Voting on Thresholds for Public Goods: Experimental Evidence

Julian Rauchdobler, Rupert Sausgruber and Jean-Robert Tyran *

February 2010

Abstract

Introducing a threshold in the sense of a minimal project size transforms a public goods game with an inefficient equilibrium into a coordination game with a set of Pareto-superior equilibria. Thresholds may therefore improve efficiency in the voluntary provision of public goods. In our one-shot experiment, we find that coordination often fails and exogenously imposed thresholds are ineffective at best and often counter-productive. This holds under a range of threshold levels and refund rates. We test whether thresholds perform better if they are endogenously chosen, i.e. whether a threshold is approved in a referendum, because voting may facilitate coordination due to signaling and commitment effects. We find that voting does have signaling and commitment effects but they are not strong enough to significantly improve the efficiency of thresholds.

Keywords: Provision of public goods, threshold, voting, experiments

JEL classification: H41, D72, C92

*Rauchdobler: University of Innsbruck, Department of Public Economics, Universitätsstr. 15, A-6020 Innsbruck. E-mail: Julian.Rauchdobler@uibk.ac.at. Sausgruber: University of Innsbruck, Department of Public Economics, Universitätsstr. 15, A-6020 Innsbruck. E-mail: Rupert.Sausgruber@uibk.ac.at. Tyran: University of Copenhagen, Department of Economics, Øster Farimagsgade 5;, DK-1353 Copenhagen K. E-mail: Jean-Robert.Tyran@econ.ku.dk. We thank Charles Noussair, Rudolf Kerschbamer, and Wolfgang Höchtl for helpful comments. We are grateful for financial support by the Austrian Science Fund (FWF) under Projects No. 17029 and S10307-G14 and to the Danish Science Foundation FSE.

1 Introduction

In some cases, public goods have a threshold, i.e. a minimal project size, for technological reasons - building half a bridge does not make much sense - but in other cases, the public good can be provided continuously - think of donations to start a community library. This paper investigates whether introducing a threshold to a public good that does not have a specific threshold value for technological reasons can increase efficiency.[1] More specifically, we ask whether such thresholds are more effective if they have been approved in a referendum rather than just imposed by some authority.

The issue we investigate is of potential relevance in both the small and the large. For example, consider a fundraising drive to start the community library. Will citizens donate more if the charity commits to a minimal size of the library than if the fundraising drive is started without any lower boundary for the size of the library? To provide an example from the other end of the scale, consider international agreements to abate greenhouse gases. Would nations be willing to contribute more to prevent global warming if the international agency commits to a minimum target for abatement? Are nations more likely to reach the threshold if they approved of the threshold in a referendum than if the threshold was imposed by some international organization? These examples illustrate that our study refers to situations in which potential contributors to the public good can be identified and, therefore, contributors can vote on the threshold but the central authority is weak and does not have the power to enforce contributions to the public good. In the examples above, the charity cannot force citizens to donate, the international organization cannot force sovereign states to reduce greenhouse gases. The mechanism we investigate is therefore a potential remedy to the free rider problem when no strong enforcement institution exists and when the players keep the full decision sovereignty on whether or not to contribute to the public good. All that is required is that an institution can credibly commit not to provide the public good if voluntary contributions are insufficient to meet the threshold.

The theoretical rationale for expecting that introducing a threshold increases efficiency is that a threshold transforms a public goods game with a unique and inefficient equilibrium into a coordination game with a set of additional equilibria in which all players are better off (Bagnoli and Lipman 1989). The intuition for why this transformation may increase efficiency is that a player's contribution can be essential for whether the public good is provided at all.[2] If a single player makes the difference between providing the

[1] Threshold public goods are also called step-level, binary, discrete, lumpy, or provision point goods.

[2] Note that we investigate the effect of introducing a threshold in the sense of the minimal project size. In a situation without a threshold, the provision of the public good is continuous and proportional to the aggregate contributions. In contrast, in a situation with a threshold, the public good is provided if contributions meet or exceed the threshold but is not provided at all if the threshold is not met. If

public good at the threshold level versus none at all, he may find it in his best interest to contribute. However, this situation typically occurs if others make contributions such that the threshold is "almost" reached, i.e. if it is feasible (and not too costly) for a particular player to "make the difference". If others already contribute more than the threshold (and the public good is provided anyway) or if they contribute too little for a single player to make the difference between provision or not (i.e. the public good is not provided anyway), free riding is optimal for a self-interested player. Thus, thresholds only improve incentives to cooperate if players manage to solve a coordination problem.

The main focus of this paper is to investigate the effect of voting on thresholds, i.e. of choosing the level of the threshold in a vote. In particular, we ask whether efficiency is higher (because coordination is more successful) when a threshold has been approved in a referendum rather than exogenously imposed. Our hypothesis is that strong electoral support for a threshold improves coordination because it serves to "signal" a willingness to contribute, i.e. widespread support is a reduced form of communication which shapes expectations. To illustrate, suppose that in a referendum all N voters support threshold level T. It seems plausible in a symmetric game (for reasons of focality and inequality aversion) to assume that, upon learning the result of the referendum, all players expect T to be reached and that all players contribute T/N.[3] In contrast, if only a small majority votes for T, players may hold the belief that some voters will contribute less than T/N each. This may result in contribution of more than T/N by player i (if he thinks he is essential) or in zero contributions (if he thinks aggregate contributions are too low for his contribution to make a difference). Thus, the signaling effect is theoretically ambiguous and the efficiency-enhancing effect of voting on thresholds is a fundamentally empirical issue which can be systematically investigated only in the experimental laboratory (e.g. Falk and Heckman 2009).

We test the hypothesis that approving of a threshold in a referendum improves the effectiveness of thresholds in a two-stage game for different threshold levels (low, intermediate or high) and for different refund rates (no, partial or full refund). In stage one, the players vote on implementing a threshold. The players know that if they accept, the public good will be provided only if their total contributions in stage two reach or exceed the threshold. In stage two, upon learning the outcome of the vote, the players make con-

contributions exceed the threshold, the public good is again continuous and proportional to aggregate contributions, as in a situation without a threshold. This property is also called the "extended benefits rule" which tends to induce relatively high contributions compared to rules in which contributions beyond the threshold are simply wasted (no rebate) or returned (full rebate) to the participants. See Marks and Croson (1998) for a comparison of these rules and Spencer et al. (2008) for a systematic discussion of alternative rebate rules.

[3]This conjecture is supported by a substantial body of research on social dilemmas showing that under conditions of strategic uncertainty, group members tacitly coordinate their choice behavior by anchoring their decisions on rules of fairness (e.g. Allison et al. 1992, Suleiman et al. 2001).

tribution decisions. If a positive threshold is approved, there is a set of equilibria in pure strategies where the threshold is exactly met. Because these equilibria Pareto-dominate the equilibrium of the standard public good game, it is a weakly dominant strategy to vote for the threshold in stage one. Variation of the threshold level is interesting because both benefits and costs of coordination increase with the threshold. Higher thresholds are more efficient if they are reached but players may be less confident that they can be reached, which increases the risk of "wasting" their contribution on a project that turns out not to be implemented.[4] Because of this trade-off, intermediate thresholds may be more efficient than high thresholds, and our experiment allows us to investigate if more ambitious thresholds are more or less efficient than low or intermediate thresholds. We also test the hypothesis at various refund rates (no, partial and full refund) in case the project is not realized. A refund provides partial or full insurance against the risk of "wasting" one's contribution on a project that is not realized. Clearly, such insurance reduces the cost of miscoordination, and we expect better coordination at higher refund rates.

Our main findings are that thresholds are counterproductive if exogenously imposed and if less than full insurance is provided. When full insurance is provided, exogenous thresholds cease to be counterproductive and become merely ineffective as efficiency is the same with no, low or high thresholds. While higher thresholds are generally associated with higher contributions, the contributions often do not increase sufficiently to match the more ambitious thresholds. Higher thresholds induce the belief that others will more generously contribute which increases contributions somewhat, but these beliefs are excessively optimistic. As a result, contributions often fail to meet the threshold.

When thresholds are endogenously chosen in a referendum, we find that thresholds are more popular with full insurance and that intermediate thresholds are more popular than ambitious ones. Thus, voting patterns reflect the (anticipated) risk of coordination failure when a threshold is ambitious and insurance unavailable. We do find that approving of a threshold serves as a coordination device in the sense that expected contributions for chosen thresholds are higher than for imposed ones, and this effect is more pronounced if the threshold receives stronger support among other voters. However, these effects are weak and cannot significantly reduce the massive coordination failure we observe in our design. As a result, overall efficiency does not increase when participants choose a threshold compared to the case when it has been imposed. Thus, introducing a threshold - whether by fiat or in a referendum - is not an effective cure for the inefficiency in the provision of public goods in our framework.

The paper proceeds as follows. Related literature is discussed in section 2, section

[4] A reason might be that the set of asymmetric Pareto-efficient equilibria shrinks with increasing threshold levels.

3 explains the experimental design. In section 4 we further discuss the predictions and hypotheses. Section 5 reports the results, and section 6 concludes.

2 Related literature

The effects of exogenously imposed threshold levels have been extensively studied in the experimental literature, but not, to the best of our knowledge, the effects of endogenous thresholds. The previous literature tends to find rather mixed results on imposed thresholds. For example, the literature review of Croson and Marks (2000) shows that results of previous studies have varied widely with success rates ranging between 10% and 82%.[5] This meta-study shows that coordination tends to be more successful with a high ratio of total benefits of the public good to its costs [6], with higher refund rates and communication. Leadership contributions and other possibilities to choose sequentially also seem to increase the effectiveness of thresholds (e.g., List and Rondeau 2003, Cadsby and Maynes 1999, Coats et al. 2009). Controlled evidence from the field seems to support these findings (List and Lucking-Reiley 2002). Rondeau et al. (2005) report higher efficiency under a threshold mechanism when contributions below the target are fully refunded.

