer to direct effects whereas the results in our experiment point to an indirect crowding-out effect. It is the external intervention that forces *others* into a particular behavior that crowds out one's own motivation to contribute to the public good in response to the increased contributions of other people. Admittedly this interpretation is to some extent speculative and more evidence is needed to see whether such indirect crowding-out indeed takes place.

Our results regarding the effect of partial coercion on contributions to a public good put some doubt on the idea pursued in legal theory that people obey laws, pay their taxes, or – more generally – contribute to public goods, because they know that others will be forced if necessary (Hart, 1961). This severely limits the range of legal government intervention in cases where coercion is not complete. Hence, the public enforcement and punishment of, for instance, caught tax dodgers may not have the desired effect on other potential tax dodgers. The government may achieve better results in terms of tax compliance by 'manipulating' the beliefs of tax payers about other tax payers compliance behavior. Indeed, survey evidence suggests that there exists a high correlation between perceived tax evasion and tax morale (Frey and Torgler, 2007). A recent experimental study by Lefebvre et al. (2011) conducted in four different European regions shows that negative information about others tax evasion behavior indeed has adverse effects on tax compliance. We believe that these are interesting observations that are also important for the theoretical modeling of coercion (Winer et al., 2008). Usually these models are silent about the fact that a large fraction of people is willing to pay their taxes as long as others do. In addition, our experiment has revealed a potentially important additional cost of coercion, namely that it may crowd out some of this intrinsic motivation. It also indicates that some of the costs of coercion – as defined in Winer et al. (2008) – could be avoided by using information policy instead of legal coercion to make citizens contributing to a public good. This is also of interest in the light of the study by Rider et al. (2011) who experimentally investigate the willingness of resisting coercion. Such costly resistance actions may also be avoided when citizens voluntarily contribute because they believe that (most) others also voluntarily contribute.

In our experiment we have also tested if self-commitment – that is, self-coercion – to a minimum contribution level can overcome the free-rider problem in public good provision. Here we find that a majority of people is indeed willing to commit themselves even if they believe that others may not do so. However, overall this inclination to self-commitment does not lead to higher voluntary contributions to the public good. The main reason for this negative result is that those who do not self-commit are not contributing anything to the public good. Hence, self-commitment serves 'only' as a selection device. Those who do not self-commit are those who prefer to fully free-ride on the contributions of others. This result resonates with the finding in Kosfeld et al. (2009) who investigate the voluntary establishment of organizations that can coerce their members to contribute fully to the public good. Importantly, in their experiment not everybody had to join an organization in order to establish it. When only a subgroup was organized, those outside the organization could choose any contribution level. Their main result is that most of the time organizations are formed only if all people join, thereby eliminating the possibility of free-riding. However, in some circumstances only subgroups formed an organization, and in these cases those outside the organization indeed showed a tendency to free-ride completely. A behavior very similar to the behavior we observe for the not self-committed subjects. However, in contrast to our experiment, Kosfeld et al. (2009) find that the voluntary establishment of coercive organizations does increase the supply of public goods. This indicates that self-commitment can increase public good provision, but only if it is embedded in an appropriate institutional environment.

References

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*, 54:1–24.

Andreoni, J., Harbaugh, W., and Vesterlund, L. (2003). The carrot or the stick: Rewards, punishments, and cooperation. *American Economic Review*, 93(3):893–902.

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

Chen, Y. and Plott, C. R. (1996). The Groves-Ledyard mechanism: An experimental study of institutional design. *Journal of Public Economics*, 59:335–364.

Coats, J. C., Gronberg, T. J., and Grosskopf, B. (2009). Simultaneous versus sequential public good provision and the role of refunds – an experimental study. *Journal of Public Economics*, 93(1-2):326 – 335.

Croson, R. T. A. and Marks, M. B. (2001). The effect of recommended contributions in the voluntary provision of public goods. *Economic Inquiry*, 39(2):328–249.

