Public Choice (2006) 130:115–128
DOI 10.1007/s11127-006-9075-3

ORIGINAL ARTICLE

# Why feed the Leviathan?

**Pablo Guillen · Christiane Schwieren · Gianandrea Staffiero**

Received: 7 June 2006 / Accepted: 28 June 2006


**Abstract** This is a study about the possibility of self-governance. We designed two versions of a step-level public good game, with or without a centralized sanctioning mechanism (CSM). In a baseline treatment participants play 14 rounds of the non-CSM game. In an automatic removal (AR) treatment participants play 7 rounds with CSM plus 7 rounds without CSM. In voted removal (VR) participants play 7 rounds with CSM followed by a voting stage to decide whether to keep CSM. All VR groups removed CSM. Contributions in AR and VR after CSM removal are dramatically higher than in the baseline. Most groups with a CSM history managed to cooperate until the last round. We do not find more cooperation in VR than in AR.

**Keywords** Public good · Step-level · Sanctioning institution · Cooperation · Education · Trust

## 1. Introduction

This is a study about the possibility of self-governance. Our inspiration comes mainly from the work of Ostrom et al. on self-governance (1992). They show how communication and/or sanctions alleviate the tragedy of the commons. In one of their treatments participants are allowed to punish partners. Fehr and Gaechter (2000) and Fehr et al. (2002) have similar experiments with general punishment opportunities. They examine experiments on public good games in which participants can invest to decrease the payoff of other players after

P. Guillen (✉)
Harvard Business School, Baker B90, Soldiers Field, Boston, MA 02163
e-mail: pguillen@hbs.edu

C. Schwieren
Pompeu Fabra University and Mannheim University

G. Staffiero
SP Research Center, IESE Business School, University of Navarra

observing their contributions. However, it is not possible to trace each player contributions from one round to the next. Hence, revenge is not possible.

In many social interactions, especially in large societies, people cannot identify others' actions, only the final result. Additionally, it might not be desirable that individuals punish each other because of the possibility of never ending cycles of revenge. This problem can be alleviated by a central punishment agency. In this context, a professional third party can be set up to observe actions and punish the individuals who do not behave according to a certain mutually beneficial rule. Notably, this was suggested by the XVII century philosopher Thomas Hobbes in his "Leviathan" Hobbes (1998).[1] If this centralized sanctioning mechanism is funded by voluntary contributions it actually becomes a public good itself. That is the situation studied by Yamagishi (1986, 1988). While this is an interesting approach, many real-world institutions are not funded by voluntary contributions but by taxes. Those are contributions people decide upon in a more indirect way, through electing politicians who will decide how to spend the money.

Among economists it is quite standard to think that institutions serving as third party contract enforcement evolve in advanced societies so that it is no longer necessary to rely on imperfect social norms or trust. According to this theory, a Leviathan has to be designed in a way that makes socially undesired behavior also individually undesired. That is, the expected value of misbehaving must be lower than the expected value of the socially desired action. It is not necessary for an efficient Leviathan to be a perfect watchdog; the size of the sanction can balance the probability of catching deviators. A mere observation of real life facts can easily challenge this approach. Let us show an example: there are no turnstiles in the subways of Vienna, Budapest or many other cities. In the case of Budapest, tickets are sometimes inspected when passengers leave the station. Nevertheless, it is announced in the trains when tickets are going to be inspected. So, if you are a free rider it is easy just get off one station after and avoid the fine, which is anyway relatively low. Therefore it is possible to free ride. However, it is clear that most people pay their tickets. If people are habituated to follow certain rules, that might help to keep cooperation high. Many social norms are learned through initial enforcement via some kind of sanctioning from parents or teachers. Probably everyone's own family is the most important educational institution. Siblings are taught to cooperate under the threat of punishment. When they grow up they pay their subway tickets and do not litter.

We want to focus on the possible educational effects of a centralized sanctioning institution. Our main hypothesis is if such an institution enforces a mutually beneficial situation this is likely to be maintained when the institution disappears. To test our hypothesis, we designed a step-level public good game. The implementation of the public good does not require all group members to contribute fully. Therefore, there are multiple equilibria in which the public good is provided plus the equilibrium in which no one contributes. We designed our game so that cooperative equilibria are asymmetric. Hence, repeated cooperation cannot be the result of the sheer repetition of the same move by all players. Moreover, someone is always slightly exploited in a cooperative equilibrium. Of course, cooperative equilibria are less likely to occur. Imperfectly coordinated actions leading towards cooperation may be closely above or closely below the step-level. We are interested to see whether slightly uncoordinated cooperation can last over time, especially after people experienced sanctions for non-cooperative moves. We therefore set up three different treatments: a baseline treatment without any sanctioning mechanism and two treatments with a centralized sanctioning

[1] See Coleman (1986) for a review of the collective action problem in social sciences.