Closely related to our study are a field and lab experiment by Rondeau and List (2008) and a theoretical contribution by Gerber and Wichardt (2008). Rondeau and List (2008) investigate (among other things) the effect of introducing a threshold into a public good that does not have a threshold for technological reasons (a fundraising drive by the Sierra Club to provide environmental education) in a field experiment under conditions very similar to our lab study (contributions are fully refunded, i.e. $r = 1$ if the threshold is not met and the "extended benefit rule" that we use in all treatments applies). In line with our results, the authors find that a higher threshold level (USD 2500 vs. USD 5000) increased donations, but the increase was insignificant and donations (USD 945 vs. USD 1375) were in both cases clearly insufficient to meet the threshold. The authors also implement a one-shot game in the lab with thresholds levels at USD 22.50 vs. USD 45. Now, contributions increase significantly (USD 5.4 vs. USD 7.5) but the increase is again less-than-proportional compared to the increase in the threshold (39% vs. 100%). We infer from these numbers (the paper does not say) that the "success rate" must have fallen with the threshold level, i.e. that higher thresholds were counterproductive in this sense. Gerber and Wichardt (2008) suggest a mechanism to provide a public good in the absence of sanctioning institutions. In their two-stage game, the players choose to pay a

[5]Our definition of a threshold as a minimum project size (allowing for project sizes exceeding the threshold) differs from some experimental studies surveyed in Croson and Marks (2000) which define the threshold as the only feasible project size.

[6]Note that we hold this ratio, which is also called the step return, constant at a level of 1.5 across all conditions.

deposit in stage one. The deposit is lost unless a player contributes to the public good. Hence, the deposit serves as a commitment device that renders contributing to the public good a dominant strategy.

3 Experimental design

Our design has 6 treatments which vary along 2 dimensions (see Table 1). We vary (i) whether thresholds T are imposed (EXO) or endogenously chosen in a majority vote (END) and (ii) the refund rate r across treatments. Each subject only participates in one treatment, i.e., a subject makes choices either in EXO or END and with only one of the refund rates. The numbers in parentheses in Table 1 show the number of participants in each condition. For example, we had 36 subjects participating in condition END0, which means that these subjects voted over threshold levels (END) and received no refund in case a threshold was not met ($r = 0$). The next two sections explain parameters and procedures in EXO and END, respectively.

Table 1: Treatments of the experiment (number of subjects per cell in parenthesis)

	Refund rate		
Decision mode	0%	50%	100%
END	END0 (#Subj.: 36)	END50 (#Subj.: 36)	END100 (#Subj.: 36)
EXO	EXO0 (#Subj.: 36)	EXO50 (#Subj.: 33)	EXO100 (#Subj.: 33)

3.1 Imposed thresholds (EXO)

Treatments in EXO implement a standard version of the threshold public goods game (see Isaac, et al. 1989). Subjects are randomly assigned to groups of $N = 3$ and group composition remains constant throughout the experiment. The subjects are endowed with $E = 20$ experimental points and decide how many of these points to keep or contribute to a public good. The payoffs are determined by

$$\pi_i = \begin{cases} E - c_i + \alpha \sum_j c_j, & \text{if } \sum_j c_j \geq T \\ E - c_i + r c_i, & \text{if } \sum_j c_j < T, \end{cases} \tag{1}$$

where π_i is subject i's payoff in points, c_i is i's contribution to the public good, and T is the threshold. If the sum of contributions within a group reaches or exceeds the threshold, each subject receives $\alpha = 0.5$ times this sum as payoff from the public good in addition to the amount kept, $E - c_i$. If the sum of contributions fails to meet the threshold, the public good is not provided and contributions are refunded at the rate r, with $0 \leq r \leq 1$. The parameter α is the marginal per capita return (MPCR) from the public good.

In EXO, participants make contribution choices for low ($T = 21$), intermediate ($T = 39$) and high ($T = 57$) thresholds. The case with $T = 0$ is a standard linear public goods game and serves as a control. The subjects make contribution choices for each of these thresholds in a randomized order. We provide no feedback about outcomes until the end of the experiment. We chose thresholds which are divisible by $N = 3$ to facilitate coordination, thus making equal contribution by all group members focal.

Each participant makes contribution choices given one refund rate, r. Refund rates vary the cost of contributing when the public good is not provided. For example, a value of $r = 0$ makes coordination failure costly because it implies that all contributions to the public goods are "wasted" if the threshold is not reached. In contrast, a value of $r = 1$ implies full insurance in the sense that contributions to the public good are fully refunded should the threshold not be met. Table 2 summarizes the parameters of the experiment.

Table 2: Parameters of the experiment

	Variable	Value
Endowment	E	20
Group size	N	3
MPCR	α	0.5
Threshold	T	$\{0, 21, 39, 57\}$
Refund rate	r	$\{0, 0.5, 1\}$

For each threshold level, participants make only one contribution decision. The one-shot nature of the game serves to investigate if participants are able to solve the difficult coordination problem absent any opportunities for communication, learning and experience. For each threshold level, subjects state their expectations on the contributions of others. The data on beliefs enables us to evaluate best-response behavior. Beliefs are elicited by rewarding a correct point prediction by an additional payment of 10 points. Incorrect beliefs were not rewarded even if these were close. Point incomes from all choices are converted into money and paid out at the end of the experiment according to the exchange rate of 10 points = 0.8 Euros. Subjects were paid for each contribution choice.

3.2 Voting on thresholds (END)

The END treatments are essentially the same as the EXO treatments except that participants vote on which threshold to implement before making contribution decisions. Voting is over pairs of thresholds T_L and T_H, with $T_L < T_H$. Participants vote on all 6 pair-wise comparisons of thresholds, i.e., $T_L = 0$ vs. $T_H = 21$, $T_L = 0$ vs. $T_H = 39$, ..., $T_L = 39$ vs. $T_H = 57$. To avoid sequence effects, we randomize the order of voting over subjects. Participants make conditional contribution decisions (i.e. according to the strategy method) for all possible outcomes of the vote. More specifically, subjects make contributions for the case that zero, one, or two of the others in the group vote for T_H. Obviously, the outcome of the referendum may depend on the subject's own vote. For example, if a subject has voted for T_H, T_H is accepted for $H_{-i} \in \{1, 2\}$, while if the subject has voted for T_L, T_H is accepted only for $H_{-i} = 2$. The decision screen in the experiment accounts for this fact (see Appendix).

Applying the strategy method has the important advantage that we observe choices for all contingencies, including the cases that are not implemented. In particular, we can analyze how the contribution behavior depends on the subject's own voting and on other group members' voting choices. This rich data allows us to investigate the effects of voting choices on contributions - the main purpose of the paper - in great detail. For example, it allows us to disentangle the signaling and commitment effects discussed below. However, the use of the strategy method has the disadvantage to make choices more complicated - participants make 18 (6 votes × 3 cases) contribution choices in END compared to 4 contribution choices in EXO - and perhaps also more cognitively demanding.

In treatment END, participants state their beliefs about the contributions by others in the group after having made voting and contribution choices. Since the subjects make conditional contribution decisions, they also state beliefs conditional on all possible voting outcomes $H_{-i} \in \{0, 1, 2\}$. In addition, we ask subjects to state their beliefs regarding H_{-i}, i.e. the number of others' votes for the higher of the two thresholds, T_H.

In total, 210 (see Table 1) undergraduate students from the University of Innsbruck participated in our computerized (z-Tree by Fischbacher 2007) experiment. A session lasted approximately 45 minutes and the average subject earned Euro 8.8.[7] Because subjects made four contribution choices in treatment EXO as compared to six in END, average payoffs were higher in END. The average earning in EXO was Euro 7.2 (incl. Euro 0.3 for rewarding correct point predictions. In END they earned Euro 10.8 Euro (incl. Euro 0.6 for correct predictions).

[7]Subjects were recruited via Email. Those with experience in public good experiments were excluded from the recruitment. The subjects were randomly assigned to treatments. We check proper understanding of the instructions (which are available from the authors on demand) in a series of control questions. The sessions did not start before all participants answered these questions correctly.

4 Predictions and hypotheses

The game without a threshold ($T = 0$) has a unique inefficient equilibrium in which all participants contribute zero, $\sum_j c_j = 0$. The threshold public goods game with $T > 0$ has multiple pure-strategy equilibria (see e.g. Isaac et al. 1996). In addition to the inefficient free-riding equilibrium, there is a set of efficient equilibria that contains all feasible combinations of contributions along the mutual best response where the threshold is exactly met, $\sum_j c_j = T$. This set contains symmetric equilibria in which each participant contributes T/N and asymmetric equilibria in which participants contribute different amounts.[8] Because $\alpha N > 0$, equilibria involving positive contributions Pareto-dominate the zero-contribution equilibrium.[9]

Figure 1 illustrates individual best responses for the low ($T = 21$), intermediate ($T = 39$) and high ($T = 57$) threshold as a function of the sum of contributions by others in the group. The figure is drawn for $\alpha = 0.5$ and $N = 3$. For example, at $T = 21$, if others' contributions are below 10 points, the individual cost to meet the threshold exceeds the individual benefit from the public good. A rational and self-interested subject therefore contributes zero to the public good. If others contribute between 11 and 20 points, the best response is to contribute just as many points as needed to reach the threshold.[10] For others' contributions above 20 points, the best response is to contribute zero because his contribution is not essential for implementation of $T = 21$. The figure shows the analogous best response functions for $T = 39$ and $T = 57$.

