

Public Goods and Voting on Formal Sanction Schemes: An Experiment

Louis Putterman^{a*}, Jean-Robert Tyran^b and Kenju Kamei^a

^a Department of Economics, Brown University, Box B, Waterman Street 64, Providence, RI 02912, United States

* Corresponding author. Louis_Putterman@Brown.Edu. Tel.: 401-863-3837, fax 401-863-1970.

^b Department of Economics, University of Copenhagen, Øster Farimagsgade, DK-1353 Copenhagen K, Denmark

Abstract

The burgeoning literature on the use of sanctions to support public goods provision has largely neglected the use of formal or centralized sanctions. We let subjects playing a linear public goods game vote on the parameters of a formal sanction scheme capable both of resolving and of exacerbating the free-rider problem, depending on parameter settings. Most groups quickly learned to choose parameters inducing efficient outcomes. But despite uniform money payoffs implying common interest in those parameters, voting patterns suggest significant influence of cooperative orientation, political attitudes, and of gender and intelligence.

JEL classification: C91, C92, D71, D72, H41

Keywords: Public good; voluntary contribution; formal sanction; experiment; penalty; voting.

1. Introduction

An important insight from the last decade of research on the provision of public goods is that the problem of voluntary cooperation can be partially resolved by allowing individuals to engage in informal sanctions, i.e. in costly punishment of other group members after learning of their contributions (Fehr and Gächter, 2000). This insight provides an important stimulus to the economic theory of collective action because a preference-based willingness to incur a cost to punish free riding can dramatically alter equilibrium outcomes in social dilemma problems, and evidence suggests that such a preference-based willingness to punish is widespread (Falk, Fehr and Fischbacher, 2005). Because of its importance, a burgeoning literature has explored what conditions are favorable to making informal sanctions effective.¹

In contrast, “formal sanctions” — by which we mean the application of sanctions by a central authority or by a group acting collectively — have received much less attention in the laboratory.² This neglect is surprising for at least two reasons. First, centralized punishment institutions are widely observed in the real world — consider fines for non-payment of taxes or for failing to buy a subway ticket or put the required amount in a parking meter. By contrast, it is difficult to think of real world counterparts to the decentralized monetary punishment included in the experiments just mentioned, unless such punishment is thought of as a representation of social sanctions like shunning, badmouthing, or ostracism. Second, formal sanctions may be

¹ The effectiveness of informal sanctions varies greatly, depending on the gains from cooperation, the ratio of punisher cost to punishee loss, and the subject population. For example, Cinyabuguma, Page and Putterman (2004) have pointed out that earnings are often lower rather than higher when punishment is introduced in experiments such as Fehr and Gächter (2000). By varying the ratio of cost to the punisher versus that to the person targeted for punishment, Nikiforakis and Normann (2008) and Egas and Riedl (2005) shed light on the conditions under which the unrestricted opportunity to punish increases efficiency, while Herrmann, Thöni and Gächter (2008) show the sensitivity of outcomes to the subject pool. Auxiliary mechanisms such as pre-play communication or sorting by subject type are also needed to increase efficiency, in many settings. For experiments that allow communication, see e.g. Brosig, Ockenfels and Weimann (2003), Bochet, Page and Putterman (2006); for sorting by subject type, see Gunnthorsdottir, Houser and McCabe (2007), Page, Putterman and Unel (2005).

² There is an extensive literature studying incentive-compatible mechanisms in the provision of public goods (e.g. Chen and Plott, 1996 or Falkinger, Fehr, Gächter and Winter-Ebmer, 2000). Such mechanisms involve tax (and subsidy) mechanisms which can be thought of as formal sanctions in the sense that they are rule-bound. However, such mechanisms are often complex and that may also be the reason why they rarely observed “in the wild.”

more reliable and equitable. Formal sanctions are rule-bound and are ideally meted out to all norm-violators uniformly. They may therefore be thought to deter free riding rather effectively. In contrast, informal sanctions often suffer from the “scattershot” nature of unconstrained punishing by individuals (e.g. Cinyabuguma, Page and Putterman, 2006 and Herrmann, Thöni and Gächter 2008), and they may invite retaliation (e.g. Nikiforakis, 2008).³ While the effectiveness of informal sanctions depends on the presence and incidence of social preferences, individuals interested solely in their own monetary payoff can in principle agree to use formal sanctions in order to eliminate free-riding incentives and achieve a Pareto-superior outcome. For example, Tyran and Feld (2006) let subjects vote to implement severe sanctions that effectively eliminate the free-rider problem and found that the large majority of groups vote for such sanctions and comply with them.

This paper provides the first study to investigate whether groups left to themselves and given no external guidance choose efficient formal sanctions by voting. A possible reason why this question has not been studied previously is that the issue is trivial from the perspective of the rational actor paradigm. When given the choice between an efficient scheme, i.e. one that effectively deters free riding, and a scheme without any sanction, i.e. one where full free riding is optimal, rational and self-interested voters will vote for the efficient sanction (see, for example, Buchanan and Tullock, 1965). The reason is that a rational, self-interested player compares his or her payoff without sanction (the full free-riding outcome) with the outcome with efficient sanctions (i.e. cooperation at the efficient level by all players). However, mounting experimental evidence suggests that such equilibrium predictions often fail because the assumptions needed to deduce them are not necessarily met for all players in practice. In particular, many people cooperate to some extent even when material incentives are stacked

³ In his *Second Treatise on Civil Government* (1689), John Locke wrote: “In the state of nature ... resistance many times makes the punishment dangerous, and frequently destructive to those who attempt it.”

against cooperation, because they have pro-social preferences. At the same time, some players free ride even when it is not in their material interest, perhaps due to bounded rationality or for ideological reasons.

Alternative preferences or bounded rationality might either strengthen the prediction of voting for efficient formal sanctions or cause it to break down. First, while subjects having preferences for cooperation (such as the conditional cooperators identified by Fischbacher, Gächter and Fehr, 2001) can be expected to favor efficient penalties, and in that respect may be no different from the maximizers of own payoff assumed by standard theory, individuals inclined to punish free riders at personal cost might pick up on the desirability of centralized penalties directed at free riders more quickly than do payoff-maximizers, because emotion might aid cognition.⁴ Second, and in contrast, the kind of individuals found to engage in what Cinyabuguma *et al.* (2006) call “perverse” and Herrmann *et al.* (2008) call “anti-social” punishment in experiments with informal sanctions may prefer to work *against* group efficiency by voting to have high rather than low contributors be penalized. Third, some individuals may oppose formal sanctions out of an ideological or other dislike of “coercive” measures. To investigate the possible influences of bounded rationality, cooperative orientation and ideology on voting, our design includes tasks generating measures of intelligence, conditional willingness to contribute to a public good, and political views.

In addition to providing a first experimental study of choice among formal sanction schemes, our paper also contributes to a recent stream of literature on endogenous formation of institutions. Several recent experiments have given subjects the opportunity to choose between allowing or not allowing informal sanctions (Botelho, Harrison, Rutström and Costa-Pinto, 2005; Ertan, Page and Putterman, 2009) or rewards (Sutter, Haigner and Kocher, forthcoming)

⁴ See Damasio (1994). Evidence that cognitively difficult problems are more easily solved when the decision-maker is focused on detecting violations of social norms is presented by Cosmides and Tooby (1992).

either by voting (the previous papers) or by choice of group (Gürerk, Irlenbusch and Rockenbach, 2006).⁵ In those papers, the vote is about whether to permit informal sanctions (rewards), leaving the decision on whether to punish (reward) in the hands of each individual. In this paper, we permit voting choice over a menu of rules under which sanctions are automatic once voted upon and triggered by the relevant contribution behavior. Although the impact of formal sanctions has been studied previously, no paper has studied voting choice over multiple dimensions of the sanction scheme, as in our design.

The choices confronting our subjects, after initially facing a standard linear voluntary contributions mechanism (VCM) with homogeneous endowments, are (a) whether contributing to the public or to the private account is subject to a penalty, (b) what (if any) level of contribution to the account in question is exempt from penalty, and (c) the maximum amount of the penalty. Together, (b) and (c) determine the penalty rate per unit contributed to the public account or the private account once the exempt contribution is exceeded. We study subjects' abilities to select efficient schemes over the course of multiple votes and multiple sets of contribution decisions. We also conduct a baseline VCM treatment without voting or penalties in order to compare behavior and its evolution over time for the same group size, return from public account, instruction phrasing, number of interactions, and subject pool.

Our main findings are that most groups in the treatment with voted penalty schemes learn to choose parameters that induce efficient outcomes, whereas in the baseline treatment contributions decay over time in the usual fashion. Analyzing votes on scheme parameters, we find that intelligence and cooperative orientation are significantly positive predictors of more efficient choices, while politically conservative or anti-incentive beliefs, and female gender, are significant negative predictors of such choices. We also find impacts of intelligence,

⁵ Kosfeld, Okada and Riedl (2008) also study choice of institutions in a public goods setting, but in a design involving neither formal nor informal sanctions.

cooperative orientation, and political outlook on contributions to the public good, although they are muted or disappear when penalties are in place. We conclude that while standard theoretical reasoning need not lead us astray about the choice of efficient sanction schemes, designing good institutions requires experience, and such institutions may fail to emerge in populations having less cognitive sophistication or more aversion to the use of penalties to achieve social goals.

The rest of the paper proceeds as follows. Section 2 provides details of the experimental design and theoretical predictions. Section 3 reports our results. Section 4 provides a summary discussion and conclusions. An Appendix, providing supporting material, is available on-line.

2. Experimental Design and Predictions

In the linear VCM that is the core of our experiment, a subject is assigned to a group, provided with a fixed endowment each period, and asked to make a series of decisions on the allocation of that endowment between a private and a public account. His/her payoff per period is given by

$$C_i + \alpha \cdot \sum_{\text{all } i \text{ in a group}} G_i = (W - G_i) + \alpha \cdot \sum_{\text{all } i \text{ in a group}} G_i \quad (1)$$

where W is the uniform per-period endowment, C_i is subject i 's contribution to his or her private account, G_i is i 's contribution to the public (or group) account, and $0 < \alpha < 1$ is the marginal per capita return (MPCR), the domain of which makes it privately optimal to set $G_i = 0$ but socially optimal to set $G_i = W$.

