

# Political Accountability and Democratic Institutions: An Experimental Assessment\*

José Gabriel Castillo

ESPOL-Escuela Superior Politécnica del Litoral  
College of Social Sciences and Humanities  
jcastil@espol.edu.ec

John Hamman

Florida State University  
Department of Economics  
jhamman@fsu.edu

This version: June, 2019

## Abstract

We study the extent to which centralized democratic institutions enhance collective action under political accountability. In a public goods game with costly punishment, we vary the appointment of one group member to enforce punishment. Specifically, we compare democratically elected punishers to those appointed exogenously, under both single- and multiple-election environments. We find that democratically appointed sanctioning authority has muted effects on group outcomes. Yet, democratically elected authorities contribute as much as other group members when facing repeated elections, as opposed to the ones in single elections or exogenously appointed. One important feature of modern governance to discipline authorities is political accountability; when in place, it offers different incentives, and in particular, we observe a *responsibility effect* reflected in higher contribution behavior. Important in our study results, this effect rises only under a democracy.

**JEL:** C92, D72, H41

**Keywords:** public goods game, democracy, political accountability.

---

\*We would like to thank fellows at LAWEBESS in Santiago de Chile for their comments. Also, we acknowledge the financial support of the project CIEC-19-2016, the Laboratory for Experimental and Behavioral Economics (L.E.E.) and the Research Funds (R-CD-FCSH-135-2017) of the School of Social Sciences and Humanities at ESPOL. Special thanks to Donald Zhangallimbay and Jonathan Parrales for excellent research assistance and help running the sessions. The corresponding disclaimer applies. All errors are our own.

# 1 Introduction

Modern societies enforce collective action through delegation of sanctioning duties, which relies on the legitimacy of authority to promote socially desired outcomes.<sup>1</sup> And yet, whether for general economic outcomes or organizational behavior and performance, we know little about “the precise causal mechanisms through which the type of governance affects individual behavior” (Hargreaves Heap et al., 2015). In particular, do democratic institutions succeed by selecting the best delegates, or does participative democracy have an effect on accountability beyond selection? In this paper, we experimentally control for this selection effect to examine whether electoral political accountability functions as a source of institutional legitimacy to promote collective action.

Much research on democratic institutions focuses on direct democracy, in which participants vote to directly implement group-wide decisions. Walker et al. (2000) and Kroll et al. (2007) find that direct democracy increases contributions to public goods, though DeAngelo et al. (2018) show that majority coalitions may use direct democracy to exacerbate inequality. A similar strand of research explores direct voting over sanctioning institutions and other consequential actions (Botelho et al., 2005; Ertan et al., 2009; Cinyabuguma et al., 2005; Tyran and Feld, 2006; Ambrus and Greiner, 2015). Sutter et al. (2010) compares the effectiveness of sanctions and rewards as determined via exogenous or endogenous selection, finding that both incentives have a larger effect when endogenously chosen by participants (see also Sefton et al., 2007).

A smaller literature examines democratic and other forms of delegation of authority. Democratic delegation of group contributions in collective action environments has been shown to effectively resolve the free-riding problem (Hamman et al., 2011; Bolle and Vogel, 2011; Fleiß and Palan, 2013). Other recent work shows that a centralized sanctioning authority can lead to more efficient outcomes than decentralized sanctioning (Andreoni and Gee, 2012; Baldassarri and Grossman, 2011). One question that arises from these studies, which we address using laboratory experiments, is to what extent the method of appointing a central authority matters for a group’s outcome.

---

<sup>1</sup>For a recent review of this literature, see Van Lange et al. (2014).

Contrary to the common belief about the causal influence of democratic institutions on collective action and economic outcomes, experimental evidence on delegated enforcement remains inconclusive. Baldassarri and Grossman (2011) and Grossman and Baldassarri (2012) use a lab-in-the-field public goods game with centralized punishment, finding that democratically appointed sanctioning agents increases public good contributions relative to randomly appointed agents. However, Castillo et al. (2018) finds no difference between elected and exogenously appointed sanctioning agents in two experimental environments with different levels of sanctions' efficacy, and Hargreaves Heap et al. (2015) similarly find no difference in group outcomes under democratic and dictatorial decision rules. Beyond known experimental nuances, the source of legitimacy in democratic institutional arrangements remains a challenge for causal inference.