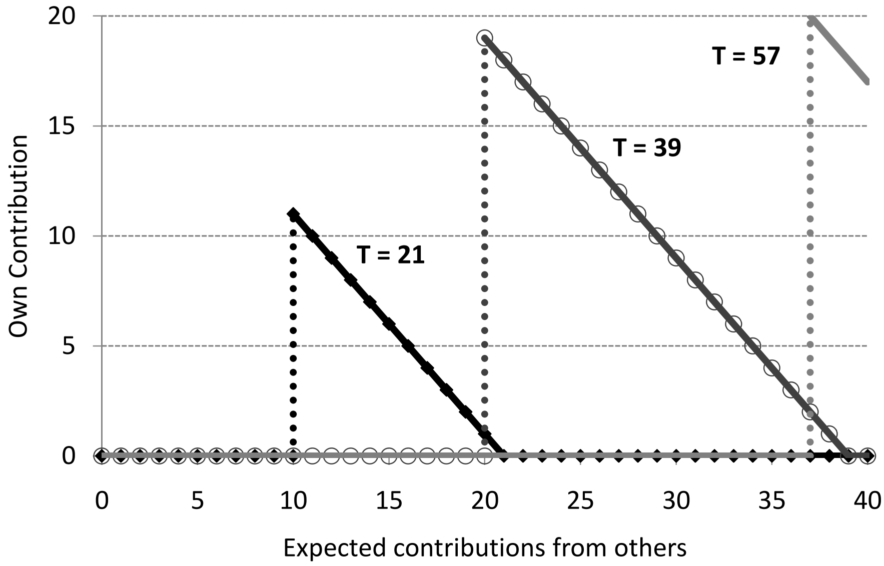
In treatment END, the game is solved by backward induction. If a majority in a group votes for a positive threshold over the "zero threshold", this decision transforms a social dilemma game with a unique inefficient Nash equilibrium into a coordination game with a set of Pareto-superior equilibria. It is therefore a weakly dominant strategy to vote for a positive threshold if the alternative is a zero threshold. Matters are more complicated when the vote is between two positive thresholds $T_L > 0$ and $T_H > 0$ because (empirically debatable) assumptions about the equilibrium selection in the contribution stage of the game must be made. Unless otherwise stated, we will therefore restrict our analysis to the decisions between $T_L = 0$ and $T_H > 0$. The within-subject variation of the threshold nevertheless enables us to evaluate which threshold level is most popular and to compare this to the empirically optimal threshold.

[8] Note that the equilibria are given by $\sum_j c_j = 0$ plus $\sum_j c_j = T$, with the additional restriction that none of the subjects contributes more than 10 points at $T = 21$ and 19 points at $T = 39$.

[9] Note that for $r = 1$, there is a larger set of inefficient equilibria. These equilibria obtain when an efficient equilibrium is not feasible, i.e., $E < (\sum_j c_j - c_i)$, and subject i is indifferent between contributing zero or any positive amount to the public good.

[10] Note that the figure only serves illustrative purposes. In the experiment subjects contribute integer numbers as points to the public good.

Figure 1: Best response as function of the sum of others' contributions for $T \in \{21, 39, 57\}$



In our one-shot design participants cannot learn from experience and the risk of miscoordination is therefore high. As illustrated in Figure 1, higher thresholds are associated with higher critical levels triggering positive contributions by rational and self-interested players. Since $\alpha N > 1$, the set of equilibria with higher thresholds contains equilibria that Pareto-dominate all equilibria with lower thresholds. This fact may make these equilibria more focal and ease the coordination problem. At the same time, if the threshold is high, the cost of miscoordination and thus deviating from equilibrium is also high for participants if $r < 1$. For example, a participant who contributes 20 points but overall contributions fail to exceed $T = 57$ ends up earning zero when $r = 0$. The net effect is therefore indeterminate and the question of which threshold level is more efficient is fundamentally empirical.

We hypothesize that voting improves efficiency by reducing the risk of miscoordination. Our hypothesis is based on three arguments. First, the number of votes for a threshold may provide a signal for others' cooperativeness. Second, voting may be determined by subjects' beliefs about others' behavior as well as their personal characteristics such as social preferences or cognitive skills. If such characteristics are relevant for the behavior in the game, voting may give rise to selection effects that influence the outcome of the game. Finally, a subject who votes in favor of a threshold may feel committed to also contribute to the successful provision of the public good. It is important to note that these arguments do not univocally support a positive effect of voting on efficiency. For instance, voters may vote strategically and send misleading signals. Moreover, depending on the expected contributions with and without a threshold, a signal of contributions from others can increase, decrease or leave unchanged optimal contributions (see Figure 1). The effect of approving a threshold in a referendum is thus theoretically indeterminate and therefore fundamentally an empirical issue.

5 Results

Section 5.1 presents the results for exogenous thresholds. Our main findings from this analysis are that exogenous thresholds are at best ineffective (with full insurance, i.e. $r = 1$) and frequently (in all other cases) counterproductive. Thus, exogenous thresholds do not increase efficiency as measured by the sum of participants' payoffs in our experiment.[11] The counterproductive effect on efficiency is most pronounced with the most ambitious threshold ($T = 57$). While higher thresholds tend to induce higher expectations and somewhat higher contributions, the increase in contributions is usually insufficient to reach the more ambitious threshold as shown in section 5.2. For example, when increasing the threshold by a factor of 2.7 from $T = 21$ to $T = 57$ with partial refund ($r = 0.5$), we find that expectations increase by a factor of 1.6, and contributions only increase by a factor of 1.4, falling clearly short of the required factor 2.7 in this example. The consequence is that the success rate, i.e. the percentage of cases in which the threshold is reached falls from 73% to 27%, and efficiency measured by average payoffs falls by 23%.

Section 5.3 shows that voting on thresholds does not improve efficiency of thresholds. Again, expectations and contributions increase with chosen thresholds, but the increase falls short of what is required to meet the more ambitious threshold. To continue the example in the previous paragraph with moving from $T = 21$ to $T = 57$ with partial refund ($r = 0.5$), efficiency falls in END by 19% which is in the same ballpark as the drop observed in EXO (23%). To explore the reasons for this result, we discuss voting behavior and the effects of voting on expected and actual contributions.

5.1 Results in EXO

In EXO the main variables of interest are the "success rate" at a given threshold, i.e. whether subjects manage to coordinate on the Pareto-superior equilibria of the game and

[11] Taking the sum of payoffs as a measure of efficiency assumes that (non-refunded) contributions are simply wasted. This is a reasonable measure of efficiency in cases where contributions to the (non-provided) public good are sunk as when half a bridge is built (and it certainly is the appropriate measure from the perspective of the average participant in our experiment by virtue of our design). Alternatively, we discuss the "surplus from the public good" which measures the efficiency of public good provision when non-refunded contributions are not counted as waste. This measure has the advantage, as one of the referees points out, of normalizing surplus across refund rates. According to this measure, introducing intermediate thresholds does have a weakly beneficial effect.

how efficiency is related to various threshold levels. In addition, we are interested in how these effects interact with the refund rate.

Figure 2: Average contributions by threshold and refund rate in treatment EXO

Figure 2 shows that average contributions monotonically increase with the threshold at all refund rates. For example, in EXO0 with no refund (upper left panel of Figure 2), average contributions are 6.4 without a threshold ($T = 0$), and increase to 8.3 ($T = 21$), 12.4 ($T = 39$), and 14.5 with the most ambitious threshold ($T = 57$). Figure 2 also suggests that contributions tend to be higher with a full refund (EXO100) than with partial or no refund. For example, at $T = 57$, contributions in EXO100 are 18.4 vs. 14.2 in EXO50 and 14.5 in EXO0. These findings are in line with the conjecture that lower refund rates increase the cost of miscoordination which makes the participants more reluctant to contribute.

Table 3 reports the results of an OLS-regression to test the significance of these results.[12] Variables T21, T39, and T57 indicate the level of the threshold (with $T = 0$ as the left-out category). Similarly, variables EXO50 and EXO100 capture the effect of a

[12]The regression uses all data from EXO, i.e. 102 subjects choosing contributions for four levels of thresholds. The regression accounts for the fact that individual choices are not independent over multiple decision rounds by using robust standard errors adjusted for clustering of individuals. We have also estimated the effects of the thresholds by accounting for individual fixed effects. The results are robust to this modification. Note that the refund does not vary within subjects implying that one cannot estimate the effects of different refunds with individual fixed effects.