mechanism. The latter differ in the way the sanctioning mechanism is removed. In the baseline treatment (BL) participants played 14 repetitions of our step-level public good game. Only one group managed to reach the step-level at least three times. Average contributions went down to being close to zero in round 14. In both the Automatic Removal treatment (AR) and the Voted Removal treatment VR groups played the 7 initial rounds under the rule of a Leviathan or "centralized sanctioning mechanism" (CSM). Under CSM an individual has to pay a fine if his or her contribution is below a certain amount. However, everybody in a group contributing that amount is not enough to ensure the provision of the public good. Cooperative equilibria are asymmetric; hence someone has to contribute one unit more. On the other hand, with CSM everybody contributing zero is also an equilibrium. In AR people know in advance they are going to play with CSM from round 1 to 7 and without CSM from round 8 to round 14. Hence, after round 7 the game becomes the same as in BL. In VR, after round 7, each group decides by a majority vote whether to maintain or to get rid of CSM. Notice the main reason to get rid of the CSM is efficiency. If people thought cooperation can be maintained without CSM there is no reason to vote for keeping it. Note that the willingness to free ride is not a valid explanation for voting against CSM. Indeed, in a step level public good game it is very hard to reduce one's own contribution without causing the public good not to be reached. On the other hand there is always a little free riding in cooperative equilibria. Our most salient result is that considering all groups in AR and VR, most of them managed to keep cooperating until the very end of the experiment. Note that according to Frey and Oberholzer-Gee (1997) and Frey (1998) our CSM may have a negative effect on intrinsic motivation to cooperate. Our results suggest that this possible effect is small relative by the educational effects.

A secondary hypothesis was that cooperation should be higher in VR than in AR because people can actually decide by themselves. We expected that in this treatment more groups should keep cooperating after having voted away the CSM. This hypothesis is based on Aquino et al. (1992) and Kroll et al. (2002). In these studies voting is perceived as fair because those who are affected by a decision have their voice heard before the decision is made. However, we could not confirm a positive effect of voting in our experiments.

Perhaps the real-life examples closest to our experimental design are the UN peacekeeping missions. The "blue helmets" arrive in countries in turmoil and stay until people learn to deal with each other in a peaceful way. Then they withdraw from the country and often the conflicts do not resume. This can be understood as if the "blue helmets" induce a mutually beneficial situation which people have no interest in breaking when the peacekeeping force leaves. In the sense of Kramer (1999), they establish deterrence-based trust to develop knowledge-based trust.

The paper is structured in five sections and two appendices. Section 2 describes our experimental design. Section 3 explains our experimental procedures. In Section 4, we describe our results and Section 5 concludes. Appendix 1 includes an English translation of the experimental instructions for the VR treatment. Finally, Appendix 2 contains three tables with overall group contributions per round.

## 2. Experimental design & theoretical analysis

A step-level public goods game is played in groups of five. Group composition is randomly determined at the beginning of the experiment and does not change from round to round. The step-level, or threshold-based structure, allows us to get a clearer definition of the sufficient level of contribution, and a related simple way to design the centralized sanctioning mechanism, which is easily understood by the participants. Each player in every round has an endowment worth 10 units of experimental currency, and has to decide his or her contribution (via the choice of an integer number between 0 and 10) to a step-level public good reached at a total contribution of 31. If 31 is reached each player equally receives 12 points. If 31 is not reached no public good is created, so that any positive contributions are lost and each player only receives what he has kept.

In the game without the CSM, player $i$'s payoff ($\pi_i$) in any given round is defined as:

$$\pi_i = \begin{cases} 10 - g_i + 12 & \text{if } \sum_{j=1}^{5} g_j \geq 31 \\ 10 - g_i & \text{otherwise} \end{cases} \tag{1}$$

where $g_j$ represents player $j$'s contribution.

In the two experimental treatments, AR and VR, we introduce a centralized sanctioning mechanism. The CSM punishes players who choose a contribution strictly lower than 6 by subtracting 6 units from their round payoff. The CSM is costly: 3 units are subtracted from each players round payoff regardless of reaching the step-level or not.

Therefore player $i$'s payoff ($\pi_i$) in any given round is defined as:

$$\pi_i = \begin{cases} 10 - g_i + 12 - 3 & \text{if } \sum_{j=1}^{5} g_j \geq 31 \quad \text{and} \quad g_i \geq 6 \\ 10 - g_i + 12 - 3 - 6 & \text{if } \sum_{j=1}^{5} g_j \geq 31 \quad \text{and} \quad g_i < 6 \\ 10 - g_i - 3 & \text{if } \sum_{j=1}^{5} g_j < 31 \quad \text{and} \quad g_i \geq 6 \\ 10 - g_i - 3 - 6 & \text{if } \sum_{j=1}^{5} g_j < 31 \quad \text{and} \quad g_i < 6 \end{cases} \tag{2}$$

The experiment consists of a baseline treatment and two different treatments. In the baseline treatment (BL), the game described by (1) is played for 14 rounds. In the AR treatment, groups play the game including the CSM structure defined by (2) for 7 rounds. Then they play 7 more rounds of the game without CSM. In VR, groups play the game including the CSM structure defined by (2) for 7 rounds. Then each group votes whether to maintain the CSM structure for the following 7 rounds or get rid of it and therefore play the game defined by (1). The result is decided by simple majority rule. Experimental instructions for AR explain that CSM is going to be removed after round 7. Instructions for VR explain the voting stage and their consequences in VR.