In the political science literature, several mechanisms for legitimacy have been analyzed. These include the type of leadership (Grossman, 2014), quality of governance and public information (Adserà et al., 2003), and political competition (Zudenkova, 2018). Less is known about the influence of the mechanisms of political accountability in building a leader's legitimacy - in fact, Ferejohn (1999) acknowledges that informational advantages that leaders have can be used to reduce responsiveness without reducing legitimacy. Under representative institutions, political accountability rises as the primary mechanism to hold leaders responsible for their political agenda and assure their actions remain aligned with the public's best interests (Grossman and Baldassarri, 2012; A. Huber and Gordon, 2004). If citizens dislike the incumbent's performance, they may seek a replacement in the following election. In line with such retrospective voting, we hypothesize that political accountability builds legitimacy only through a democratic institutional scheme, that is, only if subjects are called to act through voting; hence, they are politically involved in the leader's election process, as opposed to an automatic (exogenous) political refreshment by the end of the incumbent's term.

To study this relationship between democratic elections and political accountability, we conduct a laboratory experiment using a two-by-two between-subjects design where subjects play a public goods game with and without punishment opportunities. In one treatment dimension, we

vary whether the central sanctioning authority is elected by the group or exogenously appointed. Here we make a novel contribution to the study of Hobbesian versus democratic institutions by designing this random process - described in greater detail in section 2 - to control for the selection effects of voting (i.e., signaling). The second dimension varies the frequency with which elections or appointments are made, either once for the duration of the session or every three periods. This allows us to study one commonly attributed benefit to democratic processes, in which re-election concern incentivizes the authority to act in the electorate's best interests (Ferejohn, 1986).

We find that democratic elections impact political accountability only for an official's actions, but not for the behavior of their constituents. Specifically, when democratically elected authorities must face repeated elections, they contribute as much as constituents to the public good. In all other treatments, we observe strong free riding by the sanctioning authority. We see no such differences in contributions by other group members across treatments. In line with Castillo et al. (2018) and Hargreaves Heap et al. (2015), our results suggest muted effects of democratic systems once we control for the quality of the appointed leader.

In the following section we provide the details for our experimental design and specify our analytical approach. We report our results in section 3 and offer concluding remarks in section 4.

## **2 Design and procedures**

### **2.1 Experimental design**

The design extends the framework of Fehr and Gächter (2000), using the centralized tax/punishment environment proposed by Castillo et al. (2018) with between-subjects treatments that vary on two dimensions: the power delegation mechanism and the political accountability institution.

#### **2.1.1 General framework**

Table 1 summarizes the sample and treatments for the experimental design. In each cell subjects face two phases: i) standard "linear" voluntary contribution mechanism (VCM, henceforth), that

runs for 10 identical rounds, and; ii) the centralized punishment institution, that runs for 10 rounds on the single election treatment and 12 rounds in the multiple elections treatment. Instructions for the second phase are distributed only after the first phase finishes, to avoid strategic decisions.

Participants in the first phase receive an endowment of  $w = 20$  experimental units (EU) in each decision round. They can contribute  $c$  to a “public good” which constitutes a pool with all group members’ contributions, yielding revenue defined by a multiplier ( $m$ ); in our experiment contributions increase by a multiplier of two ( $m = 2$ ) and are divided equally among  $n$  group participants ( $n = 5$ ). This implies a marginal per capita return of 0.4 ( $MPCR(\alpha) = m/n$ ).<sup>2</sup> Each subject faces the trade off between keeping the endowment and free ride on contributions from his partners or contribute to the public account; that is, he faces three alternatives:  $c = 0$ , which represents the dominant strategy (i.e. Nash equilibrium);  $c = w$ , which constitutes the socially optimal decision (i.e. Pareto solution), and  $0 < c < w$ .

The individual’s  $i$  payoff function in period  $t$  can be summarized in the following:

$$\pi_{it} = 20 - c_{it} + 0.4C_t \quad (1)$$

where  $C_t = \sum_{j=1}^n c_{jt}$ , the sum of all members’ contributions to the group account.