Table 3: OLS Regression Treatment EXO

	(1)	(2)	(3)	(4)
Dep. Var.	c_i	π_i	PG Surplus	Belief
T21	1.060*	-0.306	0.017	0.581
	(0.582)	(0.709)	(0.025)	(1.375)
T39	1.973***	-1.517	0.102***	9.211***
	(0.716)	(0.934)	(0.034)	(1.520)
T57	2.240***	-6.984***	-0.072	14.370***
	(0.791)	(1.215)	(0.049)	(1.919)
EXO50	0.344	3.339***	0.059	1.386
	(0.787)	(1.169)	(0.055)	(1.948)
EXO100	1.640**	6.191***	0.189***	2.053
	(0.691)	(1.084)	(0.056)	(1.894)
Belief	0.396***	–	–	–
	(0.025)			
Round	-0.497***	-0.108	-0.014	-0.123
	(0.161)	(0.339	(0.014))	(0.456)
Constant	0.824	21.073***	0.342***	18.267***
	(0.807)	(1.174)	(0.045)	(1.837)
Observations	408	408	408	408
R-squared	0.881	0.895	0.600	0.831
	F(8, 101)= 359.60***	F(7, 101)= 466.06***	F(7, 101)= 53.53***	F(7, 101) = 208.81***

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

partial and full refund (as compared to no refund in EXO0). The variable "Belief" stands for expected contributions by others and ranges from 0 to 40. Finally, we include a trend for the decision round to all our regressions.[13] Column (1) confirms the visual impression from Figure 2. The first three coefficients show that higher thresholds significantly increase contributions, and the coefficient for EXO100 shows that a full refund increases contributions compared to no refund. In addition, we find that higher beliefs are overall strongly associated with higher contributions. For example, a player expecting others to contribute 40 points rather than 20 contributes about 8 points more on average.

Figure 3 shows the average payoff in points by threshold. The figure reveals that the payoffs tend to fall with thresholds with no (EXO0) and partial (EXO50) refund, and this drop is particularly pronounced for the most ambitious threshold. With full refund

[13]Note that this variable is orthogonal to our treatment variables by design because the order of decisions was randomized.

(EXO100), average payoffs are essentially constant across thresholds and the threshold level has therefore virtually no impact. The visual impression from figure 3 is again confirmed in regression analysis. Column (2) of Table 3 shows a significant negative impact of the most ambitious threshold of almost 7 points. Average payoffs are by about 6.2 points higher with full refund, which implies that overall payoffs don't drop much in EXO100 for $T = 57$, as shown in figure 3.

The average payoff is an appropriate measure of efficiency because in our design all or part of the contributions (for $r < 1$) are lost if the group fails to reach the threshold. This assumption seems reasonable to capture real-world situations, for example, where contributions go into investments that are sunk. If we choose to ignore such sunk private cost (as has been common in some of the experimental literature), we can generate a "surplus" measure reflecting how close the provision of the public good is to the efficient level of provision. Figure 4 shows the share of the maximum surplus from the public good (30 points in our design) that was actually provided. This "surplus" measure clearly paints a less bleak picture. Figure 4 shows that surplus tends to increase from $T = 0$ to $T = 39$, and to decrease again for the most ambitious threshold independent of the refund. Column (3) of Table 3 corroborates a positive effect of the threshold at level $T = 39$ for the variable "PG Surplus", but $T = 57$ is not significantly associated with negative effects. Confirming the visual impression of Figure 4, the regression (3) confirms that the surplus is significantly higher with a full refund (EXO100) than no refund.

Figure 3: Average payoff by threshold and refund rate in treatment EXO

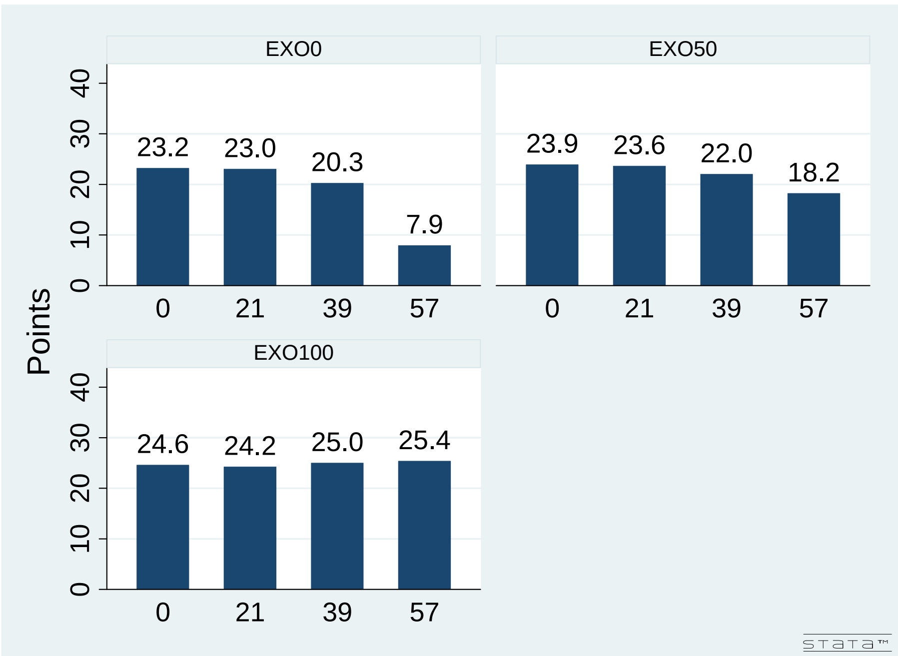
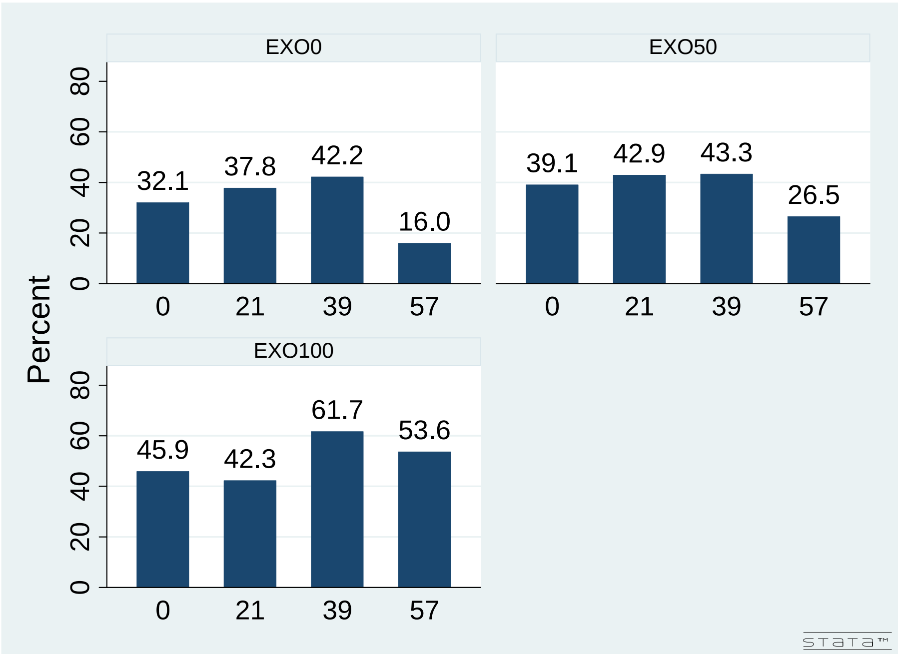


Figure 4: Average surplus (in percent of maximal) of the public good by threshold and refund rate in treatment EXO



We conclude that exogenously imposing a threshold does not increase efficiency measured by overall payoffs, and ambitious thresholds are counterproductive absent full insurance in our design. The reason why thresholds fail to increase efficiency despite the fact that contributions do increase with the threshold level is that the increase is often insufficient to match the increase in the threshold. Table 4 (upper panel) shows the "success rate", i.e. the share of groups who manage to reach or surpass the threshold. With zero or partial refund, low success rates are particularly costly because contributions are wasted in this case. The table shows that the average success rate decreases dramatically with the size of the threshold, whereas it clearly increases with the refund rate. Even a modest threshold under full refund does not significantly improve payoffs (or the surplus from the public good) compared to the standard linear public goods game (see EXO100 in Figure 3 and Figure 4) (Wilcoxon signed-rank test: $p = 0.844$ for payoffs, $p = 0.469$ for surplus).

5.2 Discussion of results in EXO

Why do the subjects increase their contributions with the threshold as shown in Figure 2 despite the fact that this behavior does not increase, and often reduces, their payoffs? The answer to this question comes in three parts. The first part is that the threshold

Table 4: Success rates by treatment

		EXO0	EXO50	EXO100	*avg.*
	21	0.75	0.73	0.82	*0.76*
Threshold	39	0.58	0.55	0.82	*0.65*
	57	0.17	0.27	0.55	*0.32*
	avg.	*0.5*	*0.52*	*0.73*	*0.53*
		END0	END50	END100	*avg.*
	21	0.45	0.67	1.00	*0.68*
Threshold	39	0.38	0.62	0.73	*0.57*
	57	0.10	0.36	0.64	*0.46*
	avg.	*0.35*	*0.58*	*0.76*	*0.57*

level provides a signal about others' contributions. If met, a high threshold improves the efficiency of the Pareto-superior equilibria. For some of these equilibria, a subject may rationally expect high contributions from others. At the same time, if the threshold is high, the cost of miscoordination and thus deviating from equilibrium is also high (for $r < 1$). This argument might induce lower expectations. A priori, the direction of the signal is therefore not clear. Column (4) of Table 3 shows the results of a regression using beliefs as the dependent variable. The results show that subjects expect much higher contributions by others if the threshold is high. On average, expectations increase by about 14.4 points if the threshold is $T = 57$ rather than $T = 0$. But these expectations were excessively optimistic. Column (1) of Table 3 shows that contributions only increased by about 2.2 points (see coefficient for $T57$).