### 2.1. Theoretical considerations

The equilibrium analysis for the stage game without the CSM is straightforward. Every combination of contributions summing up to 31 is a Nash equilibrium (NE). All players contributing zero is also NE. Notice the latter is a symmetric equilibrium whereas the former equilibria are all asymmetric because participants are allowed to contribute only integer amounts. Therefore, a coordination problem in cooperative equilibria arises. However, some equilibria are more symmetric than others: consider any combination of contributions in which four people choose 6 and one chooses 7.

Playing the same stage game Nash equilibrium in each round is a NE for the repeated game. This comes from a backwards induction logic. The asymmetric nature of the stage game makes trigger strategies impractical for the repeated game. With only one symmetric cooperative equilibrium and only one symmetric non-cooperative equilibrium a trigger strategy formulated as "play cooperative until somebody plays non-cooperative" would be NE in the repeated game. A situation where one player chooses "I play 7 while contributions sum up to 31, otherwise 0" and the other do "I play 6 while contributions sum up to 31, otherwise 0" is still a NE of the repeated game. Nevertheless, it is not likely that people manage to coordinate repeatedly in that way. The main problem is how to decide who is the player who contributes 7. A cycle implying that people take turns to be the 7-player in a repeated game seems even more far-fetched even if it would be also NE.

The CSM does not change the equilibrium analysis very much. However, not all combinations of contributions summing up to 31 are still a cooperative equilibrium. Indeed, in the stage game with CSM playing anything between 1 to 5 is strictly dominated. A participant contributing one of these amounts would be better of either changing to contribute 6 if that leads to implementation of the public good and avoiding the 6 units fine, or changing to 0 if the public good is not going to be implemented in any case. Combinations in which two or more players pick 0 and the others 6 are also Nash equilibria. In any case, (0, 0, 0, 0, 0) and any combination of four players picking 6 and one picking 7 is still an equilibrium of the stage game with CSM. The same considerations made for the repeated game without CSM are still valid for the game with CSM.

## 2.2. Hypotheses

In light of our equilibrium analysis it seems very difficult to coordinate on the asymmetric cooperative equilibrium. It would be possible, however, that groups are able to contribute closely below or above the step-level. We hypothesize that the CSM helps to coordinate to reach it.

*Hypothesis 1:* in the treatments with the centralized sanctioning mechanism (for the first 7 rounds) contribution to the public good is higher than in the baseline. As a result the step-level is reached more often.

We are mainly interested in what is going to happen after the CSM disappears. We propose that contributions will be higher in the AR and VR treatments than in the baseline even without the CSM. The reason for this is that we assume that the CSM helps people to learn an efficient way to reach the step-level.

*Hypothesis 2:* in the treatments with the CSM contributions to the public good are higher than in the baseline in periods 8 to 14. As a result the step-level is reached more often.

Following Aquino et al. (1992) and Kroll et al. (2002) we hypothesize that contributions in VR are going to be higher than in AR after CSM is removed.

*Hypothesis 3:* in VR contributions are higher than in AR in periods 8 to 14. As a result the step-level is reached more often.

### 2.3. Interpreting the vote

If a majority of members in a group develops enough trust on his or her peers to think cooperation is going to be maintained after CSM disappears they would vote against maintaining it because CSM is costly and therefore inefficient. Willingness to free-ride does not seem a valid reason to vote against CSM. Remember we look at a step-level public good game. If a subject decreases his or her contribution from a cooperative equilibrium it is not very likely that the public good is going to be produced anymore. Actually, this would happen only if at the same time one or more other participants decide to increase their contributions. If the group contribution does not reach the step-level there is no public good produced at all, resources are wasted, and there is no free-riding.

## 3. Sample and procedure

The sample consisted of 120 students of Universitat Pompeu Fabra in Barcelona, Spain, who signed-up and participated voluntarily for performance-based payment. They came from different faculties, mostly from Economics, Business and Law.

In each session, 3 stable groups of 5 students played the public goods game together. For the BL treatment, 2 sessions were run. For the other treatments, VR and AR, 3 sessions each were run. The experiment was computer-based, using the experimental software z-Tree Fischbacher (1999).