In the second phase, the centralized tax/punishment institution, each period has two stages. During the first stage, subjects face the standard VCM from phase 1. In the second stage, one subject, which we call “the manager,” administers the management account funded by a tax of two EU ( $\tau = 2$ ), automatically collected from each group member; that is, there are 10 points available in each round. The manager uses the points accumulated to decide over whether or not imposing punishment to his fellow group members, and; if so, how many points and to whom they will be directed. There is no institutional inefficiency and unused points from the management account return in equal parts to each group member. To allow for better punishment efficacy, punishment points “assigned” are transformed through a convex punishment cost function to punishment points

---

<sup>2</sup>The MPCR satisfies a basic condition:  $0 < \alpha < 1 < n\alpha$ ; hence, it is socially efficient to contribute all the endowment to the public good if  $n\alpha > 1$ .

“reduced” as follows:<sup>3</sup>

Points ASSIGNED ( $p$ )	0	1	2	3	4	5	6	7	8	9	10
Points REDUCED ( $p^*$ )	0	1	2	4	6	9	12	16	20	25	30

Note that, in extreme cases, negative earnings in a round are possible. To reduce the impact of negative payoffs, subjects are allowed to loose either the payoff result or the number of punishment points assigned (not reduced), whichever is lower in absolute terms. Yet, we do not observe any such instances.

The payoff function of subject  $i$  for the second phase looks as follows:

$$\pi_{it} = \begin{cases} \underbrace{20 - c_{it} + 0.4C_t}_{\text{VCM}} - \underbrace{2 - p_{it}^* + \frac{1}{5}(10 - p_{jt})}_{\text{tax/punishment mechanism}} & ,\text{if } \pi_{it} \geq 0 \\ \min\{|p_{it}|, |\pi_{it}|\} & ,\text{if } \pi_{it} < 0 \end{cases} \quad (2)$$

The general framework is one of *ex-post* full information. All group members, including the manager, see their actions and payoffs following each round, including any reduction in earnings resulting from punishment point. Participants also receive feedback about others’ contributions and profits, anonymously in random order, in each period. They also observe the total punishment points used but not to whom they were targeted.

Note that managers face stronger incentives to free ride since self punishment is not allowed. This allows us to observe how they trade off these benefits with long-term incentives to maintain accountability.<sup>4</sup>

<sup>3</sup>Some debate always arises among skeptics on the cost function chosen. Although Nikiforakis and Normann (2008) show its relevance, under the same decision environment as in this document, with no political accountability, Castillo et al. (2018) show that punishment efficacy is inconsequential.

<sup>4</sup>This tension is similar to that used in Cooper et al (2019), which studies leadership in situations involving incentive conflict.

### 2.1.2 Treatments

The first treatment dimension corresponds to the *centralized power delegation mechanism*. Here we analyze whether the manager selection mechanism affects the behavior of group members. In the exogenous power delegation mechanism, which we call the *Leviathan*, one group member is selected as a manager by the experimenter. The selection process is calibrated in order to isolate the potential effect of the manager's quality on the behavior of participants; based on calibrations from previous data (Castillo et al., 2018), the probability of choosing the highest phase 1 contributor is 75%.<sup>5</sup> Subjects are informed about this process and observe phase 1 contribution averages for their group. The second treatment corresponds to the endogenous power delegation, or *Democracy*. By plurality vote, subjects elect one group member, after observing their contributions during the first phase (the VCM).<sup>6</sup> Votes are cast simultaneously and anonymously, with ties randomly broken by the software.

The second dimension is the *political accountability institution*. The first treatment is the *Single Election* framework. Managers are elected by one of the described mechanisms and, once elected, they remain in-office permanently. We introduce the second treatment in order to assess the main hypothesis of this research; the *Multiple Elections* framework allows for manager selection every three rounds; to even the decision rounds we extend the periods to 12. Feedback now is based on performance of every player on the previous three rounds, again anonymously, except for the first period where information is based on the 10 periods of the first phase.

---

<sup>5</sup>This feature, we argue, is crucial in the design. In the endogenous treatment individuals act upon the contribution information provided by voting, hence this decision reflects the preference over the perceived quality of the manager; while in the exogenous selection they are only informed on the selection result. A fully random assignment in the *Leviathan* would bias behavior since subjects might perceived a higher risk of a bad quality manager in-office.