Our experiment has two treatments. In the BASELINE treatment, subjects play the standard linear public goods game for 24 periods in fixed groups. In the endogenous PENALTY treatment, subjects vote on a scheme consisting of two elements. First, they decide whether to assess fines on contributors to private accounts or to the public account. By this decision, they determine whether it is low or high contributors to the public account that are fined. Second, they vote on the penalty structure, which determines the strength of the sanction and its domain

(what, if any, contributions are exempt from fines). As in BASELINE, each subject plays a total of 24 contribution stage games in fixed groups.⁶ We chose groups of five in order to induce free-riding problems, make cooperating without communication or fines a serious challenge, and avoid tied votes. We set the endowment per period at 20 points (28 points = 1 dollar) and the MPCR (α in equation (1)) at 0.3.

The penalty structure is a linear function with two parameters: maximal penalty (described to subjects as “fine”) level, a , and penalty-free (equivalently “exemption”) level, b . Consider the case in which it is contributing to the private account that gives rise to a penalty (see Figure 1.a). If a subject contributes nothing to the public account, she has to pay a . If she contributes more than b to the public account, she pays no penalty. For each unit assigned to the private rather than public account, beyond the exemption, she is fined a/b .

Although it would damage efficiency, subjects can also elect to penalize group members for contributing to the public account, in which case the chosen b is the amount that can be put in that account without facing a penalty and a is the maximum penalty that the subject pays should she choose to contribute her full endowment to the account (see Figure 1.b). Less-than-maximum fines are levied for intermediate contribution levels, as shown by the values lying on the line segments labeled ab in the two figures.

[Figure 1]

In the PENALTY treatment, the first four periods are identical to those in the BASELINE treatment, that is, without fines. A voting opportunity is then provided after the fourth period and again every four periods thereafter. We refer to a set of four periods as a “phase,” the experiment session being composed of six of these, five of which are governed by

⁶ In both BASELINE and PENALTY treatments, before the start of the contribution game, subjects are clearly informed that the session ends in exactly 24 periods.

voted rules in the PENALTY treatment. Voting occurs privately and without opportunity for communication, in the following order:

1. Vote on whether “private account” or “public account” contributions can be penalized; learn group outcome (the option receiving a majority of votes);
2. Vote a maximal penalty level a ; learn group outcome (the median a value voted for);
3. Vote a penalty exemption level b ; learn group outcome (the median b value voted for).

Subjects learn vote outcomes but not how many voted for them or what the individual votes were. Subjects can choose to have no penalty by voting to set $a = 0$ or, when contributions to the private (public) account are penalized, setting $b = 0$ (20). To control for a “restart effect” (Andreoni, 1998), play in the BASELINE treatment is also conducted as a set of four-period phases punctuated by pauses and announcements that another four periods under the same rules will begin.⁷

The penalty possibilities are kept simple to facilitate subjects’ comprehension and to allow votes to be taken relatively quickly. Nonetheless, the voting task facing our subjects is considerably more complex than in existing experiments that permit only an up-or-down vote on a pre-arranged scheme. Instructions (included in the Appendix) are neutrally worded, with no hint as to what it might be desirable to penalize, at what level, or why one might want fines. A subject participated in one treatment only and was not informed about the existence of the other treatment.⁸

⁷ In both treatments, subjects are told at the outset that there will be six phases of four periods each, that the initial instructions cover the first four periods, and that there will be additional instructions before the next phase. Subjects thus have exactly the same information in periods 1 – 4 of the BASELINE treatment as in the corresponding periods of the PENALTY treatment. Throughout both treatments, their understanding is that there may be some new information, rule or procedure introduced in the instructions for the next four periods. With the exception of the break between Phase 1 and Phase 2 of the PENALTY treatment, the information at each break is that the next phase will have the same number of periods and follows the same rules as the previous one.

⁸ One question the reader might pose is why we have groups selecting exemption and fine levels rather than simply mandating a required contribution. One reason is that real-world tax systems couple required tax payments with penalties for non-compliance. Also, we wanted to study subjects’ choice of what to penalize over a space of options resembling that available in the decentralized or informal sanctions literature.

Before the main part of the experiment, we elicited conditional contribution schedules (see Figure 2 for the time line). We asked each subject to choose his or her contribution level to a public account conditional on 21 possible average contribution levels of other group members (0, 1, 2, ..., 20). In order that the responses constitute costly decisions, we announced an extra period between periods 1 and 2 in which each group member's earnings are calculated using the unconditional period 1 contribution of four group members and the conditional contribution of one randomly chosen member given the average period 1 contribution of the other four.⁹

[Figure 2]

After period 24, each subject took a short intelligence test consisting of 15 problems. To encourage them to give their best answers, we gave subjects 15 points for a correct response and deducted 3 points for an incorrect response. The questions are taken from Séréville and Myers (1994) and are based on Raven's progressive matrices (see the Appendix for an example). As an indicator of intelligence, we use the number of correct responses. We included these questions to test whether voting is affected by cognitive ability. The scores, which we refer to as IQ,¹⁰ also allow us to check whether contributing to the public good is affected by intelligence, and how intelligence is correlated with conditional willingness to contribute and political views.

Finally, each subject answered a questionnaire consisting of 8 questions about political preferences (see Appendix). We included these questions so that we could test whether political views including those regarding the desirability of government intervention influence votes on the penalty scheme.¹¹

⁹ Our procedure resembles that used in Fischbacher *et al.*, 2001 and Fischbacher and Gächter (forthcoming).

¹⁰ The term "IQ" is used loosely by us insofar as the scores may not satisfy some conventions with respect to standard deviation and other aspects of the score distribution.

¹¹ As with any design of this type, one may ask whether the answers to a survey administered after the main part of the experiment might be influenced by experience in it. We believe that any such influence is unlikely to vary much among subjects in a given design, so that whatever the influence of the experiment on the survey responses, response differences among subjects in a treatment will with some faithfulness reflect differences of political views.

Predictions

In line with the well-known features of the VCM, in the BASELINE treatment the dominant strategy of each stage game is $C_i = 20$, $G_i = 0$ and the payoff is 20 for all players. In the PENALTY treatment, an individual's payoff is given by

$$(20 - G_i) + 0.3 \cdot \left(\sum_{\text{all } i \text{ in a group}} G_i \right) - p_i \quad (2)$$

where $p_i = \begin{cases} \left(a - \frac{a}{b} G_i \right) \cdot \mathbf{I}_{\{G_i \leq b\}} & \text{if contributing to the private account is penalized} \\ \left[(G_i - b) \cdot \frac{a}{20 - b} \right] \cdot \mathbf{I}_{\{G_i \geq b\}} & \text{if contributing to the public account is penalized} \end{cases}$
and $\mathbf{I}_{\{G_i \leq b\}}$ ($\mathbf{I}_{\{G_i \geq b\}}$) is an indicator variable that equals 1 when $G_i \leq b$ ($G_i \geq b$) and 0 otherwise.

An individual can achieve a maximum payoff of 44 if $a = 0$, she contributes nothing to the public good, and her four teammates contribute their full endowments. However, if the individual assumes that all group members are equally rational and intent on maximizing their individual payoffs, she knows that at $a = 0$, the others will also contribute nothing and her earnings will be 20. Therefore, if she expects to be pivotal in the decision, a rational player seeking maximum earnings will vote to penalize contributions to the private account and to set $a \in (14, 20]$ and $b = 20$, thus compelling herself and her teammates to contribute $G_i = 20$ and earn 30 points. While there are many subgame perfect equilibria with non-pivotal voting, the possibility of errors suggests that the individual would reject these weakly dominated choices in favor of the three efficient choices above, selection of which constitutes a trembling-hand perfect equilibrium (Selten, 1975).¹² In this equilibrium, each subject earns 30 by contributing $G_i = b = 20$ in response to penalty rate $(a/b) > 0.7$, and no fines are in fact paid. Note that these choices of a and b implement the efficient allocation.

¹² The fact that group members have no opportunity to communicate makes estimating the likelihood of being pivotal exceedingly difficult, providing an added rationale for selecting the weakly dominating alternatives.

Although a rational, self-interested subject who believes others to be the same will always vote for the parameters indicated, experimental subjects may not vote this way for at least three reasons. First, some subjects may fail to understand the mechanism and may vote in error for penalizing contributions to the public account, for setting a low b , or for a low a . Second, subjects may prefer to indulge ideological or social tastes even at some cost in terms of potential income; for example, an aversion to fines and/or to being “coerced” by fines into contributing to the public good may motivate some to vote for $a = 0$, and individuals with such a mindset may also vote for penalizing contributions to the public account in order to deter others from introducing a contribution-inducing penalty scheme. Finally, differences in cooperative orientation may affect voting, for instance the type of subject seen to punish high rather than low contributors in public goods experiments with informal sanctions—a type usually associated with free riding at the contribution stage—may also vote for penalizing contributions to the public rather than the private account. In contrast, high contributors in voluntary contribution settings, who do most of the punishing of free riders in informal sanctions experiments (Fehr and Gächter, 2000; Herrmann et al., 2008), should be at least as inclined as are their perfectly self-interested counterparts to impose high penalties on free riders by choice of a formal sanction scheme.

3. Results

Our experiment was conducted at a computer lab at Brown University in June and July, 2009, using undergraduate subjects drawn from the full spectrum of fields.¹³ As summarized in Table 1, forty subjects in eight groups with five subjects each participated in two sessions of the BASELINE treatment and the same number in eight groups in two sessions of the PENALTY

¹³ 9 out of 80 subjects listed their major as Economics, similar to the proportion in the university’s undergraduate population as a whole. Others study a broad range of fields, from the humanities to the sciences.

treatment.¹⁴ The average earnings of subjects were \$28.52 and \$32.85 (including a \$5 show-up fee and an average of \$3.54 from the IQ portion) in the BASELINE and PENALTY treatments, respectively.