The second part of the answer concerns the subjects' reaction given their beliefs. Table 5 shows that about a third (34% to 36%) of all subjects choose exact best responses ($= BR$) to their expectations. Among these subjects, a sizeable fraction hold focal beliefs, i.e., beliefs about others' contribution equivalent to "cost sharing", i.e. $\frac{2}{3}T$ (see numbers in parentheses in column $= BR$). A majority of subjects (between 51% and 65%) contribute more than their best response (>BR). This high share of "overcontributing" subjects can be explained by three factors. First, subjects overcontribute to avert the risk of not reaching the threshold. Second, since the contributions can have positive externalities on others within the group, subjects may overcontribute due to social preferences. Finally, there may be decision errors. However, since only few subjects contribute less than the best response (<BR) the explanatory power of unsystematic decision errors is limited. The observation that many subjects hold focal beliefs to which they choose a best response, and the fact that many subjects overcontribute given their beliefs explains why an increase in expected contributions translates into an increase in own contributions (note the strong

impact of beliefs on contributions in regression (1) of Table 3).

Table 5: Number of subjects in EXO who contribute exactly (=BR), less than (<BR), and more than (>BR) best response towards own belief

	Threshold	=BR	 BR	#obs.
EXO0	$T = 21$	11 (6)	2 (0)	23 (6)	36
	$T = 39$	12 (7)	4 (1)	20 (2)	
	$T = 57$	16 (8)	3 (1)	27 (2)	
EXO50	$T = 21$	12 (8)	0 (0)	21 (0)	33
	$T = 39$	12 (10)	4 (1)	17 (0)	
	$T = 57$	15 (8)	2 (1)	16 (0)	
EXO100	$T = 21$	10 (7)	2 (1)	11 (1)	33
	$T = 39$	13 (11)	1 (0)	19 (0)	
	$T = 57$	11 (8)	1 (0)	21 (2)	

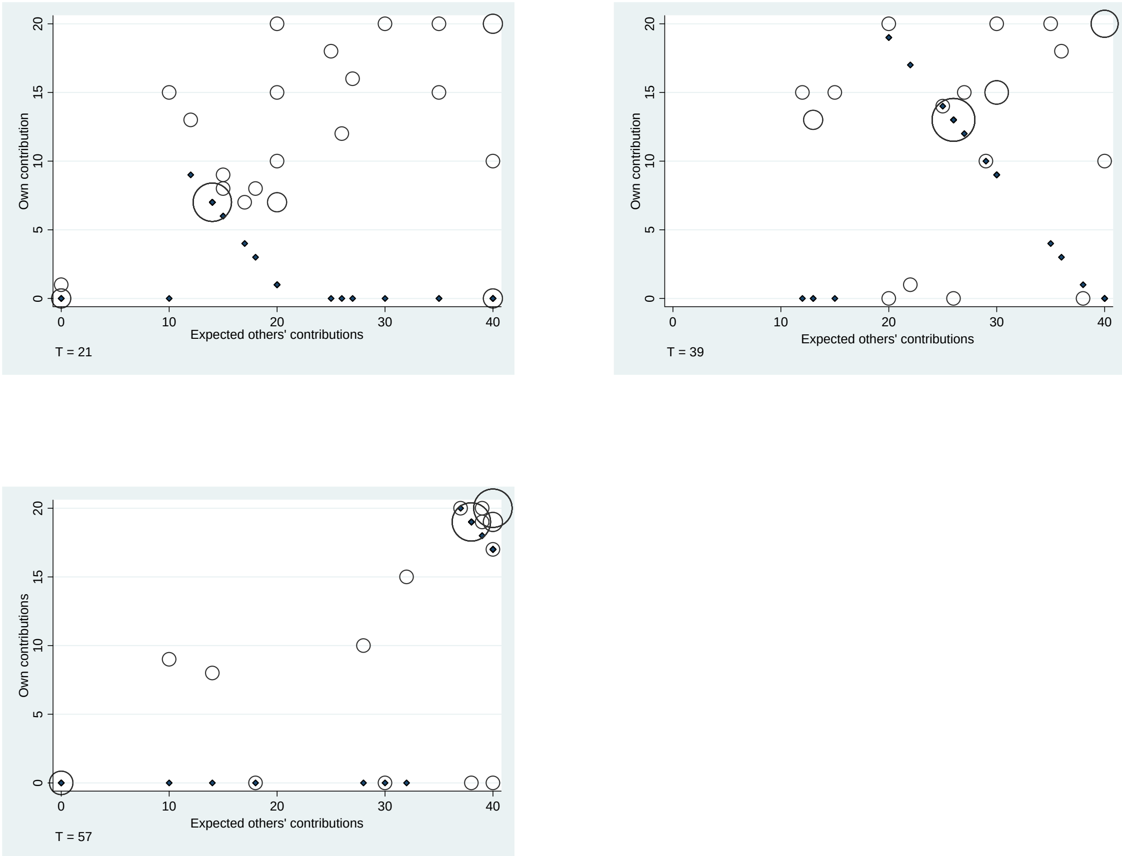
Note: The numbers in parentheses show the number of subjects with focal expectations (i.e., $\frac{2}{3}T$).

Figure 5 illustrates our discussion in the last paragraph for treatment EXO50.[14] The figure shows own contributions as a function of expected contributions by others for the low (upper left panel), intermediate (upper right) and ambitious threshold (bottom panel). The size of the circles indicates the frequency of observations for a particular combination of contributions and beliefs. The black diamonds show the theoretical best response. The figure serves to illustrate three regularities. First, contributions are positively associated with expected contributions. Second, as shown already in Table 5, approximately one third of the subjects choose best response to their beliefs and a large part of these beliefs are focal (i.e., beliefs are at $2/3T$). Finally, an increase in the level of threshold shifts up expected and actual contributions.

The third part of the answer is that, despite the positive effects of thresholds on expected contributions, coordination often fails. The most likely explanation is that expectations are imprecise and biased. The data shows a vast variation in the difference between expected and actual contributions: averaged over all thresholds and treatments, the standard deviation of this difference is 14.4 points. In addition, subjects tend to overestimate others' contributions. On average, the subjects expect others to contribute 2.1 points more than they actually do contribute ($p = 0.001$, two-sided Wilcoxon signed rank test). These figures do not come as a surprise given the findings in the previous literature. With best response functions that are kinked and decreasing for important ranges of expectations, coordination is difficult especially in the one-shot game. In fact,

[14]The figures look quite similar for EXO0 and EXO100 (not reprinted here due to space constraints).

Figure 5: Own contributions as a function of expected contributions by others in EXO50



Note: Black diamonds indicate best response. Size of circles indicates the number of observations at a particular combination of contributions and beliefs.

Isaac et al. (1989) reported success rates similar to ours' for the first period of the repeated game.

5.3 Results in END

This section reports the results for the endogenous treatments. We first ask which threshold is most popular. We then explore how voting affects choices and efficiency in the game.

Voting and aggregate outcomes: Table 6 shows acceptance of the higher of the two thresholds in each pairwise vote by treatment. The numbers in parentheses show aggregate acceptance rates. In END0 and END50, the low threshold $T = 21$ is clearly the most popular as it is the unique majority winner. In all pairwise comparisons, a majority of voters prefer $T_H = 21$ and $T_H = 39$ over $T_L = 0$, and it prefers $T_L = 21$ over $T_H = 39$ and $T_H = 57$. This result suggests that the subjects anticipate the risk of miscoordination at a high threshold. Indeed, when this risk is eliminated ($r = 1$), the majority winner is $T = 57$ in END100.

Table 6: Individual (aggregate) acceptance in percent for T_H ($T_H > T_L$)

	END0	END50	END100
0 vs. 21	56% (58%)	56% (50%)	69% (58%)
0 vs. 39	53% (67%)	61% (58%)	72% (83%)
0 vs. 57	31% (25%)	39% (33%)	64% (75%)
21 vs. 39	44% (50%)	31% (42%)	64% (58%)
21 vs. 57	44% (42%)	31% (33%)	67% (75%)
39 vs. 57	36% (17%)	31% (25%)	56% (58%)

Despite the electoral support for positive thresholds, voting does not increase aggregate efficiency (see Dal Bo et al. (forthcoming) for efficiency-increasing effects of voting in a prisoner's dilemma game). Table 7 reports regressions for treatment END, i.e. when thresholds are endogenously chosen.[15] Column (1) reports results for determinants of contributions. These results for END look very similar to those for EXO reported in column (1) of Table 3 and are indeed not statistically different from those estimates.[16] As a consequence, there is no sizable difference between exogenous and endogenous thresholds on subjects' payoffs and the surplus from the public good (compare columns (2) and (3) across tables 7 and 3). Note, however, that the positive effect of an intermediate threshold on the surplus from the public good found for EXO disappears in END. As in EXO, the reason for a lack of positive effect of thresholds is that contributions often fall short of reaching the threshold (see Table 4, lower panel). We now explore the effect of voting in more detail.

Signaling: As contributions are importantly driven by expected contributions, voting may make a difference if the voting outcome effectively signals others' contributions (see Tyran and Feld (2006) for evidence of signaling effects in voting on sanctions in public goods games). Evidence in support of a signaling effect is provided in regressions (1) and (2) in Table 8. The regression uses the data that were elicited by the strategy method. In particular, it uses the decisions from the three pair-wise votes between $T_L = 0$ and $T_H \in \{21, 39, 57\}$, where T_H has been accepted by the group. Therefore, variables T39

[15]The regression uses $3 \times 108 = 324$ data points. In principle, contributions may not only depend on the outcome of the vote but also on the baseline option in the voting choice. To assure a constant default for the voting decisions, the regression uses only data for voting choices between $T_L = 0$ and $T_H > 0$. However, the overall picture remains the same if we also include the data for choices between $T_L > 0$ and $T_H > 0$.