<sup>6</sup>We confirm this information is relevant for the decision making process by asking the participants a set of open questions at the end of the experiment. The majority of subjects focus on the number of points contributed to the public good as the reference for the election process and the manager's quality. Some other interesting expectations over high contributors are: subjects expect high contributors to manage better the public account and to think more on others' wellbeing; also, they are attributed some personal traits such as intelligence and generosity, giving the sense of deservedness of the appointment.

## 2.2 Data

Sessions were conducted in the Laboratory for Experimental and Behavioral Economics (L.E.E.) at ESPOL-Polytechnic University, in Guayaquil-Ecuador, between January and September 2018. We used O-Tree (Chen et al., 2016) as the computer interface, and the recruitment process was performed through ORSEE (Greiner, 2015).

Table 1 summarizes the sample distribution in each cell. A total of 435 subjects were recruited; 200 subjects for the single election treatment and 235 subjects for the multiple elections treatment. Participants are inexperienced subjects; 45% are women; the mean age is 21 and ages fluctuate between 17 and 32. Thirty five percent are economics majors, with the rest distributed among careers in engineering and STEM.

Sessions lasted for around 90 minutes and subjects received a show-up fee of USD 2.00, for an average total earnings of around USD 13.00.

## 2.3 Empirical analysis

To provide a complete empirical analysis, we employ both nonparametric tests and formal econometric methods.

Since each treatment is randomly administered by session, we analyze mean differences between treatments and phases, directly through the Mann-Whitney U-test (Wilcoxon-Mann-Whitney) at group level. This is the main approach for our results on the differences between the two political accountability institutions.

We extend the analysis econometrically to control for possible confounds within each treatment of the power delegation mechanism. Since the design allows to collect longitudinal data, we include a fully saturated specification with several controls and fixed effects at various levels. This is a Difference-in-Difference approach; to analyze the contribution determinants we estimate an equation as follows:



$$C_{igt} = \alpha_1(D * P)_{it} + \alpha_2 Democracy(D)_i + \alpha_3 Punishment(P)_t \\ + X'_{ig}\Lambda + Z'_i\Gamma + \phi_g + \tau_t + \varepsilon_{igt}$$

where  $C_{igt}$  is the contribution level of subject  $i$ , in group  $g$ , in period  $t$ . The coefficient of interest for the average treatment effect (ATE) of the endogenous power distribution under centralized punishment institution is  $\hat{\alpha}_1$ .  $\alpha_2$  represents the average effect of the democratic power delegation;  $\alpha_3$  is the average effect of a centralized punishment institution;  $\Lambda$  is a vector controls for individual behavior within each group;  $\Gamma$  is the vector of individual controls (individual fixed effects in the most flexible case);  $\phi_g$  are group fixed effects;  $\tau_t$  are dynamic time trends within each phase; and  $\varepsilon_{igt}$  the *i.i.d.* idiosyncratic error term.

### 3 Results

#### 3.1 Do democratic elections incentivize collective action?

The first thing to note in our analysis is that an endogenous (i.e., democratic) power delegation does not trigger intrinsic motivation to improve contribution behavior in a centralized management environment, in line with other recent work (Hargreaves Heap et al., 2015; Castillo et al., 2018).

Figure 1 shows the average contribution dynamics of our experiment. Note first that the centralized punishment institution yields results similar to the decentralized case, which is well documented in prior literature (Fehr and Gächter, 2000; Fehr and Gächter, 2000; Ledyard, 1995; Puterman et al., 2011; Nikiforakis and Normann, 2008). Once imposed, the mechanism promotes higher and more stable levels of cooperation than without punishment opportunities. Second, we observe no significant differences between power delegation mechanisms, whether or not the political accountability institution is imposed.<sup>7</sup>

---

<sup>7</sup>In the graph, the average contribution switches between treatments, but no conclusion can be extract due to statistical insignificance.

Table 2 summarizes the main results. Each cell of Panel A shows the average difference between the centralized punishment institution and the VCM. We answer the first question by comparing vertically. The Democracy-Leviathan row shows the average difference between the power delegation mechanisms. We observe that differences between the democratic scheme and the leviathan are not statistically significant, regardless of the possibility for political accountability within the institutional framework. In Table 3 we extend the analysis econometrically to show that results are consistent over multiple specifications. As in the seminal paper of Fehr and Gächter (2000), a significant effect comes from the punishment institution imposed, but there are no differences of the power delegation mechanism whether in a single election or a multiple election environment. Also, the higher the past contributions of others within a group, the higher the observed contribution, which again aligns with prior findings of conditional cooperation as an emergent social phenomenon.