[Table 1]

3.1 Characteristics of contributions, earnings, and voting

Contributions and earnings

Figure 3.a shows that contributions in BASELINE, i.e. absent any penalty, conform closely to those seen in the literature: average contributions begin at around 40% of the endowment and decline over time,¹⁵ with clear “restart effects” evident in periods 5, 9, 13, 17 and 21. The trend is similar in the PENALTY treatment during periods 1 – 4, when no fines existed.¹⁶ Average earnings parallel average contributions.¹⁷

Result 1. When formal sanctions are not available by design, average contributions and earnings replicate the standard pattern, beginning around 40% of endowment and declining with repetition.

¹⁴ Twenty subjects took part in each session. Subjects were randomly assigned to groups, and thus, could not identify which others present were members of their own group. Anonymity was further preserved in that subjects could not associate contributions with specific others in their group. No communication was permitted between subjects. Participants were recruited from the general student population using the BUSSEL (Brown University Social Science Experiment Lab) on-line recruitment tool and the experiment was programmed in z-tree (Fischbacher, 2007).

¹⁵ The stylized summary of previous VCM experiments is that subjects contribute around 40 - 60% of their endowments to the public account in the first period (Ostrom, 2000), with the exact level depending in part on the MPCR and group size (Isaac and Walker, 1988; Isaac, Walker and Thomas, 1984; Isaac, Walker and Williams, 1994). Since we use a relatively low MPCR, average 1st period contributions near the lower end of the stated range are to be expected. A Wilcoxon Signed Ranks Test reveals that the average contribution in periods 13 – 24 at the group level is weakly significantly different from that in periods 1 – 12 (p -value = 0.093 < 0.10, 2-tailed), confirming the declining trend.

¹⁶ The average period 1 contribution of 7.9 in BASELINE treatment and 6.5 in PENALTY treatment are not significantly different according to an individual-level Mann-Whitney test (p -value = 0.313, 2-tailed). A Wilcoxon Signed Ranks Test using group-level observations from the PENALTY treatment reveals that the average contribution in periods 1-2 and that in periods 3 – 4 are significantly different, with p -value = 0.0497 < 0.05 in a 2-tailed test, confirming the apparent declining trend.

¹⁷ As can be confirmed from equation (1), group average earnings are a monotonic transformation of average contribution to the public account when there are no fines.

[Figure 3]

Voting

In the PENALTY treatment, the large majority of votes by individuals — 89% — were cast in favor of penalizing contributions to the private rather than the public account. Table 2 reports the voting outcomes at group level and shows that only two votes (the 4th and 5th votes by Group 13) resulted in a majority for punishing contributions to the public account. The proportion of individuals voting to penalize private rather than public account contributions rises from 80% in the first vote to 90% or more in each subsequent vote, although showing no clear trend during votes 2 – 5 (See Appendix Figure B.1).

Turning to the choice of a and b , Table 2 shows that optimal parameters were chosen by all but one group in the last three of five votes, with two setting optimal parameters in votes 2 – 5 and two achieving optimal settings in all five votes. Choosing an efficient penalty scheme was not automatic, however, as shown by Group 13’s difficulty agreeing on high a and high penalty rate a/b . After experiencing low a/b and lower earnings than it had achieved in Phase 1 during Phases 2, 3 and 4, this group chose in its 4th and 5th votes to penalize contributions to the public account, further exacerbating a deteriorating trend in its members’ contributions.¹⁸ In spite of this outlier group, the overall tendency is for a and b to increase with successive votes, suggesting learning.

[Table 2]

Result 2. In the PENALTY treatment, there is almost uniform support for penalizing

¹⁸ Since only one of eight groups failed to construct an efficient penalty scheme, we checked to see what about Group 13’s composition was distinctive. We found that in the eight PENALTY treatment groups, the number of female subjects has observed values 1, 2, 3 and 4, each characterizing two groups. Group 13 had four female members, placing it at the high end of the distribution. Also, the members of group 13 were 2nd lowest among the eight groups in terms of members’ average IQ scores, which is 7.2 in Group 13 and 8.35 on average in the PENALTY treatment. Group 13’s members did not have unusually conservative political ideology nor did it contain an unusually large number of subjects classified as free riders or “other” (the latter classification being treated as a possible indication of confusion in Fischbacher and Gächter (forthcoming)).

noncontribution to the group account, and voters in most groups quickly learn to choose optimal sanctioning schemes.

In spite of failure to achieve optimal penalty rules in some sets of periods and especially in one of the eight groups, Figure 3 shows that average contributions in the PENALTY treatment were considerably higher in periods 5 – 24, when penalty schemes operated, than in periods 1 – 4, when they did not.¹⁹ Notably, Figure 3 shows that once voting began and contributions to the private account were selected for penalization in all but one group, the average contribution in the PENALTY treatments rose, reaching the full endowment during the last three phases in the seven successful groups,²⁰ as compared to the declining trend in the BASELINE treatment and other VCM experiments. Both contributions and earnings are significantly higher in the PENALTY than in the BASELINE treatment in periods 5 – 24, and the median contribution over all groups in periods 5-24 is 0 in BASELINE treatment versus 20 in PENALTY.²¹

Result 3. The chosen penalty schemes increase contributions immediately, and contributions are sustained or rise with repetition in the PENALTY treatment. Contributions and earnings are significantly higher when there is an opportunity to vote on formal sanctions than when that opportunity is absent.

¹⁹ Dividing the data of the PENALTY treatment into periods 1 – 4 and 5 – 24 and comparing the average contribution of each group in the two sets of periods, Wilcoxon Signed Ranks Tests confirm that they are significantly different, with p -value = 0.017. Period-by-period contribution averages at the group level reveal that all PENALTY treatment groups except Group 13 registered strong increases in contribution levels after the first vote; see Appendix Table B.1.

²⁰ In these groups, 419 of 420 contributions decisions in phases 4, 5 and 6 were to contribute exactly 20, with the remaining decision being a contribution of 19. That contributions rose over time during the phases under voted penalties is shown by the fact that a Wilcoxon Signed Ranks Test finds average contribution in periods 5 – 8 is lower than that in periods 21 – 24, using group level observations, with p -value 0.025.

²¹ We performed Mann-Whitney tests using group-level average contributions in periods 5 – 24 for PENALTY versus BASELINE treatment groups, obtaining a p -value of $0.0063 < 0.01$ in a 2-tailed test. Mann-Whitney tests using group level average earnings for periods 5 – 24 also find earnings in the PENALTY treatment significantly different from those in the BASELINE treatment, with a p -value of $0.0063 < 0.01$ in a 2-tailed test. For the median contributions, see Appendix Table B.1.

Further evidence that the higher contributions to the public account in the PENALTY treatment are attributable to groups' penalty choices comes from analyzing individual-level decisions with multi-variate regressions. We estimate random effects Tobit regressions²² in which contribution to the public account is the dependent variable and the explanatory variables are the penalty per unit below the exemption level, a/b , the exemption level b itself, IQ, the vote number ($v = 1, 2, \dots, 5$), which also identifies the phase, and the round (period) number within the phase ($= 1, 2, 3, 4$). We exclude the observations for members of Group 13 during the two phases in which contributions to the public account were penalized in that group, since the impact of a and b should be quite different with such a penalty regime. We also exclude the observations in which $b = 0$ since then $a/b = \infty$.²³ We note here that the results of this exercise ought properly to be treated as provisional, pending investigation of the possible endogeneity of the voted parameters, a question we return to at the end of Section 3.

The results, found in Table 3, show that, regardless of the specification of the regression, the higher is the per-unit penalty level and the higher is the exemption level b , the more did individuals contribute to the public account, effects that are highly significant. To test whether the impact of higher penalty levels depends only on whether penalty rate a/b reaches the threshold level of 0.7 per unit, we also estimate regressions (shown in columns (4) – (6)) substituting for a/b the dummy variable “Binding.”²⁴ The results indicate that the threshold

²² All of our regressions that include multiple observations for given subjects need to be interpreted with caution, since these observations are not strictly speaking independent. We use Tobit regressions for contribution decisions, since these are potentially censored at 0 and 20. We use random effects estimates (using the xtobit command in Stata) because a sufficient statistic allowing fixed effects to be conditioned out of the likelihood of a Tobit model does not exist.

²³ We thus exclude the 40 observations of subjects in groups 10 and 14 in Phase 2, during which $b = 0$. To check the sensitivity of our regression results to this exclusion, we also estimated variants of the regressions including an “ $a/b = \infty$ ” dummy variable and setting $a/b = 0$ if $a/b = \infty$, instead of dropping the observations. These regressions produce highly significant positive coefficients on the added dummy variable and little change in other coefficients or their significance levels. Results are omitted to conserve space.

²⁴ “Binding” = 1 if $a/b \geq 0.7$, otherwise “Binding” = 0.

dummy variable has at least as much explanatory power as the continuous variable a/b . Vote number has significant positive coefficients, indicating that contributions are rising from phase to phase, even after controlling for the penalty parameters, while round within phase has significant negative coefficients, a finding that Figure 3 suggests is attributable mainly to the downward trends within Phases 2 and 3, during which some groups had not yet set a and/or b at efficient levels (Table 2).²⁵ The other columns and other variables in Table 3 are discussed below.²⁶

Result 4. Contributions are well predicted by the voted penalty parameters.

[Table 3]

3.2 Heterogeneous social preferences, voting choice and contribution behavior

We elicited measures of subjects' conditional willingness to cooperate, intelligence, and political attitudes in order to obtain controls for factors we hypothesized might hamper the choice of an efficient sanctioning scheme. Before using the information thus obtained, we briefly introduce it.

Conditional Contribution Schedule

Following Fischbacher et al., (2001), we measure subjects' conditional willingness to voluntarily contribute to a public good using their decisions in the "strategy method" portion of the experiment. As Appendix Figure B.2.a shows, our subjects' decisions closely resemble those of the subjects in Fischbacher *et al.*, with quite similar proportions classifiable as "conditional cooperators" (50%), "free riders" (31%), "hump-shaped contributors" (14%), and "other"

²⁵ Group 13 alone exhibits downward trends within Phases 4, 5 and 6, and the observations of that group from phases 5 and 6 are excluded from the regressions.