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

Falk, A., Fehr, E., and Fischbacher, U. (2008). Testing theories of fairness-Intentions matter. *Games and Economic Behavior*, 62(1):287–303.

Falkinger, J. (1996). Efficient private provision of public goods by rewarding deviations from average. *Journal of Public Economics*, 62(3):413–422.

Falkinger, J., Fehr, E., Gächter, S., and Winter-Ebmer, R. (2000). A simple mechanism for the efficient provision of public goods: Experimental evidence. *American Economic Review*, 90:247–264.

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

Fehr, E. and Gächter, S. (2002). Altruistic punishment in humans. *Nature*, 415:980–994.

Fischbacher, U. and Gächter, S. (2010). Social preferences, beliefs, and the dynamics of free riding in public goods. *American Economic Review*, 100(1):541–556.

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

Frey, B. S. (1997). The cost of price incentives: An empirical analysis of motivation crowding-out. *American Economic Review*, 87(4):746–755.

Frey, B. S. and Jegen, R. (2001). Motivation crowding theory. *Journal of Economic Surveys*, 15(5):589–611.

Frey, B. S. and Torgler, B. (2007). Tax morale and conditional cooperation. *Journal of Comparative Economics*, 35(1):136 – 159.

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(1514):791–806.

Gächter, S., Nosenzo, D., Renner, E., and Sefton, M. (2010). Sequential vs. simultaneous contributions to public goods: Experimental evidence. *Journal of Public Economics*, 94(7-8):515 – 522.

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

Galbiati, R. and Vertova, P. (2007). Behavioral effects of obligations. Econpubblica Working Paper 120, Bocconi University.

Galbiati, R. and Vertova, P. (2008). Obligations and cooperative behavior in public good games. *Games and Economic Behavior*, 64:146–170.

Gneezy, U. and Rustichini, A. (2000). A fine is a price. *Journal of Legal Studies*, XXIX(1):1–18.

Groves, T. and Ledyard, J. (1977). Optimal allocation of public goods: A solution to the 'free rider problem. *Econometrica*, 45(4):783–809.

Hart, H. L. A. (1961). *The Concept of Law*. Oxford University Press, Oxford, U.K.

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

Holmås, T. H., Kjerstad, E., Lurås, H., and Straume, O. R. (2010). Does monetary punishment crowd out pro-social motivation? a natural experiment on hospital length of stay. *Journal of Economic Behavior and Organization*, 75(2):261–267.

Keser, C. and Winden van, F. (2000). Conditional cooperation and voluntary contributions to public goods. *Scandinavian Journal of Economics*, 102:23–39.

Konow, J. (2000). Fair shares: Accountability and cognitive dissonance in allocation decisions. *American Economic Review*, 90:1072–1091.

Kosfeld, M., Okada, A., and Riedl, A. (2009). Institution formation in public goods games. *American Economic Review*, 99(4):1335–1355.

Kosfeld, M. and Riedl, A. (2007). Order without law? Experimental evidence on voluntary cooperation and sanctioning. *KritV - Kritische Vierteljahresschrift fur Gesetzgebung und Rechtswissenschaft*, 90(1-2):140–155.

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

Laffont, J.-J. (1988). *Fundamentals of Public Economics*. MIT Press, Cambridge, MA; London, England.

Ledyard, J. O. (2010). Non-coercion, efficiency and incentive compatibility in public goods decisions. Mimeo, California Institute of Technology.

Lefebvre, M., Pestieau, P., Riedl, A., and Villeval, M. C. (2011). Tax evasion, welfare fraud, and "The Broken Windows" effect: An experiment in Belgium, France and the Netherlands. Working Paper 3408, CESifo, Munich, Germany.

Lindahl, E. (1958 [1919]). Die Gerechtigkeit der Besteuerung, translated (in part) as "Just taxation: A psoitive solution. In Musgrave, R. and Peacock, A., editors, *Classics in the Theory of Public Finance*, pages 168–176. Macmillan.