Observed results for the first dimension of the analysis can not be explained by differences in punishment behavior. Figure 2 shows various punishment measures for both treatments, which reveal no evidence of significant differences. Panel B of Table 2 and Table 4 support this conclusion. The only significant difference observed is on the extensive margin; that is, the probability of being punished slightly diminishes in the single election treatment under a democratic scheme; however, this result disappears in the presence of political accountability opportunities (see Figure 2e and Figure 2f). There are neither differences in the intensive margin, punishment points used are not statistically different between power delegation mechanisms (see Figure 2a and Figure 2b); nor in the manager's use of punishment (see Figure 2c and Figure 2d). Deviations from the social norm (i.e. the group's average contribution) intervene in the probability of been punished in the expected way; negative deviations increase the probability and intensity of punishment in both treatments, the single and multiple elections institutions.

Putting things together, the welfare measure of our framework can be summarized in the subject's profit, that is, the net payoff received after punishment. Panel C of Table 2 shows these results. Conclusions remain.

### 3.2 Does political accountability of sanctioning authority affect behavior?

This question can be sliced into two different aspects of the framework's incentives: contribution behavior of the group, and; contribution behavior of the managers.

To answer the first part, we return to the main results in Table 2, only this time we concentrate on the comparison across columns. On the previous section we show that the democratic scheme has no effect on the general contribution behavior regardless of the single or multiple election scheme; in other words, opportunities for political accountability (i.e. multiple election) does not add any differential incentive to the democratic process to promote collective action. Results in the third column in Panel A show whether there are behavioral differences across treatments on the second dimension of the analysis, that is, the institution of political accountability. Reinforcing the previous conclusions, observed differences between the single and multiple treatment are not statistically significant, regardless of an exogenous power delegation mechanism (i.e. the Leviathan) or an endogenous mechanism (i.e. the Democracy).

The sequential selection process, in the multiple elections' case, adds little to the contribution dynamics. Every three periods, contributions tend to increase on the electoral period; this is more clearly seen in the Democracy (Figure 1b). Given the design conditions, subjects in both treatments can signal, either their peers or the experimentalist selection rule, to increase their selection probabilities by increasing their contributions.<sup>8</sup> Again, observed differences are not statistically significant, in particular once netting out the first phase behavior (i.e. VCM).

The second part of the question sheds some light on the relationship of the political accountability institution and the manager's selection mechanism. Figure 3 decomposes the contribution dynamics of manager's and non-manager's by each dimension of the experimental design. As expected, on the first phase (i.e. the VCM) subjects selected as manager's are usually the highest contributors in both power delegation schemes (see Figure 4); on the other hand, manager's contribution differences, observed in the first phase, do not translate into the second phase for a central-

---

<sup>8</sup>Note that this is also the case in the Leviathan treatment, given the 75% probability of choosing the highest contributor.

ized punishment environment. The contribution dynamics of all treatments shrinks, regardless of the roles of the group members. Table 5 offers a formal tests of the mean changes in contribution between phases within the two dimensions of the design, by roles subject's play within the group. The take-up from the table's results is that manager's, on average, exercise their edge in free riding opportunities and contribute significantly less in all treatments, except in the democratic power delegation under the political accountability institution; mean differences between managers and non-managers are not statistically significant in this single treatment (last row). In other words, although political accountability opportunities do not have an effect on constituents (i.e. other group members), it does matter in terms of the manager's behavior, conditional on being in a democratic power delegation environment.

## 4 Conclusion

There is a rich empirical literature examining the benefits of democratic institutions in providing public goods, often through direct democracy. We contribute to this literature by examining the effects of democratic accountability on sanctioning authority. In contrast to democratically determined contributions, we find that democratically appointed sanctioning authority has muted effects on group outcomes. When we control for the quality of the appointed authority, we find no difference in group outcomes between democratic and exogenous mechanisms.

However, we do find that democratically elected authorities facing repeated elections no longer free ride. Instead, their contributions are in line with those of other group members. In contrast, democratically elected authorities who do not face repeated election (i.e., in the absence of political accountability), as well as exogenously appointed authorities in both single- and multiple-election environments, contribute significantly less to the public good than their fellow group members.