²⁶ To check that our results for a/b , b and "binding" are not dependent on the inclusion of the other regressors shown in Table 3, we also estimated regressions excluding all other terms apart from the intercept and period terms. Results for the variables discussed above remain largely unchanged.

(6%).²⁷ In our regression analysis, we use a more continuous statistic of the contribution schedules rather than type dummy variables, which sacrifice degrees of freedom. We considered for this purpose the average conditional contribution and the slope of conditional contribution as a function of others' average contribution. The two variables are highly correlated in our data (as in Fischbacher and Gächter, forthcoming), and average conditional contribution achieves the best fit in the regressions that follow.²⁸ The average conditional contribution (ACC) of those of our subjects fitting the conditional cooperator criterion is 8.22, that of hump-shaped contributors is 4.10, that of free riders 0.01, and that of "other" (not classifiable) 6.49.²⁹

Cognitive Ability Measured by IQ Test

Appendix Figure B.3 provides a histogram of our IQ measure. There appears to be ample dispersion, making it possible to identify associations between behaviors in the experiment and intelligence, should these exist (the standard deviations are 3.32 in BASELINE treatment and 2.58 in PENALTY treatment). Kolmogorov-Smirnov tests fail to reject that the IQ scores of the subjects in each treatment and in the combined sample are distributed normally. The mean (median) IQ measure is 7.28 (7.0) for BASELINE treatment subjects and 8.35 (8.0) for those in the PENALTY treatment.³⁰

Political preferences

²⁷ Appendix Figure B.2.b shows the conditional contribution schedule and our classification of it for each subject, individually.

²⁸ When both slope of conditional contribution and average conditional contribution are included in a regression, there is a problem of multicollinearity and often neither is significant. We also performed estimates using the first principal component of the two variables and found that the results are almost indistinguishable from those using average conditional contribution alone.

²⁹ On the conditional contribution schedules of each subject in Appendix Figure B.2.b, we also provide the value of that subject's ACC.

³⁰ The difference in IQ between subjects in the two treatments is significant according to Mann-Whitney but not according to Kalmogorov-Smirnov tests.

The eight political preference questions asked at the end of the session are taken from the World Values Survey. The first, labeled General, asks the subject to place him/herself on a spectrum from 1 for Left to 10 for Right. The average response in our subject pool was 3.71. Other questions concern whether incomes should be made more equal or whether larger income differences are needed as incentives for individual effort (labeled Income); whether government ownership of business and industry should be increased or whether private ownership of business and industry should be increased (Ownership); whether government should take more responsibility to ensure that everyone is provided for or people should take more responsibility to provide for themselves (Responsibility); whether competition is harmful because it brings out the worst in people or is good because it stimulates hard work (Competition); whether hard work does not generally bring success or usually brings a better life (Hard Work); whether getting rich is always at the expense of others or wealth can grow so there's enough for everyone (Wealth); and whether everything is determined by fate versus people shaping their fate themselves (Destiny). Although a “more conservative” and “more liberal” answer might appear to exist for each question, responses to the different items are far from perfectly correlated, so they cannot be viewed simply as alternative measures of one underlying orientation. Average responses are found in part (a) of Appendix Table B.4.³¹ Correlations between the ACC, IQ, political attitude responses and Gender are also shown in the Appendix, in part (b) of Table B.4.

Subject-specific determinants of contributions

We begin our investigation of whether ACC, IQ, or political orientations are predictive of behaviors in our experiment with a further analysis of contribution decisions. The BASELINE

³¹ The table also displays the average response to each political preference variable by subject type. Note that for ease of interpretation, we redefine the response on the Ownership, Competition and Hard Work questions as 11 minus the subject's response, so that for all items, what is usually considered a “more conservative” response obtains a higher value.

treatment is a standard VCM not complicated by sanctioning schemes, and Table 4 reports a set of random-effects Tobit regressions in which a subject's contribution in a period in that treatment is the dependent variable and ACC, IQ, general political orientation, a female gender dummy and the period number (1,2,...,24) are entered as explanatory variables, with alternative specifications in which either an IQ*Period interaction term, an ACC*Period interaction term, or both are also included. A final specification includes also a General*Period interaction term.

Looking at the results, the coefficients on both ACC and IQ are always positive, both reaching weak statistical significance in the last two specifications, which include an ACC*Period interaction term and an IQ*Period interaction term. The estimated coefficients on those terms are negative and significant, at the 1% level for ACC*Period and the 10 or 5% level for IQ*Period, depending on specification. These results indicate that more conditionally cooperative subjects and higher-IQ subjects contributed more on average but also reduced their contributions more rapidly than others.³² Female fails to obtain a significant coefficient,³³ and period has a significant negative coefficient in the first two specifications that disappears with the addition of the interaction terms, suggesting that these account for most of the effect. Decline of contributions with repetition is of course a standard finding in VCM experiments lacking sanctioning opportunities or other added elements.

General political orientation obtains a significant negative coefficient when entered without a Period interaction in columns (2) and (3). While its coefficient becomes insignificant

³² The IQ-contribution association may result from the fact that more intelligent subjects are more aware of the possibility of fostering cooperation in early periods, given the fixed membership nature of grouping in our design. Page *et al.* (2005) comment on the mimicking of conditional cooperation by more sophisticated pay-off maximizers but have no measure of IQ with which to investigate their conjecture. A meta-analysis of prisoners' dilemma experiments by Jones (2008) finds that experiments conducted at schools whose student populations have higher SAT scores are significantly more cooperative. Burks *et al.* (2008) find a positive association between individuals' cognitive skills and their tendency to cooperate as first-movers in a sequential prisoners' dilemma game.

³³ A series of experiments have attempted to examine gender differences in the VCM; however, their results haven't generated a clear stylized result, and there exist results demonstrating significantly higher contributions by male subjects, significantly higher contributions by female subjects, and ones with insignificant gender differences. See Croson and Gneezy (2009) for a survey.

when we include the General*Period interaction in column (4), the interaction term itself is negative and significant. Together, these coefficients suggest that more politically conservative subjects contribute *less* to the public good, *ceteris paribus*, and that their contributions decay more quickly than those of other subjects.

[Table 4]

Table 5 reports the same regression specifications as Table 4 but using the pooled observations of both treatments for the initial period (columns (1) and (2)) and the initial four periods (columns (3) – (6), during all of which no penalty schemes existed in either treatment.³⁴ The results are qualitatively similar, with highly significant positive partial correlations of ACC and IQ with contribution in both the first period and the first phase as a whole. Apart from the greater significance of those two variables, the main difference from the Table 4 results is that General political orientation and its interaction with Period are insignificant. These results strengthen the conclusion that the conditional contribution schedule picks up an underlying willingness to cooperate and that more intelligent subjects are more likely to attempt to establish cooperation in early interactions, with the contributions of both high-ACC and high-IQ subjects decaying significantly more rapidly over time.

[Table 5]

We also investigate whether ACC, IQ, General political attitude and gender affected contributions when penalty schemes were in effect. In the regressions of Table 3, which we discussed above with reference to the effects of penalty parameters only, all specifications include a Female dummy, vote number, and round within phase. In regressions (2), (3), (5) and (6), General is also included, and in columns (3) and (6), we further add ACC and IQ score. As in Tables 4 and 5, gender obtains insignificant coefficients, but now the coefficients of IQ and

³⁴ Recall that the information of subjects in both treatments regarding the nature of the experiment was identical until after period 4.

General are also insignificant. The insignificant coefficient on ACC as well in the specification using “binding” (column (6)) suggests that personal characteristics are unimportant to contribution decisions when penalty schemes operate, although ACC does obtain a positive coefficient significant at the 5% level when we use *a/b* instead of “binding” (column (3)).

Result 5a. When the opportunity to vote on formal sanctions is absent, more conditionally cooperative and more intelligent subjects contribute more to the public good but reduce their contributions more rapidly. More politically conservative subjects contribute less and reduce their contributions more rapidly. Gender has no effect on contributions, after controlling for these factors.

Result 5b. When subjects have the opportunity to vote on formal sanctions, individual characteristics have qualitatively similar but weaker effects than when the choice of formal sanctions is unavailable. Penalty parameters are the main determinants of contributions.

Voting on PENALTY Schemes

Each subject had to vote first at each voting stage on whether contributions to the public account or to the private accounts should be penalized. Although 38 of the 40 votes by groups (95%) came out in favor of penalizing contributions to the private accounts, 10% of individual votes cast were for penalizing contributions to the public account. Do intelligence, cooperative orientation, and political views help to predict voting, as we conjectured? We estimated Random Effects Probit, Probit, and Logistic regressions, shown in Table 6, in which the dependent variable is one if a subject voted to penalize contributions to the private account, zero otherwise,

and explanatory variables are ACC, IQ, the vote number, gender, and in some specifications three political survey responses that contribute to the regressions' explanatory power.

In columns (1), (3), and (6), which don't include the survey responses, only four coefficients achieve significance at the 10% level or better. They are (a) one of the coefficients for vote number, which is consistently positive, weakly suggestive of learning, and (b) all three coefficients for Female, which are negative, indicating that females were significantly less likely to vote to penalize contributions to the private account, and vice versa for the public account. Of the political preference responses, we found those to the General political preference and the Income and Hard Work questions to be most helpful to explaining subject voting.³⁵ The influence of these variables on voting is indicated not only by the fact that all three obtain significant coefficients under each of the three regression forms and by the improved log likelihood and other summary statistics, but also by the fact that ACC, the coefficient on which was positive and insignificant in the simpler specifications, becomes significant at the 5 or 10% levels. The positive effect of ACC means that more conditionally cooperative subjects were more likely to vote for the efficient choice of penalizing contributions to the private accounts, supporting our conjecture that tastes for cooperation might influence voting on penalty scheme parameters even though an efficient penalty scheme is a rational choice for a strictly self-interested subject. The answer on the General question obtains negative coefficients, suggesting that subjects with more conservative political views voted less often for the efficient scheme, as we had conjectured might be the case. In contrast, the coefficients on the Income (Inequality needed?) and Hard Work variables obtain positive coefficients, indicating that not all dimensions of conservatism cut the same way—those believing that inequalities are needed to provide incentives and those believing that hard work produces rewards are less hesitant to favor

³⁵ Adding other survey responses changed results little.

penalizing private account contributions in order to achieve an efficient outcome. IQ obtains positive coefficients but these are insignificant in all specifications.