McAdams, R. H. (2000). A focal point theory of expressive law. *Virginia Law Review*, 86:1649–1729.

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

Reuben, E. and Riedl, A. (2009). Cooperation and punishment in privileged groups. *Journal of Conflict Resolution*, 53(1):72–93.

Rider, M., Ackert, L. F., and Gillette, A. B. (2011). Cooperating to resist coercion: An experimental study. Research Paper Series 11-06, Andrew Young School of Policy Studies.

Rosen, H. S. and Gayer, T. (2010). *Public Finance*. McGrawhill, 9 edition.

Samuelson, P. (1954). The theory of public expenditure. *Review of Economics and Statistics*, 36(4):387–389.

Schlag, K. and van der Weele, J. (2009). Efficient interval scoring rules. Economics Working Papers 1176, Universitat Pompeu Fabra.

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

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

Tyran, J.-R. and Feld, L. P. (2006). Achieving compliance when legal sanctions are non-deterrent. *Scandinavian Journal of Economics*, 108:135–156.

Winer, S. L., Tridiman, G., and Hettich, W. (2008). Social welfare and coercon in public finance. Working Paper 2482, CESifo.

A   Additional Results

Table A.1 and Figure A.1 show the group level contributions of non-coerced (low and high co-ercion treatments) and not self-committed subjects. A Kruskal-Wallis test indicates statistically significant differences between treatments ($\chi_3^2 = 27.311$, $p = 0.0001$, two-sided) and Mann-Whitney tests show that in the self-commitment treatment contributions of not self-committed subjects are significantly lower than contributions of non-coerced subjects in the other three treatments ($p < 0.0001$, two-sided). All other pair-wise comparisons of average contributions return insignificant results ($p \geq 0.6540$, two-sided).

Table A.1: Group level contributions of non-committed subjects

treatment	average	st.dev.	median	no. of obs.
baseline	15.61	11.18	16.67	24
low coercion	15.71	11.80	12.75	24
high coercion	17.48	13.28	18.25	20
self-commitment	2.43	4.01	0.0	21

Note: in baseline unit of observation is group average contri-bution; in low and high coercion the average contribution of non-coerced subjects in a group; in self-commitment the av-erage contribution of not self-committed subjects in groups with at least one self-commitment.
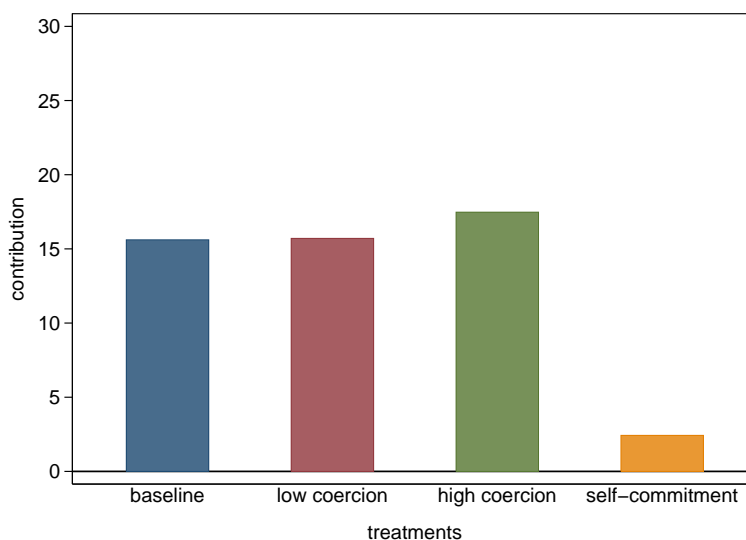
Note: In self-commitment treatment also those groups where nobody committed are taken into account.

Figure A.1: Group level average contributions of non-committed subjects