We conjecture that strong beliefs over the advantages of democratic institutions in centralized power environments rely on features that either act jointly or are independent of the power delegation mechanism. One important feature of modern governance is political accountability; when

in place, it offers different incentives to the authorities, in particular, what we refer to as a *responsibility effect* reflected in higher contribution behavior. Important in our study results, this effect rises only under a democracy.

## References

- A. Huber, G. and S. C. Gordon (2004). Accountability and coercion: Is justice blind when it runs for office? *American Journal of Political Science* 48(2), 247–263.
- Adserà, A., C. Boix, and M. Payne (2003, 10). Are you being served? political accountability and quality of government. *The Journal of Law, Economics, and Organization* 19(2), 445–490.
- Ambrus, A. and B. Greiner (2015). Democratic punishment in public good games with perfect and imperfect observability. *Economic Research Initiatives at Duke (ERID) Working Paper* (183).
- Andreoni, J. and L. Gee (2012). Gun for hire: Delegated enforcement and peer punishment in public goods provision. *Journal of Public Economics* 96(11-12), 1036–1046.
- Baldassarri, D. and G. Grossman (2011). Centralized sanctioning and legitimate authority promote cooperation in humans. *Proceedings of the National Academy of Sciences* 108(27), 11023–11027.
- Bolle, F. and C. Vogel (2011, Sep). Power comes with responsibility—or does it? *Public Choice* 148(3), 459–470.
- Botelho, A., G. W. Harrison, L. Pinto, and E. E. Rutström (2005). Social norms and social choice.
- Castillo, J., Z. Xu, P. Zhang, and X. Zhu (2018). The effects of centralized power and institutional legitimacy on collective action. Technical report, ESPOL. Available at SSRN: <https://ssrn.com/abstract=3127775>.

- Chen, D. L., M. Schonger, and C. Wickens (2016). otree- an open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance* 9(Supplement C), 88 – 97.
- Cinyabuguma, M., T. Page, and L. Putterman (2005). Cooperation under the threat of expulsion in a public goods experiment. *Journal of Public Economics* 89(8), 1421 – 1435. The Experimental Approaches to Public Economics.
- DeAngelo, G. J., D. Dubois, and R. Romaniuc (2018). The perils of democracy. *Journal of Economic Behavior & Organization*.
- Ertan, A., T. Page, and L. Putterman (2009). Who to punish? individual decisions and majority rule in mitigating the free rider problem. *European Economic Review* 53(5), 495–511.
- Fehr, E. and S. Gächter (2000). Cooperation and punishment in public goods experiments. *American Economic Review* 90(4), 980–994.
- Fehr, E. and S. Gächter (2000). Fairness and retaliation: The economics of reciprocity. *Journal of economic perspectives* 14(3), 159–181.
- Ferejohn, J. (1986, Jan). Incumbent performance and electoral control. *Public Choice* 50(1), 5–25.
- Ferejohn, J. (1999). Accountability and authority: toward a theory of political accountability. *Democracy, accountability, and representation* 131, 133.
- Fleiß, J. and S. Palan (2013). Of coordinators and dictators: A public goods experiment. *Games* 4(4), 584–607.
- Greiner, B. (2015). Subject pool recruitment procedures: Organizing experiments with orsee. *Journal of the Economic Science Association* 1(1), 114–125.
- Grossman, G. (2014). Do selection rules affect leader responsiveness? *Quarterly Journal of Political Science* 9(1), 1–44.