Interestingly, Female has more significant negative coefficients when the political attitude variables are included, meaning that female subjects were less likely to vote for the efficient option of fining private account contributions even though females show more cooperative orientations in our data (the simple correlation between gender and ACC is significant and positive). Because some votes to penalize contributions to the public account may be meant to discourage the imposition of fines altogether,³⁶ a possible explanation of the result could be a gender-biased dislike for penalties.

[Table 6]

Result 6. Subjects reporting more conservative general political attitudes are less likely to favor penalizing contributions to the private accounts, with political orientation having considerable explanatory power. Subjects with more pro-incentive orientations, belief in the importance of hard work for success, more cooperatively oriented subjects, and male subjects are more likely to favor penalizing such contributions.

Finally, we turn to the votes on the maximum penalty and the penalty-free contribution level. We estimated Tobit regressions to see whether individual votes on the maximum penalty a and on the penalty-exempt contribution level b can be explained by difference in cooperative orientation and intelligence, by gender, by learning or experience, and by political views.³⁷ We

³⁶ That is, subjects cannot know while casting their votes between the public and private option that an a of zero will be chosen in the impending vote on that ballot item, but they may believe the group more likely to select an a of zero if it is contributions to the public account that would be subject to penalty.

³⁷ The Tobit form is used here since the dependent variables each take a range of values having upper and lower bounds that are observed with some frequency. We do not show separate specifications that leave out the political variables to save space and because the other coefficients show less change from their inclusion than in Table 6. We note that the log likelihoods improve with the inclusion of the political variables, but by a smaller margin than in that table.

separately estimated one set of regressions for the very first vote, when there was no information on how others might vote or what the effect of the policy would be, and a second set for the remaining four votes (combined). Since group members knew what a had been chosen when selecting b , the chosen a was added as a control in the regressions for b .³⁸ For the regression studying votes 2 – 5, the last voted values of both a and b are controlled, and random effects estimators are used.³⁹ Observations for the 4th and 5th votes by members of Group 13 are left out since they followed decisions to penalize contributions to the public good, making the impacts of the two parameters (a and b) and the incentives to vote for high or low values of each incomparable to those facing other groups. A substitution of one of the three political preference variables in the Table 6 regressions was found to add more explanatory power.

Higher conditional willingness to cooperate, measured by ACC, is a significant positive predictor of preferred maximum penalty, a , but not of the penalty-exempt contribution, b . IQ has a significant positive effect on 1st but not later votes on a and also shows no significant effects on choice of b . Cooperative orientation and intelligence thus seem to increase voting for more efficient penalty levels in the same manner as they increase contributions to the public account, but the effect is absent with respect to the exemption level.

The coefficients on last-voted a and b indicate persistence and/or consistency where they are significant, i.e. (i) a higher a (higher b) in the previous phase is associated with a higher a (higher b) now, according to estimates (3) and (4); (ii) for votes 2 - 5 (estimate (4)), groups that select more binding penalty levels a in the current period also select higher b 's, usually levels that cause penalties to be assessed unless the whole or almost the whole endowment is contributed to the public account.

³⁸ This structural model has the potential problem of endogeneity since the chosen a is correlated with the error term, and actually we confirmed the high correlation. As one possible resolution to avoid this econometric problem, we also estimated another variant of regressions (reduced form) including only vote number variables instead of chosen a , and we obtained qualitatively the same result. In our paper, we report only the result mentioned in the text.

³⁹ The remarks in footnote 20 apply here as well.

The coefficients on the General political preference and Income (Inequality needed?) variables have roughly the same pattern as in the regressions for the vote on what to penalize, with more generally conservative subjects favoring significantly lower penalties and smaller exemptions (allowing larger contributions to the private accounts), but those declaring stronger beliefs in the necessity of incentives and inequality favoring stronger penalties and higher exemption levels. For these two variables, five of eight coefficients are significant at the 5% level, including one significant at the 1% level. The Destiny variable, taking a higher value the stronger the belief that people can shape their own destinies, has positive coefficients and adds to goodness of fit for the 1st vote regressions, but it is itself insignificant in those regressions and it makes no contribution and is thus left out entirely in the regressions for 2nd to 5th votes. Female again obtains negative, significant coefficients for the votes on both a and b , when the political attitude values are controlled for, suggesting that male subjects were more likely to vote for more efficient level of a as well as b , perhaps again because females tend to look less favorably on the use of strong penalties.

[Table 7]

Result 7. Subjects with more conservative general political views are less likely to vote for high fines and penalty-exemption levels, while subjects with more pro-incentive orientations and male subjects are more likely to vote for such levels of the two parameters. Cooperative orientation and IQ are positively associated with choice of maximum penalty.

We conclude with a brief discussion of an issue raised earlier, in our discussion of Table 3. In this section, we've used regression analysis to examine first the determinants of subjects' contribution choices and then their votes on penalty scheme parameters. But if individual characteristics explain both contributions and votes for penalty parameters, it may be more

appropriate to treat the penalty rate (a/b) and exemption level (b) in the analysis of contributions as endogenous rather than exogenous variables.

It turns out to be difficult to address this issue, because the regressions explaining voting include observations both of votes for the winning median values and votes for other values of a and b . Possible responses include (a) switching the analysis to group level observations and treating the penalty parameters as endogenous in that setting, (b) conducting individual-level two-stage least squares analysis with the observations of only those individuals who voted for the winning parameters of each phase, and (c) treating group penalty parameters as exogenous but adding to the contribution regressions also the parameter values for which the individual voted, which can be treated as endogenous. Approach (b) is closest to the spirit of our analysis above, and is helped by the fact that the large majority of individual votes, about 63%, were for the winning parameter values, so that sample size falls by only about a third when applying this method. However, when only the two survey measures used in Table 7 but not Table 3 are used, those measures fail standard instruments tests and few or no coefficients can be estimated with precision in the 2nd stage regressions. Adding a third, fourth, or even all seven survey measures (apart from General political preference) improves the estimates of 2nd stage regression coefficients, including those on b and a/b which become positive and significant as in Table 3. But the Wu-Hausman and Durbin-Wu-Hausman tests still fail to reject the null hypothesis that those variables are exogenous, with p -values well above 10% (see Appendix Table B.5).⁴⁰

⁴⁰ We also performed estimates following approach (c), adding two new terms—preferred b of individual i and preferred a/b of individual i —to specifications (1) – (3) of Table 3, using the full set of individual observations. (For a/b we use the value of a selected by the group in the previous step, since i selects b knowing that group choice.) In Tobit estimates, the individual preferred values obtain insignificant coefficients while the group-selected values of b and a/b obtain positive significant coefficients. 2SLS estimates instrumenting the individually preferred values with up to seven survey response measures produced few significant coefficients apart from those on group-selected b and a/b in some versions of the column (2) specification, and Wu-Hausman and Durbin-Wu-Hausman tests failed to reject the null hypothesis that the instrumented variables are exogenous.

4. Summary and Conclusion

We conducted two treatments, one a simple and one a modified voluntary contribution mechanism experiment, each having 40 student subjects playing in five person partner groups. In one treatment, PENALTY, group members played a standard VCM for four periods and then were asked five times, at intervals, to vote on three parameters defining a scheme to penalize either contributions to their private accounts or those to the public account. The other treatment, BASELINE, is a standard VCM, adding only periodic breaks at the points where voting occurred in the PENALTY treatment. In the PENALTY treatment, groups almost always (in 38 of the 40 group votes) chose to penalize contributions to the private accounts, and this choice was efficiency-enhancing. Efficiency-enhancement is demonstrated both within and between treatments. In 30 of the 38 instances in which groups voted to penalize contributions to private accounts, the group in question achieved higher contributions (in the same 30 cases, also higher earnings) under it than in the initial, penalty-free periods. In 33 of the 38 instances, the group achieved higher contributions (in the same 33 cases, also higher earnings) than did the average group in the BASELINE treatment during the same periods of the experiment.

Our results indicate that subjects were able to devise efficient or nearly efficient solutions to a simple collective action dilemma when presented with basic tools that could be used to effectively resolve but just as easily to exacerbate that dilemma, and with no guidance on a proper course of action and no possibility of communication. Individual confusion and some departures from efficient voting existed, but those tendencies were usually overridden at the group level.

In addition to the core contributions and voting portions of our experiment, we elicited proxies for conditional willingness to cooperate, intelligence, and political orientation. We found that not only do the cooperativeness data have predictive power with respect to

contributions in the public goods decisions, but also cognitive ability and political orientation, not previously considered in this context, help to predict these decisions. As expected, more conditionally cooperative subjects initially contributed more but also reduced their contributions more rapidly over time, in the BASELINE and penalty-free periods of the PENALTY treatment. An interesting finding is that while IQ and conditional cooperation are negatively correlated in our subject pool, higher IQ subjects also initially contribute more and then reduce contributions more rapidly, perhaps because they are more likely to see strategic value in encouraging contributions in early periods. Subjects who self-classified as more politically conservative tended to both contribute less and to reduce their contributions more rapidly.

The main purpose of our study was to see whether individuals provided with no guidance on the matter and no opportunity to communicate would recognize the nature of the social dilemma posed by a voluntary contribution mechanism and would fashion three available parameters, by voting, into a penalty scheme capable of efficiently resolving the dilemma. We were interested not only in the implicit cognitive challenge posed by our PENALTY treatment, but also in the possibility that despite the fact that the efficient parameters are in the interest of all who focus on maximizing their money payoffs, other preferences might also influence voting. Our main findings are that most subject groups were able to select efficient parameters after short learning periods. Nonetheless, preferences besides those for money earnings do appear to have influenced voting.