[16]To test, we estimated a single equation using the data both from EXO and END and employed a Chow test. None of the coefficients in column (1) is different between the treatments.

Table 7: OLS Regression Treatment END

	(1)	(2)	(3)
Dep. Var.	c_i	π_i	PG Surplus
T21	0.724	-1.883	0.021
	(0.802)	(1.184)	(0.047)
T39	1.962*	-3.897***	0.050
	(0.999)	(1.092)	(0.051)
T57	3.001**	-4.232***	0.025
	(1.412)	(1.459)	(0.072)
END50	-0.093	1.448	0.059
	(1.092)	(1.284)	(0.047)
END100	3.265***	6.649***	0.310***
	(1.002)	(1.330)	(0.049)
Belief	0.322***	–	–
	(0.042)		
Round	-0.376***	-0.369*	-0.024***
	(0.140)	(0.188)	(0.008)
Constant	2.032**	22.712***	0.334***
	(0.946)	(1.397)	(0.048)
Observations	324	324	324
R-squared	0.829	0.919	0.680
	F(8, 107) = 191.69***	F(7, 107) = 391.25***	F(7, 107) = 77.87***

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

and T57 indicate the level of the threshold at $T = 39$ and $T = 57$ as compared to $T = 21$. Furthermore, to keep the aggregate voting outcome constant, the regression uses only the cases where the decision maker plus at least one of the other subjects in the group have voted for $T_H > 0$.[17] The variable $\sum vote_j = 1$ indicates whether both of the others in the group have voted yes.

The coefficient for $\sum vote_j = 1$ in regression (1) shows that expected contributions strongly increase with others' acceptance. In particular, beliefs were on average about 3.8 points higher if the threshold was accepted with the approval of all others. The coefficient for the variable Belief in column (2) shows that contributions increase with expected contributions (and the effect is very similar to EXO, compare (1) in table 3),

[17] For the three pair-wise votes between $T_L = 0$ and $T_H \in \{21, 39, 57\}$, we observe 145 cases in which T_H is accepted and the subject voted yes. Due to the strategy method, the subject takes two choices conditional on the number of yes-votes by others. Hence, the number of observations used in the regressions is 290.

but contributions do not increase with the number of others' yes votes. Together, these estimates indicate that the expected support for a threshold provides is a signal for how much others are expected to contribute. Higher expectations, in turn, induce higher contributions.

Selection and commitment: Voters who approve of a positive threshold may do so because they expect sufficiently high contributions from others. If so, voters would rationally select into a threshold regime according to their expectations. To evaluate this argument empirically, we regress subjects' beliefs on their voting behavior. Column (3) of Table 8 shows the result. We use the data elicited from the strategy method analogously to columns (1) and (2) of Table 8. However, to keep the aggregate voting outcome constant, we now only consider the cases where two of the other voters in the group have voted for T_H.[18] Column (3) establishes a potential selection effect in that yes-voters expect higher contributions of others. However, the effect is significant only at the 10%-level.

Selection into a threshold level may also happen due to heterogeneous social preferences. To obtain a measure of cooperativeness, we let the subjects play a one-shot standard linear public goods game prior to playing the threshold public goods game. The payoffs were determined according to $\pi_i = E - c_i + \alpha \sum_j c_j$, without mention of any threshold.[19] The subjects did not receive feedback on the outcome of this choice until the end of the entire experiment. The parameters and procedures were the same as the ones described above. Using the individual-level contributions from this game as a proxy for the subjects' cooperativeness, we find no correlation between this variable and voting.[20] We have also included this proxy for cooperativeness into regression (3) (see variable $c_i PG$). The inclusion of this variable does not affect the impact of the own vote ($vote_i$).[21] To conclude: our results suggest at most modest selection effects that are not strong enough to improve the outcomes of endogenously chosen over exogenously imposed thresholds.

In addition to selecting into a threshold level, subjects may raise their contributions because they feel committed to their vote. In column (4) of Table 8, we regress subjects contributions on subjects' voting decision. The results show that yes-voters chose higher contributions than no-voters. Notice that this effect exists despite the regression accounts for potential selection based on beliefs. These results indicate that subjects feel committed

[18]The number of observations consequently amounts to 324, i.e. 3 pair-wise voting decisions $\times$ 108 subjects in END confined to the case that 2 other group members vote yes.

[19]Subjects earnings from playing this game were Euro 2.08 in EXO and Euro 2.14 in END. This money was paid out together with the other payments at the end of the experiment.

[20]A linear probability model including a constant, the refund rate, and the threshold the subjects' contributions to the public good does not explain any variation: the estimated parameter is 0.001 at $p = 0.684$.

[21]This implies that the interaction is insignificant: using the specification of regression (3), the estimate of the interaction $vote_i \times c_i PG$ is -0.115 at $p = 0.529$.

to their vote.

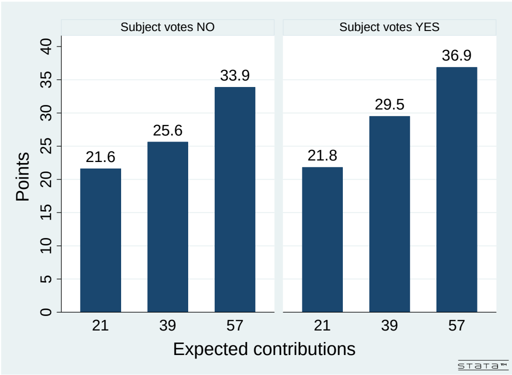
Table 8: OLS Regression Treatment END: Signaling, Selection, and Commitment

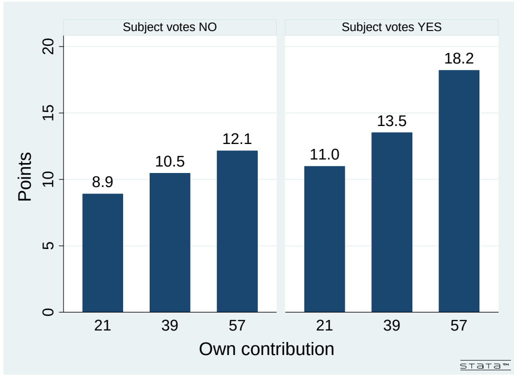
	(1)	(2)	(3)	(4)
Dep. Var.	Belief	c_i	Belief	c_i
T39	7.734***	-0.192	6.256***	0.223
	(1.253)	(0.620)	(1.167)	(0.703)
T57	15.324***	1.170	13.812***	0.977
	(1.751)	(1.131)	(1.251)	(1.084)
END50	1.444	-0.401	3.428**	-0.723
	(2.234)	(0.929)	(1.574)	(1.068)
END100	2.557	1.384*	2.425	1.877**
	(1.735)	(0.746)	(1.496)	(0.825)
Belief	–	0.383***	–	0.307***
		(0.044)		(0.048)
$\sum(vote_j = 1)$	3.841***	-0.293	–	–
	(0.715)	(0.373)		
$vote_i = 1$	–	–	2.157*	2.683***
			(1.112)	(0.746)
c_i PG	–	–	0.358***	0.126*
			(0.108)	(0.071)
Constant	12.124***	2.652**	14.353***	0.038
	(1.887)	(1.079)	(1.297)	(1.181)
Observations	290	290	324	324
R-squared	0.377	0.546	0.318	0.384
	F(5, 71) = 36.33***	F(6, 71) = 79.54***	F(6, 107) = 36.40***	F(7, 107) = 49.33***

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Figure 6 illustrates these effects graphically. The left panel shows that YES-voters, i.e. those who themselves approve of the high threshold, expect slightly higher contributions from others than NO-voters. However, the difference in expectations is small and insignificant for $T_H = 21$ ($p = 0.990$), and $T_H = 39$ ($p = 0.252$, according to a two-sides Wilcoxon rank-sum test). The only significant effect occurs for $T_H = 57$ ($p = 0.012$). At least for a high threshold, this observation indicates that subjects' approval is partially due to fact that they expect high contributions from others. The right schedule of Figure 6 shows the average own contribution conditional on the subject's vote. It shows that YES-voters

Figure 6: Average expected (left) and own contributions (right) by vote





contribute more than No-voters for $T_H = 21$ ($p = 0.063$), $T_H = 39$ ($p = 0.042$), and $T_H = 57$ ($p = 0.001$). While both effects seem to partially play a role, the relatively large effect of the voting behavior on the own contributions suggests that the commitment effects are larger than any selection effects. Further support for this conjecture comes from analysis of best-response behavior. Presumably, subjects who choose a threshold based on their expectations are more likely to choose contributions in accordance with their best response. This link is broken if a subject raises own contributions because of a commitment with the own vote. To test, we calculate the difference between own contributions and the best response towards expected contributions from others. We find that NO-voters deviate by less from their best response than YES-voters (2.76 vs. 5.78 points, $p = 0.030$), thus providing further evidence for a commitment effect.

6 Discussion and conclusion

We have studied the effect of introducing a threshold into a public goods game by voting. A priori, this is a promising approach when there is no predefined threshold for technological reasons, participants can be identified (and can vote) but where there's no central authority with the power to enforce contributions or punish non-compliance. Examples range from the small (e.g. a fund-raising drive to start a community library only if a sufficient amount is raised to buy some minimal number of books) to the large (e.g. a voluntary agreement of nations to reach a minimum abatement of greenhouse gases).