- Grossman, G. and D. Baldassarri (2012). The impact of elections on cooperation: Evidence from a lab-in-the-field experiment in Uganda. *American Journal of Political Science* 56(4), 964–985.
- Hamman, J. R., R. A. Weber, and J. Woon (2011). An experimental investigation of electoral delegation and the provision of public goods. *American Journal of Political Science* 55(4), 738–752.
- Hargreaves Heap, S. P., K. Tsutsui, and D. J. Zizzo (2015). An experiment on democratic versus dictatorial collective decision making. Technical report, UEA, U Bath, U Queensland.
- Kroll, S., T. L. Cherry, and J. F. Shogren (2007). Voting, punishment, and public goods. *Economic Inquiry* 45(3), 557–570.
- Ledyard, J. O. (1995). Public goods: A survey of experimental research. In J. Kagel and A. Roth (Eds.), *Handbook of Experimental Economics*, Handbook of Experimental Economics. Princeton University Press.
- Nikiforakis, N. and H.-T. Normann (2008). A comparative statics analysis of punishment in public-good experiments. *Experimental Economics* 11(4), 358–369.
- Persson, T., G. Roland, and G. Tabellini (1997). Separation of powers and political accountability. *The Quarterly Journal of Economics* 112(4), 1163–1202.
- Putterman, L., J.-R. Tyran, and K. Kamei (2011). Public goods and voting on formal sanction schemes. *Journal of Public Economics* 95(9), 1213–1222.
- Sefton, M., R. Shupp, and J. Walker (2007). The effect of rewards and sanctions in provision of public goods. *Economic Inquiry* 45(4), 671–690.
- Sutter, M., S. Haigner, and M. G. Kocher (2010). Choosing the carrot or the stick? endogenous institutional choice in social dilemma situations. *Review of Economic Studies* 77(4), 1540–1566.
- Tyran, J.-R. and L. P. Feld (2006). Achieving compliance when legal sanctions are non-deterrent\*. *Scandinavian Journal of Economics* 108(1), 135–156.

Van Lange, P. A., B. Rockenbach, and T. Yamagishi (2014). *Reward and Punishment in Social Dilemmas*. Oxford University Press.

Walker, J. M., R. Gardner, A. Herr, and E. Ostrom (2000). Collective choice in the commons: Experimental results on proposed allocation rules and votes. *The Economic Journal* 110(460), 212–234.

Zudenkova, G. (2018, August). Electoral accountability and interdistrict competition. *American Economic Journal: Microeconomics* 10(3), 143–76.



## Tables and figures

Table 1: Experimental design and (preliminary) sample description

Treatments	Single election <i>n=subjects (groups)</i>	Multiple elections <i>n=subjects (groups)</i>
Leviathan	100 (20)	115 (23)
Democracy	100 (20)	120 (24)
<b>Total sample</b>	<b>200 (40)</b>	<b>235 (47)</b>

Table 2: Average Performance Comparison

<b>Panel A: Contributions (points)</b>			
<b>Punishment-VCM</b>			
Treatments	Single election	Multiple elections	Single-Multiple
Leviathan	2.863 (1.035) [ $p = 0.029$ ]	2.599 (0.880) [ $p = 0.006$ ]	0.264 (0.772) [ $p = 0.961$ ]
Democracy	3.046 (0.967) [ $p = 0.005$ ]	1.931 (1.013) [ $p = 0.046$ ]	1.114 (0.870) [ $p = 0.195$ ]
Democracy-Leviathan	0.183 (0.892) [ $p = 0.850$ ]	-0.667 (0.760) [ $p = 0.395$ ]	

<b>Panel B: Punishment (points)</b>			
<b>Democracy-Leviathan</b>			
Experiments	Single election	Multiple election	Single-Multiple
	-0.076 (0.147) [ $p = 0.560$ ]	0.027 (0.125) [ $p = 0.831$ ]	-0.060 (0.208) [ $p = 0.664$ ]

<b>Panel C: Profits (points)</b>			
<b>Punishment-VCM</b>			
Treatments	Single Election	Multiple Election	Single-Multiple
Leviathan	0.576 (0.983) [ $p = 0.956$ ]	0.295 (0.846) [ $p = 0.684$ ]	0.426 (0.417) [ $p = 0.733$ ]
Democracy	0.833 (0.873) [ $p = 0.304$ ]	-0.474 (1.043) [ $p = 0.327$ ]	1.307 (0.853) [ $p = 0.157$ ]
Democracy-Leviathan	0.256 (0.916) [ $p = 0.903$ ]	-0.770 (0.787) [ $p = 0.371$ ]	

Notes: Panels A and C report within subjects differences between phase 2 (punishment) and phase 1 (VCM). Panel B reports differences between treatments.

Group-clustered standard errors in parentheses.

p-values are reported in brackets for a Mann-Whitney U tests. Two-sided t-tests report similar results.