Consistent with our anticipation of a cognitive challenge, more intelligent subjects tended to vote for more efficient parameter settings, but the difference was only significant in the first vote for the maximum penalty. With respect to tastes, we conjectured that those with tastes for cooperation, indicated by greater conditional willingness to contribute to the public good, would be more likely to vote for efficient penalty parameters, much as high contributors are the

most likely punishers of free riders in informal sanction experiments. This conjecture was largely supported.

We also wondered whether subjects' ideological orientations, which might touch on freedom from interventions or penalties, would help to explain their votes. This turned out to be so. Despite the availability of an intervention capable of resolving the free-riding problem inherent in the VCM, more politically conservative subjects voted significantly less often for efficient penalty parameters. Strength of belief in the need for material inequalities and incentives was also associated with voting for efficient penalty variables, even though such beliefs are sometimes associated with conservatism. Interestingly, female subjects, although more cooperative in the VCM and conditional cooperation task, voted for significantly less efficient penalty parameters after controlling for these other factors, perhaps because of a correlation between female gender and reluctance to impose penalties.⁴¹

Our study extends the recently opened field of investigation of choice of sanctioning institutions in collective action dilemmas from the domain of informal sanctions to that of formal ones. As pointed out at the outset, formal sanctions may be preferred to informal ones in many settings because of their greater predictability and the greater control they afford over anti-social (or for that matter confused) individuals. Ours is the first study in which subjects are not simply offered the choice of a well-designed sanction scheme, but have to construct one for themselves by voting on simple components without communication.

Directions for future research are numerous. One obvious extension would be to investigate choices between formal and informal sanctions when both are on offer. Another

⁴¹ Although our findings are based on the decisions of just 80 undergraduate subjects at one university, they are in most respects consistent with findings in pilot sessions conducted with 70 subjects at the University of Copenhagen in early 2007. There, too, most groups learned quickly to choose efficiency-inducing parameters in the PENALTY treatment, whereas contributions showed the usual decaying trend in the BASELINE treatment. There, too, expressed political views were the strongest predictors of voting, with subjects showing more conservative political orientations being significantly less likely to vote for penalizing contributions to the private accounts and for setting high penalty rates.

would be to investigate more complex choice environments, for example ones in which the social optimum lies at the interior rather than the boundary of the choice set. Third, either sanctioning or voting or both could be made costly. The last possibility links up to the critical point that whereas in our experiment, the channels and implementation of democratic rule are granted to subjects as a free gift of the experimenters, in the real world such institutions must be constructed and renewed continually by organizations and societies that are to function effectively. It is also related to the realization that both informal and formal sanctions have costs and benefits that require careful assessment when designing actual institutions.

Acknowledgement: We thank Helene Bohn-Jespersen, Thomas Alexander Stephens and Ken Ettinger for effective research assistance. We thank the Danish research council (FSE) for financial support under project title “Cooperation and Institutions.”

References

- Andreoni, James, 1988, "Why Free Ride? Strategies and Learning in Public Goods Experiments," *Journal of Public Economics* 37: 291-304.
- Bochet, Olivier, Talbot Page and Louis Putterman, 2006, "Communication and Punishment in Voluntary Contribution Experiments," *Journal of Economic Behavior and Organization* 60: 11-26.
- Botelho, Anabela, Glenn Harrison, Ligia M. Costa Pinto and Elisabet E. Rutström, 2005, "Social Norms and Social Choice," unpublished paper, Dept. of Economics, University of Central Florida.
- Brosig, Jeannette, Axel Ockenfels and Joachim Weimann, 2003, "The Effect of Communication Media on Cooperation," *German Economic Review* 4, 217-42.
- Buchanan, James M. and Gordon Tullock, 1965, *The Calculus of Consent*. University of Michigan Press: Ann Arbor.
- Burks, Stephen V., Jeffrey P. Carpenter, Lorenz Goette and Aldo Rustichini, 2008, "Cognitive Skills Affect Economic Performance, Strategic Behavior, and Job Attachment," *Proceedings of the National Academy of Sciences* 106 (19): 7745-50.
- Chen, Yan and Charles R. Plott, 1996, "The Groves-Ledyard Mechanism: An Experimental Study of Institutional Design," *Journal of Public Economics* 59: 335-64.
- Cinyabugama, Matthias, Talbot Page and Louis Putterman, 2006, "Can Second-Order Punishment Deter Perverse Punishment?" *Experimental Economics* 9: 265-79.
- Croson, Rachel and Uri Gneezy, 2009, "Gender Differences in Preferences," *Journal of Economic Literature* 47 (2): 448-74.
- Damasio, Antonio, 1994, *Descartes' Error: Emotion, Reason, and the Human Brain*. New York: Putnam.
- Egas, Martijn and Arno Riedl, 2005, "The Economics of Altruistic Punishment and the Demise of Cooperation," Tinbergen Institute Discussion Paper No. 05-065/1.
- Ertan, Arhan, Talbot Page and Louis Putterman, 2009, "Can Endogenously Chosen Institutions Mitigate the Free-Rider Problem and Reduce Perverse Punishment?" *European Economic Review* 53: 495-511.
- Falk, Armin, Ernst Fehr and Urs Fischbacher, 2005, "Driving Forces Behind Informal Sanctions," *Econometrica* 73 (6): 2017-30.

- Falkinger, Josef, Ernst Fehr, Simon Gächter and Rudolf Winter-Ebmer, 2000, "A Simple Mechanism for the Efficient Provision of Public Goods: Experimental Evidence." *American Economic Review* 90 (1): 247-64.
- Fehr, Ernst and Simon Gächter, 2000, "Cooperation and Punishment," *American Economic Review* 90 (4): 980-94.
- Fehr, Ernst and Simon Gächter, 2002, "Altruistic Punishment in Humans," *Nature* 415: 137-40.
- Fischbacher, Urs, 2007, "z-Tree: Zurich Toolbox for Ready-made Economic Experiments," *Experimental Economics* 10 (2): 171-8.
- Fischbacher, Urs and Simon Gächter, forthcoming, "Social Preferences, Beliefs and the Dynamics of Free Riding in Public Good Experiments," *American Economic Review* (in press).
- Fischbacher, Urs, Simon Gächter, and Ernst Fehr, 2001, "Are People Conditionally Cooperative? Evidence from a Public Goods Experiment," *Economics Letters* 71 (3): 397-404.
- Gunthorsdottir, Anna, Daniel Houser and Kevin McCabe, 2007, "Disposition, History and Contributions in Public Goods Experiments," *Journal of Economic Behavior and Organization* 62 (2): 304-15.
- Güerke, Özgür, B. Irlenbusch and B. Rockenbach, 2006, "The Competitive Advantage of Sanctioning Institutions," *Science* 312: 108-110, April 7 2006.
- Herrmann, Benedikt, Christian Thöni and Simon Gächter, 2008, "Antisocial Punishment Across Societies," *Science* 319: 1362-7.
- Houser, Daniel and Robert Kurzban, 2002, "Revisiting Kindness and Confusion in Public Goods Experiments," *American Economic Review* 92 (4): 1062-9.
- Isaac, R. Mark and James M. Walker, 1988, "Group Size Effects in Public Goods Provision: The Voluntary Contribution Mechanism," *Quarterly Journal of Economics* 53: 179-200.
- Isaac, R. Mark, James M. Walker and Susan H. Thomas, 1984, "Divergent Evidence on Free Riding: An Experimental Examination of Possible Explanation," *Public Choice* 43 (2): 113-49.
- Isaac, R. Mark, James M. Walker and Arlington W. Williams, 1994, "Group Size and Voluntary Provision of Public Goods: Experimental Evidence Utilizing Large Groups," *Journal of Public Economics* 54: 1-36.
- Jones, Garrett, 2008, "Are Smarter Groups More Cooperative? Evidence from Prisoner's Dilemma Experiments, 1959-2003," *Journal of Economic Behavior and Organization* 68: 489-97.
- Kosfeld, Michael, Akira Okada, and Arno Riedl, 2008, "Institution Formation in Public Goods Games," *American Economic Review* 99 (4): 1335-55.

- Ledyard, John, 1995, "Public Goods: a Survey of Experimental Research," pp. 111-194 in Kagel, J., Roth, A.E. (Eds.), *Handbook of Experimental Economics*. Princeton, NJ: Princeton University Press.
- Nikiforakis, Nikos, 2008, "Punishment and Counter-punishment in Public Goods Games: Can we Really Govern Ourselves?" *Journal of Public Economics* 92: 91-112.
- Nikiforakis, Nikos and Hans-Theo Normann, 2008, "A Comparative Statics Analysis of Punishment in Public Good Experiments," *Experimental Economics* (in press).
- Ones, Umut and Louis Putterman, 2007, "The Ecology of Collective Action: A Public Goods and Sanctions Experiment with Controlled Group Formation," *Journal of Economic Behavior and Organization* 62: 495-521.
- Ostrom, Elinor, "Collective Action and the Evolution of Social Norms," *Journal of Economic Perspectives* 14 (3): 137-58.
- Page, Talbot, Louis Putterman and Bulent Unel, 2005, "Voluntary Association in Public Goods Experiments: Reciprocity, Mimicry, and Efficiency," *Economic Journal* 115: 1032-53.
- Selten, Reinhard, "Re-examination of the Perfectness Concept for Equilibrium Points in Extensive Games," *International Journal of Game Theory* 4: 25 – 55.
- Séréville, Chantal and Bernard Myers, 1994, *Nøglen til Success I alle IQ-tester [The Key to Success in all IQ Tests (Danish)]*. Viborg: Nørhavn A/S.
- Sutter, Matthias, Stefan Haigner, and Martin Kocher, forthcoming, "Choosing the Carrot or the Stick? – Endogenous Institutional Choice in Social Dilemma Situations" *Review of Economic Studies* (in press).
- Tyran, Jean-Robert and Lars P. Feld, 2006, "Achieving Compliance when Legal Sanctions are Non-deterrent," *Scandinavian Journal of Economics* 108 (1): 135-56.