When students arrived at the lab, they were randomly assigned to a1 computer and received the instructions. The instructions do not mention anything like “sanctioning” or “punishment”. Instead, two sets of rules were described as Rule Set A (which includes the CSM) and Rule Set B (without the CSM).

BL consisted of 14 rounds of the step-level public good game. AR consisted of 7 rounds of the game with the CSM followed by another 7 rounds of the game without CSM. Participants were informed about this in the beginning of the experiment and then warned after period 7 that the rules of the game changed for the rounds to come. VR consisted of 7 rounds of the game with CSM followed by a voting phase. In the voting phase, participants had to indicate whether they wanted to continue with Rule Set A or change to Rule Set B. Participants were then informed about the result of the simple majority vote, i.e., whether the valid set of rules in the next 7 rounds was A or B. In both treatments, participants knew already at the beginning that there would be some change of rules during the game and how this change would be achieved.

After the experiment, each player was privately paid according to the sum of payoffs across all rounds (3 EUR cent per “payoff unit” were paid) plus a show-up fee of 2 EUR. Average earnings in the experiment were about 6,20 EUR.

## 4. Results

Our experimental design allows us to study the effect of having recently played under a sanctioning mechanism. As we can see inTable 1 the percentage of times the step-level is reached is indeed much higher in treatments AR and VR than in BL. This is true when the sanctioning mechanism is active (rounds 1–7). However, it is also true when it is not, in rounds 8 to 14 (in the case of VR every group voted in favor of getting rid of the CSM).

**Table 1** Percentage of times the public good is provided

| | 1–7 | 8–14 |
|---|---|---|
| BL | 7% | 0% |
| AR | 89% | 65% |
| VR | 87% | 48% |
| AR and VR | 88% | 56% |

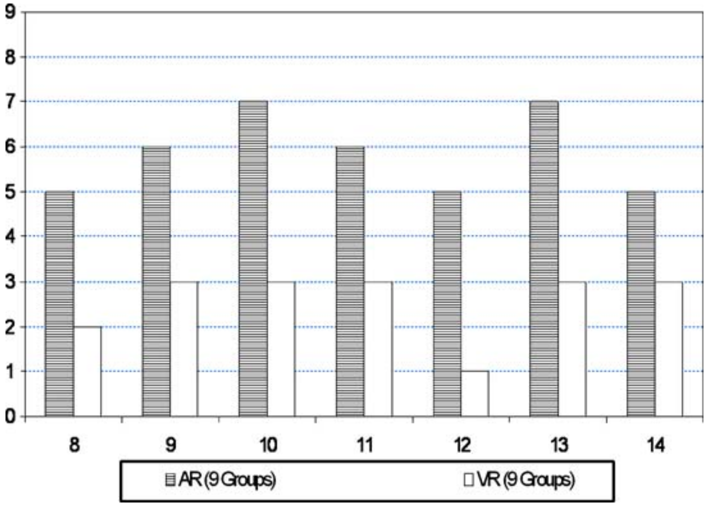


**Fig. 1** Number of groups reaching the step-level

Notice that the CSM is not working perfectly, that is, the percentage of times the threshold is reached is lower than 100% in periods 1–7 for both AR and VR.

Figure 1 shows the number of AR and VR groups reaching the step-level from round 8 to round 14. In AR five groups reach 31 in both round 8 and 14; 7 groups out of nine still reach the step-level in round 13. As you can see in the AR overall results table included in Appendix 2 7 out of nine groups manage to contribute closely above of below the step-level throughout all periods, so sometimes they reach it sometimes not. In VR the number of groups reaching 31 is always lower than in AR. Nevertheless, three groups reach the step-level almost every round, and also in the last one.Figure 2 shows the average contribution per period in the three different treatments. BL's average is never above the step-level. AR and VR are always above the step-level before round 7. After this round a decline in the average contribution is observed in both treatments. However, as Figure 1 indicates, this decline does not affect every group involved, but rather those which fail to stay above the step-level in the absence of the sanctioning mechanism.

Table 2summarizes the statistical results that support the facts we analyzed so far. We compare average success rates in the baseline and the two treatments. That is, we construct a variable equal to 1 if group contributions sum 31 or above and 0 otherwise. We compute the average of this variable across rounds when we want to compare what happens across rounds, for instance from round 1 to 7 or from round 8 to 14. We use a non-parametric Man-Whitney U test (in brackets the round or rounds taken into account).

The centralized sanctioning mechanism works. So, hypothesis 1 is confirmed by our data, see rows (1) and (2).