We find that accepting a threshold has significant signaling and commitment effects, but these effects are insufficient to improve the efficiency of public-good provision in our design. In both endogenous and exogenous conditions, participants in our experiment are challenged to solve a difficult coordination problem absent previous experience, opportunities for communicating or learning from mistakes within the game, and in a rather complex (we use the strategy method) and context-free (we use neutral wording in the presentation of the situation to participants). These design aspects suggest that we provide a demanding test for the efficiency-improving effect of voting on thresholds. However, our design can also be considered to be favorable for coordination since the players were symmetric, i.e. the endowments, costs of contributing and benefits from the provision of the public good were the same for all participants, and the symmetry was common information. In naturally occurring contexts the number of players may be large and unknown, and information about important parameters of the game may be incomplete which may further exacerbate coordination problems because contribution according to equal cost share are less focal (as suggested by Bagnoli and McKee 1991. See Rondeau et al. 1999 for a discussion). In addition, the thresholds to choose from were divisible by N in our design which should have made equal contributions to reach the threshold focal. In most naturally occurring examples such as donations for a community library or greenhouse gas abatement on a global level, players are not symmetric. For example, in the context of abatement, some countries are large (i.e. their emissions are large) and some are poor (i.e. their opportunity cost of abatement is high), and to some both or neither applies. In this situation, coordination can be expected to be more difficult than in our experiment because equal contributions are neither focal nor fair.

In summary, we provide the first study to show that introducing thresholds, both imposed or approved in a referendum, is no sure cure for the inefficient voluntary provision of public goods when no other, more intervention-intense, mechanism is available. An interesting alley for further research is to investigate the robustness of this result by adding more context (e.g. in a field experiment), allowing for learning (e.g. in a repeated laboratory experiment), making contributions sequentially, or by adding opportunities for communication.

References

Allison, S.T., McQueen, L.R. and Schaerfl, L.M. (1992): Social Decision Making Processes and the Equal Partitionment of Shared Resources. *Journal of Experimental Social Psychology* 28: 23-42.

Bagnoli, M. and Lipman, B. (1989): Provision of Public Goods: Fully Implementing the Core through Private Contributions. *Review of Economic Studies* 56: 583-601.

Bagnoli, M. and McKee, M. (1991): Voluntary Contribution Games: Efficient Private Provision of Public Goods. *Economic Inquiry* 29: 351-66.

Cadsby, C. and Maynes, E. (1999): Voluntary Provision of Threshold Public Goods with Continuous Contributions: Experimental Evidence. *Journal of Public Economics* 71(1): 53-73.

Coats, J., Grosskopf, B., and Gronberg, T.J. (2009): Simultaneous versus Sequential Public Good Provision and the Role of Refunds: An Experimental Study. *Journal of Public Economics* 93(1-2): 326-35.

Croson, R. and Marks, M. (2000): Step Returns in Threshold Public Goods: A Meta- and Experimental Analysis. *Experimental Economics* 2: 239-59.

Dal Bo, P., Foster, A. and Putterman, L. (forthcoming): Institutions and Behavior: Experimental Evidence on the Effects of Democracy. *American Economic Review* forthcoming.

Falk, A. and Heckman, J.J. (2009): Lab Experiments Are a Major Source of Knowledge in the Social Sciences. *Science* 326 (5952): 535-8.

Fischbacher, U. (2007): Z-Tree: Zurich Toolbox for Readymade Economic Experiments. *Experimental Economics* 10(2): 171-8.

Gerber, A. and Wichardt, P.C. (2008): Providing Public Goods in the Absence of Strong Enforcement Institutions. *IEW Working Paper* No. 303. University of Zurich.

Isaac, M., Schmidtz, D. and Walker, J. (1989): The Assurance Problem in a Laboratory Market. *Public Choice* 62: 217-36.

List, J.A. and Lucking-Reiley, D. (2002): The Effects of Seed Money and Refunds on Charitable Giving: Experimental Evidence from a University Capital Campaign. *Journal of Political Economy* 110(1): 215-33.

List, J.A. and Rondeau, D. (2003): The Impact of Challenge Gifts on Charitable Giving: An Experimental Investigation. *Economics Letters* 79(2): 153-159

Marks, M. and Croson, R. (1998): Alternative Rebate Rules in the Provision of a Threshold Public Good: An Experimental Investigation. *Journal of Public Economics* 67: 195-220.

Rondeau, D. and List, J.A. (2008): Matching and Challenge Gifts to Charity: Evidence from Laboratory and Natural Field Experiments. IZA discussion paper no. 3278, January 2008.

Rondeau, D., Poe, G.L., and Schulze W.D. (1999): Voluntary Revelation of the Demand for Public Goods Using a Provision Point Mechanism. *Journal of Public Economics* 24: 131-55.

Rondeau, D., Poe, G.L., and Schulze W.D. (2005): VCM or PPM? A Comparison of the Performance of Two Voluntary Public Goods Mechanisms. *Journal of Public Economics* 89: 1581-92.

Spencer, M.A., Swallow, S.K., Shogren, J.F. and List, J.A. (2008): Rebate Rules in Threshold Public Good Provision. NBER Working Paper No. 14559, December 2008.

Suleiman, R., Budescu, D.V. and Rapoport, A. (2001): Provision of Step-Level Public Goods with Uncertain Provision Threshold and Continuous Contribution. *Group Decision and Negotiation* 10(3): 253-74.

Tyran, J.-R. and Feld, L. (2006): Achieving Compliance when Legal Sanctions are Non-Deterrent. *Scandinavian Journal of Economics* 108(1): 135-56.

Appendix I: Instructions for treatment "END"

(Instructions and screenshots are translated from German. Instructions for EXO are available from the authors on request. Welcome to the experiment. If you read the instructions carefully and follow the rules, you can earn money. Your earnings not only depend on your decisions but also on those of the other participants. Your earnings will be paid to you in cash immediately after the experiment. During the experiment you earn points. These points will be converted to Euros according to the following exchange rate: **10 Points = 80 Cents (0,8 Euros)**.

This experiment has 7 phases. We now explain phase 0. For phase 1 to 6 you will get additional instructions at the end of phase 0.

In this experiment, the participants are randomly divided into groups of three, i.e. you are in a group with two other participants. The groups remains the same during the whole experiment, i.e., the groups have the same members in all phases.

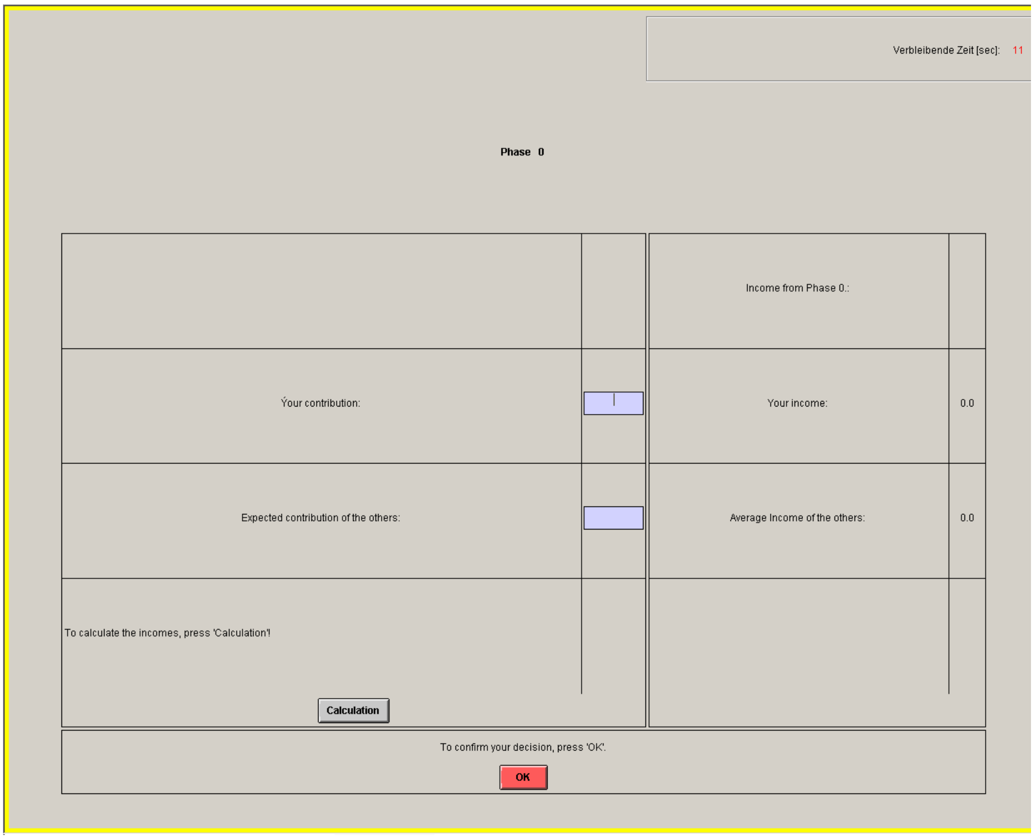
At the beginning of each phase, every member is endowed with 20 points. Your task is to decide how many of your 20 points you want to contribute to a project and how many you want to keep for yourself. Likewise, every other group member has to decide how to use his endowment.

Your income comes from two sources: (1) the points you keep for yourself. (2) the "income from the project". This income is calculated as follows: **Your income from the project = 0.5 × Sum of the contributions of all group members to the project**.

The income of the other members is calculated in the same way, i.e., each group member gets the same income from the project. If, for example, the sum of contributions is 60 points, you and the other group members earn an income from the project of $0.5 \times 60 = 30$ points. If you and the other two group members together contribute only 10 points to the project, you and the other group members each get $0.5 \times 10 = 5$ points income from the project.