Figure 1: The penalty structure

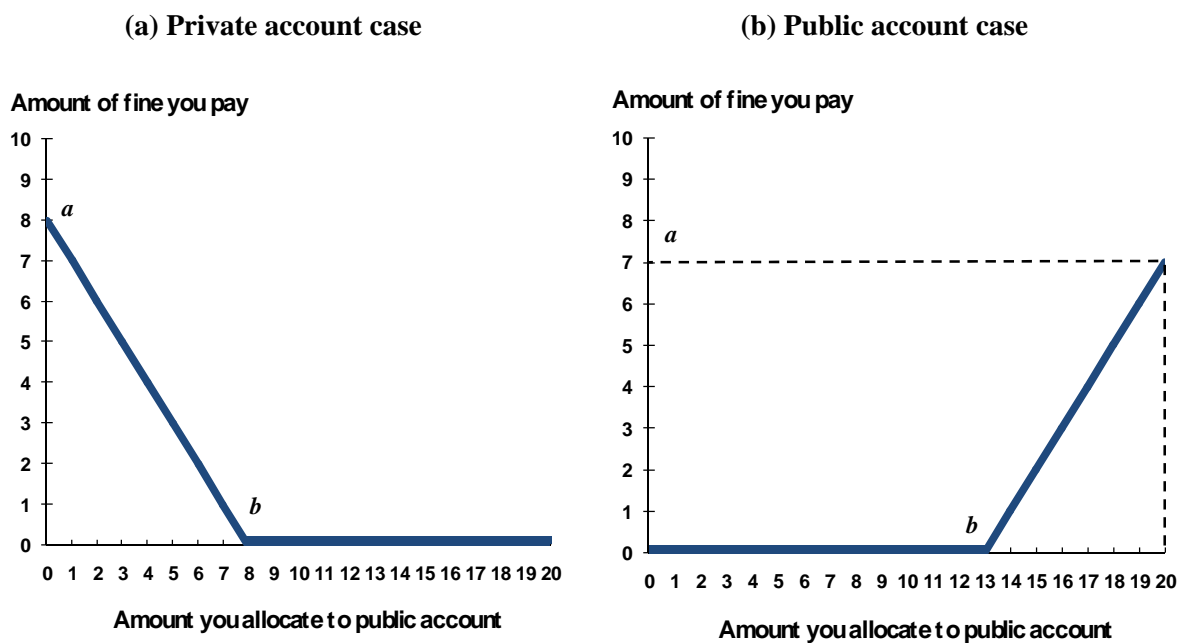


Table 1: Summary of treatments, sessions, and subjects

Treatment	Total number of sessions	Total number of groups	Group numbers	Total number of subjects
BASELINE treatment	2	8	1 – 8	40
PENALTY treatment	2	8	9 – 16	40
Experiment as a whole	4	16		80

Figure 2: Experimental Design

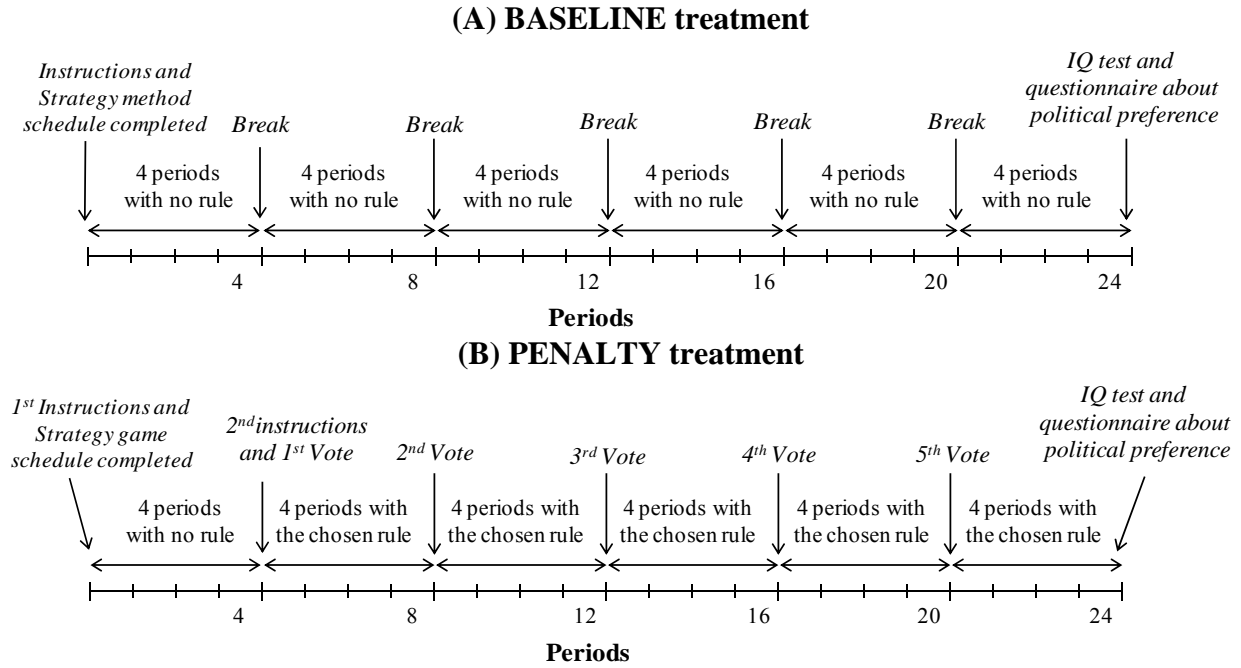
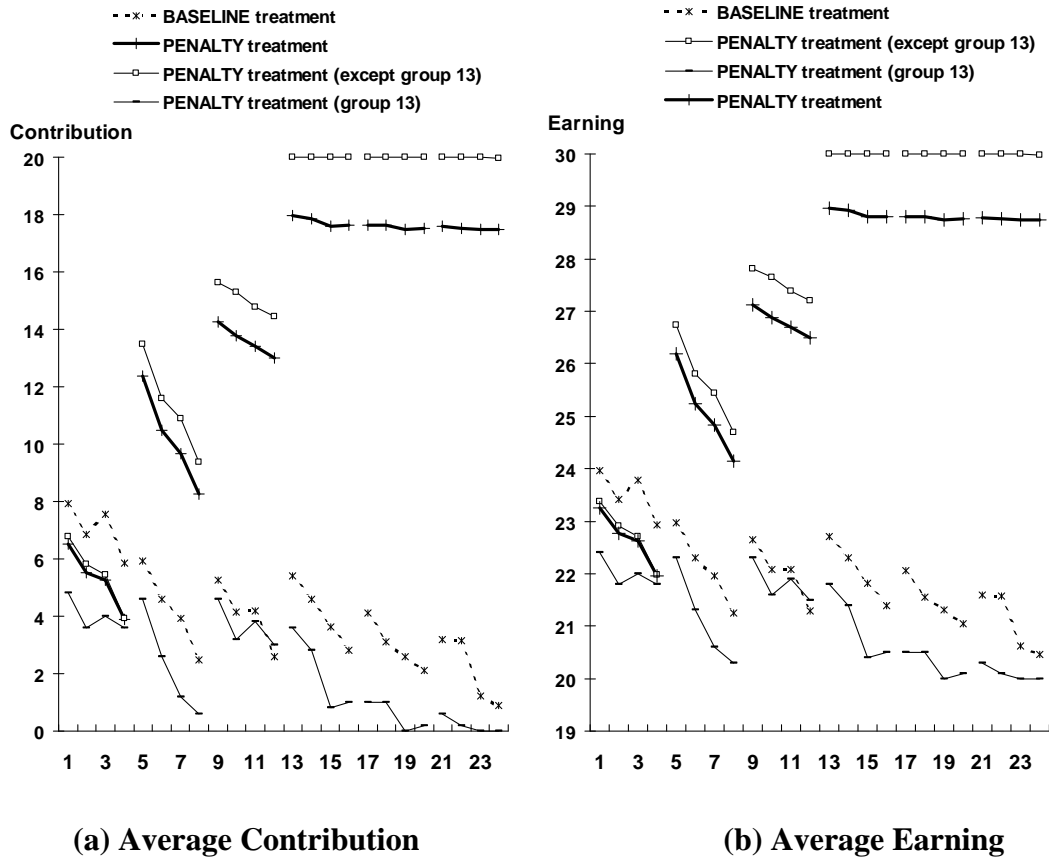


Figure 3: The trends of average contribution to the public account and average earnings



Note: The earning in period 1 in Figure 3(b) does not include the earning in the extra period between periods 1 and 2.

Table 2: Results of voting choice (account to be fined, a , b and a/b)

Period	Group 9				Group 10				Group 11				Group 12			
	G or P ¹	a	b	a/b	G or P	a	b	a/b	G or P	a	b	a/b	G or P	a	b	a/b
5	P	14	20	0.70	P	12	0	infinity	P	5	20	0.25	P	10	18	0.56
9	P	20	20	1.00	P	20	20	1.00	P	3	15	0.20	P	20	20	1.00
13	P	20	20	1.00	P	20	20	1.00	P	20	20	1.00	P	20	20	1.00
17	P	20	20	1.00	P	20	20	1.00	P	20	20	1.00	P	20	20	1.00
21	P	20	20	1.00	P	20	20	1.00	P	20	20	1.00	P	20	20	1.00

Period	Group 13				Group 14				Group 15				Group 16			
	G or P	a	b	a/b	G or P	a	b	a/b	G or P	a	b	a/b	G or P	a	b	a/b
5	P	2	10	0.20	P	20	0	infinity	P	5	3	1.67	P	20	20	1.00
9	P	5	10	0.50	P	20	15	1.33	P	8	3	2.67	P	20	20	1.00
13	P	3	15	0.20	P	20	20	1.00	P	20	20	1.00	P	20	20	1.00
17	G	6	5	-0.40 ²	P	20	20	1.00	P	20	20	1.00	P	20	20	1.00
21	G	10	5	-0.67 ²	P	20	20	1.00	P	20	20	1.00	P	20	20	1.00

- Notes:
1. G and P refer to the public account (G) and the private account (P) respectively.
 2. Because a/b is the fine per point contributed to the private account when private contributions are fined, in this case we list $-a/(20 - b)$, where $a/(20 - b)$ is the fine that would be paid for contributing a point to the *public* account (see equation (2)).