**Table 2** Statistical results summary

| | Comparison | Z | *p*-Value |
|---|---|---|---|
| (1) | BL (1–7) – VR (1–7) | −3.16 | 0.00 |
| (2) | BL (1–7) – AR (1–7) | −3.31 | 0.00 |
| (3) | BL (8–14) – VR (8–14) | −2.11 | 0.04 |
| (4) | BL (8–14) – AR (8–14) | −3.00 | 0.00 |
| (5) | BL (1–7) – AR (8–14) | −2.76 | 0.01 |
| (6) | BL (1–7) – VR (8–14) | −1.46 | 0.14 |
| (7) | AR (1–7) – VR (1–7) | −0.71 | 0.48 |
| (8) | AR (8) – VR (8) | −0.46 | 0.65 |
| (9) | AR (8–14) – VR (8–14) | −2.07 | 0.04 |

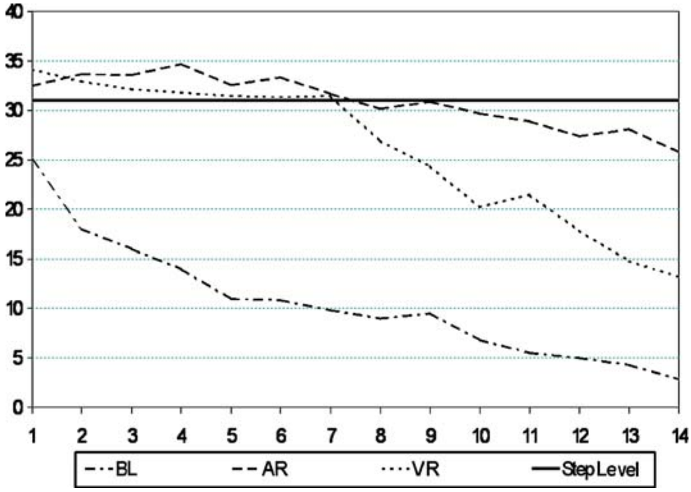


**Fig. 2** Average contribution per period

*Result 1:* in both AR and VR with CSM (rounds 1 to 7) contributions are significantly higher than in BL.

Every group in VR decided to play the game without CSM after round 7. The result of the vote varied from 5–0 to 3–2. However, people were not informed about the size of the majority, just about which was going to be the new set of rules according to the result of the vote.

*Result 2:* all nine groups in VR voted for removing the CSM after round 7.

Rows (5) and (6) show the test for the educational effect. The test supports that in AR subjects contribute significantly more after removal of the CSM than in the first 7 rounds of BL. This is not the case for VR. Remember, however, this is a test on averages and there are three groups in VR which managed to implement the public good until the last round. Note also that if we compare BL(8–14) to both AR(8–14) and VR(8–14), rows (3) and (4) we do find a significant difference. The comparison with BL (1–7) is however the most important one, as for subjects in AR and VR rounds 8–14 are the first rounds of playing without the CSM, as it is for subjects in BL in rounds 1–7.

*Result 3:* there is a clear educational effect if we compare AR with the baseline. Subjects learn to coordinate on an efficient way to produce the public good during the periods with the CSM and manage to keep doing this without the CSM.

Voting does not have a positive effect on contributions. The average success rate is actually higher in AR than in VR from rounds 8 to 14, row (9). However, AR and VR do not differ immediately after CSM disappears, round 8, see row (8). Note AR and VR average success rates are not significantly different under the CSM regime, rounds 1 to 7, row (7).

*Result 4:* contrary to Hypothesis 3, voting does not seem to have a positive effect in keeping contributions higher after CSM disappears.

Since groups in AR are actually very successful it would be difficult to obtain better results in VR. However, VR is significantly worse than AR. Possibly, participant who voted for and against the CSM behaved in a different way after CSM disappeared and that drove contributions down. This is not the case, a Man-Whitney U test on contributions of participants who voted to keep CSM and those who voted to remove CSM does not detect a significant difference at all ($Z = -0.39$, $p$-value $= 0.70$).

Another explanation for the relatively low contributions in VR after round 7 could be the influence of failures while the CSM is working. Indeed, if participants see his or her group is not able to reach the step-level even with CSM they might be not confident of reaching it without the CSM. So, a difference in failures to reach the step-level in VR compared to AR might explain a difference when the CSM is not present. We define a dummy variable indicating whether the step-level has always been reached ($Slar$ = step level always reached) in rounds 1–7 with the centralized sanctioning mechanism. It takes value 1 if group contributions had always been higher or equal than 31 in rounds 1 to 7. We consider data coming from both AR and VR, so $Slar$ equals 1 eight times and 0 ten times. The correlation between $Slar$ and the average success in rounds 8 to 14 is positive and quite high (.70). A simple OLS model yields the following results: $avg(8-14) = 0.243 + 0.507 * Slar$. Both the constant and the slope are significant. This regression can be understood as "failure to reach the step-level with CSM" increases the chance of failure in rounds 8 to 14 by 50.7%. We also look at whether even a single failure in rounds 1 to 7 with CSM has an effect. We define the variable $Jof$ (just once failed) which takes value 1 in case there had been just one failure in rounds 1 to 7 and 0 in case there had been no failure during rounds 1 to 7. We consider data coming from AR and VR, but note that now the domain is restricted by the fact that we are not taking into account groups with more than one failure in rounds 1 to 7. Therefore $Jof$ equals one six times and zero eight times. The correlation between $Jof$ and the average success in 8 to 14 is again positive and high (.74) . Again, we fit an OLS model: $avg(8-14) = 0.750 - 0.536 * Jof$. Both coefficients are significant. This regression can be understood as just one failure with the CSM increases the chance to fail in rounds 8 to 14 by 53.6%.