For every point you keep for yourself, you earn an income of 1 point. If you add this point to the project instead, the sum of contributions to the project will increase by one point and your income from the project increases by $0.5 \times 1 = 0.5$ points. However, the income of all the other group members also increases by 0.5 points, such that the entire income from the project increases by $0.5 \times 3 = 1.5$ points. Thus, your contributions to the project generate income for the other members of your group. Conversely, you earn money for every point the other group members add to the project. For every point some other member of your group adds to the project, you get $0.5 \times 1 = 0.5$ points.

You make your decision on the computer: At the beginning of phase 0 the following decision-monitor is displayed.



In the upper right corner you see how much time is left for you to decide. In the row in the middle you enter "Your Contribution" to the project. Please enter a number from 0 to 20 here. In the row below below, we ask you to state your "Expected contributions from others". Please enter a number from 0 to 40 here. We ask you to think carefully about how much you expect the others to contribute. If your expectations matches the actual contributions from others, you will earn 10 points in addition to your other payments.

Once you have entered your decisions, we ask you to click on the grey button labeled **"Calculation"**. As a decision aid, the computer then calculates the outcomes given the numbers you (hypothetically) typed. In particular, the computer calculates the income for you and the others given your contribution and assuming your estimated contributions from others are correct. You can change your input and push the button "Calculate" as many times you like.

Example:

- Your contribution is 10 points. You expect the contributions by others to be 30 points. In this case your calculated income will be 30 points and the expected average income of the others will be 25 points.
- Your contribution is 0 points. You expect the contributions of the others to be 30 points again. Your calculated income is now 35 points and the expected average income of the others is 20 points.

You confirm your final decision by pushing the **OK-button**. You will be informed about your income only at the end of the experiment, after phase 6 is completed.

If you have any question, please raise your hand and an experimenter will come to you and offer help.

Instructions for Phases 1 to 6

Like in phase 0, in the phases 1 to 6 you decide how many of your 20 points you want to contribute to a project and how many you want to keep for yourself. Like in phase 0 your income consists of the "income from the project" and the points that you keep for yourself.

Unlike in phase 0, however, in phases 1 to 6 you earn an income from the project only if the sum of the contributions of all group members exceeds a **minimal project size**. The minimal project size is a number. Specifically, this number will be 0, 21, 39 or 57 points, depending on the phase you are in.

1. What happens if the sum of contributions of all group members is **smaller than** the minimal project size? In this case you and the other group members **do not get an income from the project.** You will be refunded **50%** [**0% / 100%**] of you contributions to the project.
2. What happens if the sum of the contributions of all group members is **equal or larger than** the minimal project size? In this case you and the other members of the group get **an income from the project**. You **will not** be refunded your contributions to the project. (Note: Like in phase 0 your income from the project is 0.5 $\times$ sum of the contributions of all group members to the project.)

Example: Assume that the minimal project size was 39 points. The following table (*not shown*) illustrates some examples.

Example, row I: You contribute 13 points and the others contribute 26 point to the project. In this example the minimal project size is exactly reached. You earn an income of 19.5 ($39 \times 0,5$) points from the project plus 7 points, which you have kept for yourself. Thus your income is 26.5 points.
Example, row II: You contribute 12 points and the others contribute 26 points to the project. In this example the minimal project size is not reached. Thus you do not earn income from the project. **0%** [**50%/ 100%**] of your contribution is refunded. This leads to an income of 14 [8/ 20] points.
Example, row III: You contribute 20 points and the others contribute 26 points to the project. 7 points more than the minimal project size are contributed. You earn an income of 23 (46×0.5) points from the project and have not kept points for yourself. Thus your income is 23 points.
Example, row IV: You contribute 13 points to the project and the others contribute 20 points to the project. The minimal project size is not reached. Note that while your contribution is as big as in example in row I, your income just amounts to 13.5 [7/ 20] points.

Voting on the minimal project size

You and the other group members **choose the minimal project size** relevant to your decision. At the beginning of each phase you and the other two members vote between two alternatives A and B. Every alternative stands for a minimal project size. The alternative that receives the majority of votes (two or more) is realized and determines the minimal project size for the relevant phase.

You cast your vote over the computer: at the beginning of each phase the following decision screen is displayed:

On the upper panel, you choose between alternatives A and B. In the example shown on the screen, alternative A is a minimal project size of 0 points and alternative B is a minimal project size of 21 points. That means that the vote is between a minimal project size of 0 and one of 21 points. Choose your preferred alternative.

In the panel below we ask you to state your **expectation** of the voting behavior of the other two group members. Please enter whether you expect 0, 1 or 2 other group members to vote for alternative A.

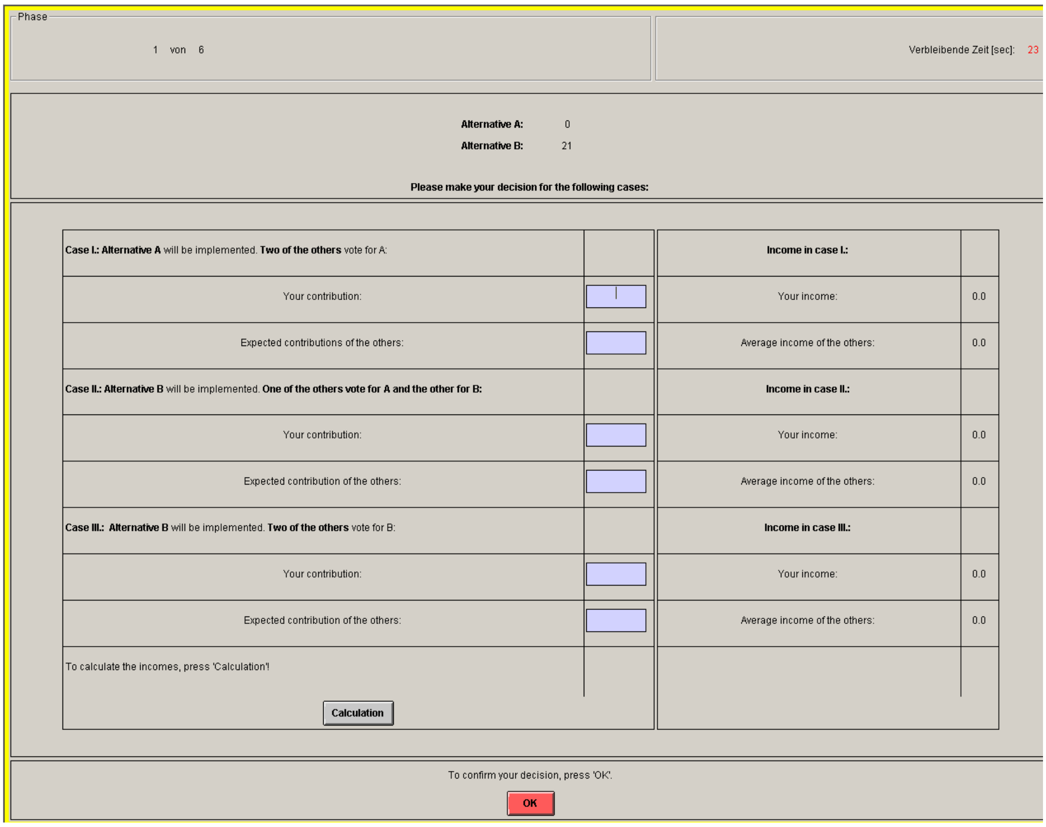
Again, we ask you to think carefully about your decision. If you correctly predict the actual voting behavior of the others you earn an additional 10 points. After the vote you choose your contribution on the following decision screen:

In this example the minimal project sizes are 0 for alternative A and 21 for alternative B. We ask you to take your decision for three cases:
Case 1: You choose given that both other members vote for alternative A and hence - in this example - a minimal project size of 0 (alternative A) is implemented.
Case 2: You choose given that one of the other group members votes for alternative A and one for alternative B. Hence - in this example - a minimal project size of 0 is implemented if you voted for A, and the minimal project size of 21 is implemented if you voted for B. (note that the screen shows the outcome assuming that you voted for alternative B)
Case 3: You make your decision, given that both of the other group members vote for alternative B and hence - in this example - a minimal project size of 21 (alternative B) is implemented.

Depending on the voting behavior of the others only one of these cases will occur. Your income therefore is determined by the case that occurs. The other two cases are not relevant for your income. At the time of your decision you do not know which case will occur; we therefore ask you in your own interest to treat all three cases as if they actually occur.

For each of the three cases you take two decisions:

1. In the row "Your contribution" you enter how many points you want to contribute to the project. Please enter a number between 0 and 20 here.
2. In the row "Expected contribution of the others" we ask you to state your expectation of how



much the other two group members will contribute to the group project. If your expectation is correct, you earn an additional income of 10 points.

Note: When you take your decision, please keep in mind the relevant minimal project size! Please further mind, that 50% [0%/ 100%] of your contribution to the project will be rebated, if the minimal project size was not reached.

Like in phase 0, pressing the calculate-button will calculate your income and the average income of the others based on your contributions and given that your expected contributions of others are correct. To confirm your final decision, press the OK-button. Then, the next phase will start, which is essentially the same as phase 1 but with different alternatives A and B.

You will learn the outcome of the voting and your income only at the end of phase 6. Before that you will not receive any information about the contributions or voting outcomes for a particular phase.

The experiment will start soon. If you now have any questions, please raise your hand and wait until an experimenter will come by.