Table 3: Determinants of contributions in PENALTY treatment

Independent variable	(1)	(2)	(3)	(4)	(5)	(6)
Per unit penalty a/b	10.9*** (1.98)	11.0*** (2.01)	11.3*** (2.00)	---	---	---
“Binding”	---	---	---	23.3*** (2.80)	23.4*** (2.82)	23.1*** (2.78)
Exemption level b	2.38*** (0.22)	2.39*** (0.22)	2.52*** (0.23)	1.70*** (0.13)	1.69*** (0.13)	1.72*** (0.14)
Conditional Contribution	---	---	0.79** (0.32)	---	---	0.49 (0.36)
IQ	---	---	0.24 (0.49)	---	---	-0.39 (0.59)
General (Political Preference)	---	-0.23 (0.69)	-0.33 (0.65)	---	-0.37 (0.76)	-0.31 (0.74)
Female	-0.43 (2.42)	-0.44 (2.39)	-1.48 (2.33)	2.95 (2.76)	2.87 (2.76)	1.29 (2.84)
Vote number	3.19*** (0.69)	3.17*** (0.69)	2.98*** (0.69)	2.29*** (0.67)	2.30*** (0.67)	2.25*** (0.67)
Round within phase	-1.34*** (0.47)	-1.34*** (0.47)	-1.33*** (0.47)	-1.86*** (0.46)	-1.85*** (0.46)	-1.86*** (0.46)
Constant	-30.4*** (4.95)	-29.7*** (5.32)	-36.6*** (6.99)	-22.6*** (3.42)	-21.2*** (4.45)	-19.9*** (6.70)
Number of observations	720	720	720	760	760	760
Log likelihood	-559.6	-559.5	-556.4	-576.3	-576.1	-574.8
Wald χ^2	197.1	198.3	200.6	262.4	263.2	262.9
Prob > χ^2	0.000	0.000	0.000	0.000	0.000	0.000

- Notes:
1. Random effects Tobit regressions.
 2. *** indicate significance at the .01 level.
 3. Vote number $\in \{1, 2, 3, 4, 5\}$ and Round within phase $\in \{1, 2, 3, 4\}$.
 4. The dummy variable “binding” equals 1 when $a/b \geq 0.7$ and 0 otherwise.
 5. In columns (1) to (3), observations in which $b = 0$ are excluded since then $a/b = \infty$.
 6. For numbers of left and right censored observations, see the Appendix.

Table 4: Determinants of contributions in BASELINE treatment

Independent variable	(1)	(2)	(3)	(4)
Conditional Contribution	0.21 (0.25)	0.15 (0.23)	0.42* (0.25)	0.45* (0.25)
IQ	---	0.42 (0.34)	0.66* (0.37)	0.71* (0.37)
General (Political Orientation)	---	-1.58** (0.76)	-1.59** (0.75)	-0.23 (0.83)
Female	---	2.77 (2.45)	2.82 (2.42)	2.85 (2.45)
Period	-0.47*** (0.04)	-0.47*** (0.04)	-0.17 (0.13)	0.34* (0.18)
ACC*Period	---	---	-0.03*** (0.01)	-0.03*** (0.01)
IQ*Period	---	---	-0.02* (0.01)	-0.03** (0.01)
General*Period	---	---	---	-0.13*** (0.03)
Constant	4.11** (1.85)	5.95 (4.54)	2.72 (4.67)	-2.48 (4.89)
Number of observations	960	960	960	960
Log likelihood	-1821	-1816	-1812	-1803
Wald chi ²	113.9	121.4	130.4	136.7
Prob > chi ²	0.000	0.000	0.000	0.000

Notes:

1. Table shows random effects Tobit regressions.
2. *, ** and *** indicate significance at the .10 level, at the .05 level and the .01 level, respectively.
3. The number in the parenthesis is the standard deviation.
4. Period $\in \{1, 2, \dots, 24\}$.
5. For numbers of left and right censored observations, see the Appendix.

Table 5: Determinants of contributions in first period and first phase, both treatments

Independent variable	(1)	(2)	(3)	(4)	(5)	(6)
Conditional Contribution	1.19*** (0.26)	1.36*** (0.26)	0.77*** (0.20)	0.81*** (0.20)	1.34*** (0.28)	1.35*** (0.28)
IQ	---	1.05*** (0.35)	---	0.60** (0.29)	1.35*** (0.40)	1.37*** (0.41)
General (Political Orientation)	---	-0.61 (0.63)	---	-0.097 (0.51)	-0.10 (0.51)	-0.49 (0.72)
Female	---	-2.35 (2.15)	---	0.66 (1.78)	0.63 (1.78)	0.62 (1.78)
Period	---	---	-0.94*** (0.34)	-0.94*** (0.34)	2.61*** (1.10)	2.13* (1.27)
ACC*Period	---	---	---	---	-0.22*** (0.08)	-0.22*** (0.08)
IQ*Period	---	---	---	---	-0.30*** (0.11)	-0.31*** (0.11)
General * Period	---	---	---	---	---	0.15 (0.20)
Constant	-0.20 (1.80)	-5.81 (4.15)	2.60 (1.59)	-2.18 (3.48)	-11.0** (4.40)	-9.79** (4.68)
Number of observations	80	80	320	320	320	320
Log likelihood	-199.0	-193.5	-763.9	-761.7	-755.5	-755.2
Wald chi ²	20.8	32.0	22.9	27.8	40.0	40.5
Prob > chi ²	0.000	0.000	0.000	0.000	0.000	0.000

Notes:

1. Random effects Tobit regressions.
2. Observations of first period only, columns (1) and (2), and of periods 1 – 4, columns (3) – (6).
3. *, ** and *** indicate significance at the .10 level, at the .05 level and the .01 level, respectively.
4. The figure in the parenthesis is standard deviation.
5. Period $\in \{1, 2, 3, 4\}$.
6. For numbers of left and right censored observations, see the Appendix.

Table 6: Determinants of voting to penalize contributions to private vs. public account

Independent variable	Random Effects Probit		Probit		Logistic	
	(1)	(2)	(3)	(4)	(5)	(6)
Conditional Contribution	0.08 (0.07)	0.12* (0.07)	0.05 (0.04)	0.10** (0.04)	0.10 (0.07)	0.21** (0.08)
IQ	0.13 (0.12)	0.12 (0.10)	0.05 (0.06)	0.08 (0.06)	0.08 (0.10)	0.14 (0.12)
General (Political Orientation)	---	0.39*** (0.14)	---	0.33*** (0.08)	---	0.61*** (0.15)
Income (Inequality needed?)	---	0.22** (0.10)	---	0.19*** (0.07)	---	0.38*** (0.13)
Hard Work	---	0.19* (0.11)	---	0.15** (0.07)	---	0.27** (0.12)
Female	-0.97* (0.58)	-1.44** (0.61)	-0.62** (0.27)	-1.11*** (0.37)	-1.22** (0.55)	-2.03*** (0.69)
Vote number	0.18* (0.11)	0.17 (0.11)	0.13 (0.09)	0.14 (0.09)	0.24 (0.17)	0.28 (0.18)
Constant	0.41 (1.15)	-0.46 (1.10)	0.58 (0.61)	-0.27 (0.75)	1.05 (1.13)	-0.72 (1.33)
Number of observations	200	200	200	200	200	200
Log likelihood	-57.6	-52.1	-64.3	-54.1	-60.3	-54.2
Wald chi ²	6.34	12.42	---	---	---	---
LR chi ²	---	---	10.1	30.46	10.0	30.14
Prob > chi ²	0.175	0.087	0.039	0.000	0.040	0.000
Pseudo R ²	---	---	0.073	0.220	0.072	0.217

Notes:

1. The dependent variable equals 1 if the subject voted to penalize contributions to the private accounts, 0 if s/he voted to penalize contributions to the public account.
2. *, **, and *** indicate significance at the .10 level, at the 0.05 level and at the .01 level, respectively.
3. The figure in the parenthesis is standard deviation.
4. Vote number $\in \{1, 2, 3, 4, 5\}$

Table 7: Determinants of votes on *a* and *b*

Independent Variable	Initial vote		2nd to 5th vote	
	(1) Vote on <i>a</i>	(2) Vote on <i>b</i>	(3) Vote on <i>a</i>	(4) Vote on <i>b</i>
Conditional Contribution	2.71*** (0.92)	-0.59 (1.04)	1.14* (0.66)	0.13 (0.45)
IQ	2.75** (1.27)	-0.30 (1.55)	0.40 (1.03)	-0.68 (0.73)
(Last-voted group <i>a</i>)	---	-0.52 (0.60)	1.71*** (0.37)	1.13*** (0.26)
Previous group <i>b</i>	---	---	0.31 (0.31)	1.06*** (0.24)
General (Political Orientation)	-3.26** (1.61)	-3.84 (2.46)	-3.92*** (1.39)	-1.86** (0.91)
Income (Inequality needed?)	2.07 (1.25)	2.79 (1.86)	2.33** (1.12)	1.65** (0.78)
Destiny	1.15 (1.72)	3.42 (2.47)	---	---
Female	-17.0** (6.46)	-17.4* (8.90)	-14.3*** (5.55)	-8.01** (3.83)
Constant	-19.3 (17.2)	7.86 (22.3)	5.21 (12.0)	3.85 (8.80)
Number of observations	40	40	150	150
Log likelihood	-79.8	-77.6	-129.0	-118.7
LR chi ²	20.32	9.28	---	---
Wald chi ²	---	---	33.45	37.91
Prob > chi ²	0.002	0.233	0.000	0.000
Pseudo R ²	0.113	0.056	---	---

- Notes: 1. Tobit regression (columns (1) and (2)) and Tobit regression with random effects (columns (3) and (4)).
2. Number in parenthesis is standard deviation.
3. * and ** indicate significance at the .10 level and the .05 level, respectively.
4. In regression (3), *a* is the value chosen in the previous vote, while in regressions (2) and (4), *a* is the value selected in the previous voting step of the current vote process.
5. For numbers of left and right censored observations, see the Appendix.