Comparing now the two treatments AR and VR (Table 3),we find that when no failure occurs in rounds 1 to 7, the number of times the step-level is reached does not differ between

**Table 3** Statistical results summary

| Comparison | $Z$ | $p$-Value |
|---|---|---|
| ARnf – VRnf (8–14) | −.900 | 0.37 |
| ARf – VRf (8–14) | −1.93 | 0.05 |

AR and VR in rounds 8 to 14. When failures occurred in rounds 1 to 7 however, success rates differ significantly between VR and AR in rounds 8 to 14. Both in AR and VR success rates are lower after removal of the punishment institution when in previous rounds failure occurred. However, in VR this effect is much stronger than in AR. Note also that failures are more frequent in VR (11) than in AR(7).

*Result 5:* a higher rate of failures in reaching the threshold during the CSM regime explains the lower average contributions in VR with respect to AR from round 8 to 14.

## 5. Conclusions

We designed two versions of a step-level public good game, with or without a centralized sanctioning mechanism (CSM). Cooperative equilibria are asymmetric in both games. In a baseline treatment participants played 14 rounds of the non-CSM game. In an automatic removal (AR) treatment participants played 7 rounds of CSM followed by 7 rounds of non-CSM. In a voted removal treatment (VR) participants played 7 rounds of CSM followed by a voting stage in order to decide by majority whether to play 7 more CSM rounds of 7 non-CSM rounds. Summarizing, we found:

- in both AR and VR with CSM (rounds 1 to 7) contributions are significantly higher than in BL
- all nine groups in VR voted for removing CSM after round 7
- there is a clear educational effect in the sense of learning to coordinate if we compare AR with the baseline
- voting does not seem to have a positive effect on contribution levels compared to automatic removal
- a higher rate of failures in reaching the step-level during the CSM regime explains the lower average contributions in VR with respect to AR from round 8 to 14.

Our results indicate that it is possible to learn how to coordinate around asymmetric cooperative equilibria by starting out with an external centralized sanctioning mechanism which is later removed. Secondly, it is important that this CSM induces a high success rate, something that is not easy to achieve in games with asymmetric cooperative equilibria. If the mechanisms fails to induce mutually beneficial outcomes the benefits in the long run are much smaller.

Overall, it appears there is a need for further clarification of the effects of the various sanctioning systems implementable in different social and economic contexts. Our experimental results support the possibility of self-governance, at least in small communities in which people had previously learned "good habits" from a Leviathan whether this role is served by family, school or any other. So, why feed the Leviathan? Besides the classical answer that enforcement institutions make life easier even if they are costly, we have a new one: the Leviathan may have educational effects leading to mutually beneficial results. In the sense of Kramer (1999), it can help to establish knowledge-based trust via deterrence-based trust. Once knowledge-based trust and efficient cooperation are established, a costly Leviathan is not necessary anymore. To obtain this result, however, the Leviathan has to be very good at its task.

## Appendix 1

Instructions for police removal (translated into English)

This is an economic decision making experiment. It is completely forbidden any kind of communication between you and the other participants once you have started to read these instruction sheet and until the end of the experiment.

Each of you will join a five people group. A computer will decide randomly which is the group you are going to join. This group will not change during the whole experiment.

The experiment will have 14 rounds. At the beginning of each round you will have 10 Experimental Count Units (ECUs). You must decide how many units to deposit in a fund and how many to keep. In order to make your decision your must take into account that:

(1) the contribution (C) to the fund must be an integer number between 0 and 10
(2) there exists a set of rules called "set of rules A". We are going to describe it:

  (a) if the sum of the contributions (SC) for your group of five people is bigger or equal to 31 then each individual will get 12 units, independently of their contribution
  (b) if your contribution to the fund is less than 6 your payoff will be reduced in 6 units
  (c) in any case all individual payoffs will be decreased by 3 units in their final payoff per round
  (d) therefore the payoff will be:

$$10 - C - 3 + 12 \quad \text{if } SC \geqslant 31 - 6 \quad \text{if } C < 6$$

(3) There exists a set of rules called "set of rules B". We are going to describe it:

  (a) if the sum of the contributions (SC) for your group of five people is bigger or equal to 31 then each individual will get 12 units, independently of their contribution. Therefore your payoff will be:

$$10 - C + 12$$

  (b) sum of the contributions (SC) for your group of five people is less than 31 then your payoff will be:

$$10 - C$$

(4) during the first 7 rounds the "set of rules A" will work
(5) at the end of the seventh round the members of each group will decide by majority whether to use the "set of rules A" or the "set of rules B" during the next 7 rounds

During the experiment you can see five types of screens:

  (i) the contributions screen where you may write your contribution in the corresponding field
  (ii) the results screen where you will see your contribution, the sum of the contributions, a message telling whether you got the 6 unit discount because of your low contribution and your payoff
  (iii) the voting screen (at the end of the seventh round) in order to choose among A or B
  (iv) the voting results screen

(v) the waiting screen

Your will receive 2 EUR just for showing up, plus 3 cent of EUR per UCE. If you have any doubt you may raise your hand now or during the experiment, we will attend you particularly.

## Appendix 2

**Table A.1** BL treatment data summary

| | G1 | G2 | G3 | G4 | G5 | G6 | Average |
|---|---|---|---|---|---|---|---|
| 1 | 27 | 24 | 19 | 25 | 29 | 26 | 25.0 |
| 2 | 9 | 10 | 24 | 30 | 14 | 21 | 18.0 |
| 3 | 0 | 11 | 16 | 34 | 12 | 23 | 16.0 |
| 4 | 0 | 6 | 10 | 35 | 11 | 22 | 14.0 |
| 5 | 0 | 1 | 8 | 31 | 7 | 19 | 11.0 |
| 6 | 1 | 0 | 13 | 26 | 8 | 17 | 10.8 |
| 7 | 2 | 1 | 11 | 22 | 11 | 12 | 9.8 |
| 8 | 1 | 0 | 5 | 15 | 22 | 11 | 9.0 |
| 9 | 10 | 0 | 4 | 15 | 22 | 6 | 9.5 |
| 10 | 6 | 0 | 1 | 10 | 16 | 8 | 6.8 |
| 11 | 0 | 0 | 1 | 18 | 7 | 7 | 5.5 |
| 12 | 0 | 0 | 1 | 13 | 5 | 11 | 5.0 |
| 13 | 0 | 0 | 1 | 9 | 13 | 3 | 4.4 |
| 14 | 0 | 0 | 2 | 10 | 3 | 2 | 2.8 |
| Avg. 1–7 | 5.6 | 7.6 | 14.4 | 29.0 | 13.1 | 20.0 | |
| Avg. 8–14 | 2.4 | 0.0 | 2.1 | 12.8 | 12.6 | 6.9 | |
| Average | 4.0 | 3.8 | 8.3 | 20.9 | 12.9 | 13.4 | |

**Table A.2** AR Treatment data summary

| | G1 | G2 | G3 | G4 | G5 | G6 | G7 | G8 | G9 | Average |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 32 | 36 | 35 | 33 | 28 | 26 | 34 | 34 | 34 | 32.4 |
| 2 | 30 | 37 | 35 | 34 | 31 | 29 | 34 | 37 | 36 | 33.7 |
| 3 | 33 | 37 | 35 | 32 | 32 | 33 | 31 | 35 | 34 | 33.6 |
| 4 | 31 | 32 | 35 | 34 | 36 | 41 | 34 | 35 | 34 | 34.7 |
| 5 | 33 | 33 | 34 | 32 | 28 | 34 | 31 | 33 | 35 | 32.6 |
| 6 | 32 | 33 | 36 | 32 | 32 | 36 | 33 | 33 | 33 | 33.3 |
| 7 | 31 | 31 | 30 | 33 | 33 | 28 | 32 | 34 | 33 | 31.7 |
| 8 | 23 | 33 | 33 | 26 | 24 | 29 | 33 | 34 | 36 | 30.1 |
| 9 | 24 | 32 | 34 | 31 | 32 | 30 | 32 | 30 | 33 | 30.9 |
| 10 | 10 | 32 | 37 | 32 | 31 | 29 | 33 | 31 | 32 | 25.7 |
| 11 | 4 | 30 | 36 | 31 | 31 | 30 | 31 | 33 | 34 | 28.9 |
| 12 | 2 | 27 | 32 | 32 | 30 | 31 | 32 | 29 | 31 | 27.3 |
| 13 | 0 | 29 | 32 | 32 | 32 | 31 | 33 | 32 | 32 | 28.1 |
| 14 | 0 | 18 | 26 | 31 | 32 | 32 | 26 | 33 | 34 | 25.8 |
| Avg. 1–7 | 31.7 | 34.1 | 34.3 | 32.9 | 31.4 | 32.4 | 32.7 | 34.4 | 34.1 | |
| Avg. 8–214 | 9.0 | 28.7 | 32.9 | 30.7 | 30.3 | 30.3 | 31.4 | 31.7 | 33.1 | |
| Average | 20.4 | 31.4 | 33.6 | 31.8 | 31.4 | 31.4 | 31.4 | 33.1 | 33.6 | |

**Table A.3** VR Treatment data summary

| | G1 | G2 | G3 | G4 | G5 | G6 | G7 | G8 | G9 | Average |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 33 | 33 | 39 | 34 | 32 | 34 | 33 | 32 | 37 | 34.1 |
| 2 | 27 | 34 | 36 | 32 | 32 | 31 | 31 | 34 | 39 | 32.9 |
| 3 | 35 | 33 | 32 | 33 | 32 | 33 | 26 | 33 | 32 | 32.1 |
| 4 | 25 | 34 | 28 | 34 | 32 | 32 | 32 | 31 | 38 | 31.8 |
| 5 | 27 | 32 | 30 | 32 | 32 | 31 | 33 | 33 | 33 | 31.4 |
| 6 | 28 | 31 | 34 | 33 | 31 | 33 | 31 | 30 | 31 | 31.3 |
| 7 | 30 | 27 | 35 | 31 | 33 | 33 | 32 | 32 | 30 | 31.4 |
| 8 | 15 | 26 | 24 | 33 | 30 | 32 | 26 | 29 | 27 | 24.9 |
| 9 | 9 | 14 | 16 | 28 | 32 | 29 | 32 | 26 | 33 | 24.3 |
| 10 | 3 | 5 | 5 | 33 | 33 | 24 | 31 | 22 | 26 | 20.2 |
| 11 | 0 | 6 | 20 | 33 | 32 | 38 | 29 | 26 | 19 | 21.4 |
| 12 | 10 | 3 | 1 | 30 | 32 | 29 | 29 | 17 | 9 | 17.8 |
| 13 | 0 | 3 | 0 | 31 | 32 | 35 | 18 | 7 | 7 | 14.8 |
| 14 | 0 | 1 | 0 | 33 | 33 | 35 | 11 | 0 | 6 | 13.2 |
| Avg. 1–7 | 29.3 | 32.0 | 33.4 | 32.7 | 32.0 | 32.4 | 31.1 | 32.1 | 34.3 | |
| Avg. 1-14 | 5.3 | 8.3 | 8.0 | 31.5 | 32.0 | 31.7 | 25.1 | 18.1 | 18.1 | |
| Average | 17.3 | 20.1 | 20.7 | 32.1 | 32.0 | 32.0 | 28.1 | 25.1 | 26.2 | |

**Acknowledgements** The authors thank Pablo Brañas-Garza and Jordi Brandts for their useful comments and kind help.

## References

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

Coleman, J. (1986). *Individual interest and collective action*. Cambridge University Press, Cambridge.

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

Fehr, E., Fischbacher, U., & Gaechter, S. (2002). Strong reciprocity, human cooperation and the enforcement of social norms. *Human Nature, 13*, 1–25.

Fischbacher, U. (1999). z-Tree: Zurich toolbox for readymade economic experiments. *Working Paper No. 21,* Institute for Empirical Research in Economics, University of Zurich.

Frey, B. (1998). *Not just for the money. An economic theory of personal motivation*. Edward Elgar Publishing Limited, Cheltenham 1998, 156 S.

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

Hobbes, T. (1998). *Leviathan*. Oxford World's Classics, Oxford.

Kramer, R. M. (1999). Trust and distrust in organizations: Emerging perspectives, enduring questions. *Annual Review of Psychology*, January 1999.

Kroll, S., Cherry, T.L., & Shogren, J.F. (2002). Voting, punishment, and public goods, an experimental investigation. *Working Papers from Department of Economics, Appalachian State University,* http://econpapers.repec.org/paper/aplwpaper/05-04.htm.

Lewicki, R.J., & Bunker, B.B. (1996). Developing and maintaining trust in work relationships. In R.M. Kramer., & T.R. Tyler. (Eds.), *Trust in organizations* (pp. 114–139). Sage Publications, Thousand Oaks.

Ostrom, E. (1998). A behavioral approach to the rational choice theory of collective action. *American Political Science Review, 92*(1).

Ostrom, E., Walker, J., & Gardner, R. (1992). Convenants with and without a Sword: self governance is possible. *American Political Science Review, 86*(2).

Rapoport, A., & Au, W.T. (2001). Bonus and penalty in common pool resource dilemmas under uncertainty. *Organizational Behavior and Humand Decision Processes, 85*(1), 135–165.

Yamagishi, T. (1986). The provision of a sanctioning system as a public good. *Journal of Personality and Social Psychology, 51*(1), 110–116.

Yamagishi, T. (1988). Seriousness of social dilemmas and the provision of a sanctioning system. *Social Psychology Quarterly, 51*(1), 32–42.