Table 3: Determinants of Contributions

Dep. Variable=Contributions (points)	Single election		Multiple elections	
	FEgt	FEgtc	FEgt	FEgtc
Democracy vs. Leviathan (P*D)	0.1785 (0.9035)	0.1636 (0.5920)	-0.6672 (0.7671)	-0.4069 (0.5224)
Punishment	2.8675*** (0.6080)	2.1510*** (0.3940)	3.9856*** (0.4184)	3.1360*** (0.3144)
Other member's Av. Contribution (t-1)		0.3883*** (0.0479)		0.3558*** (0.0406)
Punishment Received (t-1)		0.3658 (0.4247)		-0.0816 (0.1109)
Punishment in the group (t-1)		-0.1574 (0.1316)		0.1586 (0.3283)
Other controls	No	No	No	No
Individual FE	Yes	Yes	Yes	Yes
Group FE	Yes	Yes	Yes	Yes
Trend within phase	Yes	Yes	Yes	Yes
R.squared	0.4233	0.4712	0.4369	0.4736
Observations	4000	3800	5170	4935

Notes: Dummy variable for the Democracy treatment (D) excluded since it is time invariant; hence it has a null coefficient for a Fixed Effect (FE) estimation.

Standard errors clustered at group level in parentheses.

\* Significant at the 10 percent level.

\*\* Significant at the 5 percent level.

\*\*\* Significant at the 1 percent level.

Table 4: Punishment Decision

Dependent variable:	Single election		Multiple elections	
	(1) Punished=1	(2) Punishment Points	(3) Punished=1	(4) Punishment Points
Democracy (D)	-0.0843** (0.0380)	-0.0190 (0.2791)	-0.0361 (0.0254)	0.1173 (0.1962)
OMC negative deviation	0.0425*** (0.0033)	0.3795*** (0.0315)	0.0300*** (0.0031)	0.3474*** (0.0303)
OMC positive deviation	-0.0294*** (0.0042)	0.0448 (0.0542)	-0.0229*** (0.0036)	0.0079 (0.0245)
R.squared (overall)		0.444		0.3450
Observations	1999	698	2820	883

Notes: Coefficients in models 1 and 3 report the marginal effects (at means) of the probability of being punished for a Panel Data Probit model to capture the within individual correlation.

OMC=Other members' Average Contribution.

Standard errors clustered at group level in parentheses for models 2 and 4. Robust standard errors for models 1 and 3.

\* Significant at the 10 percent level.

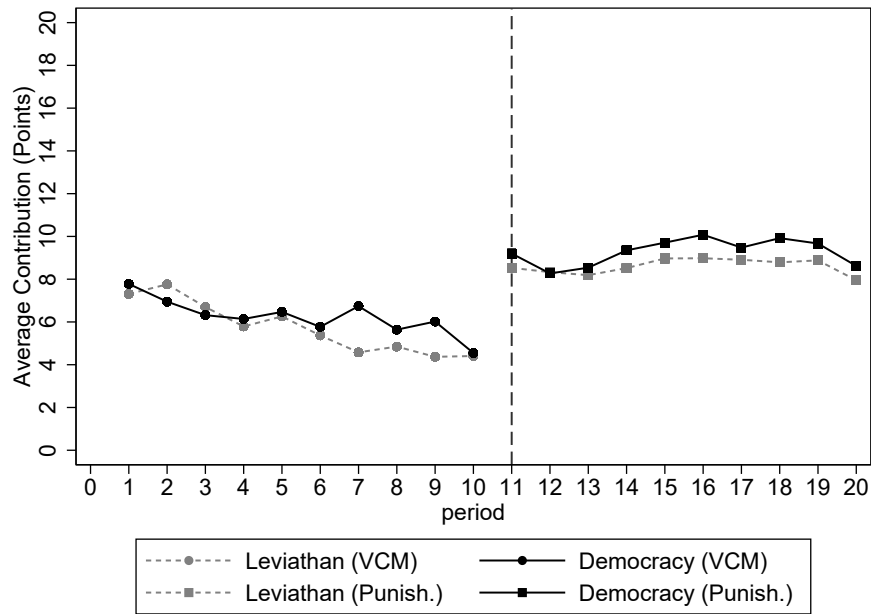
\*\* Significant at the 5 percent level.

\*\*\* Significant at the 1 percent level.

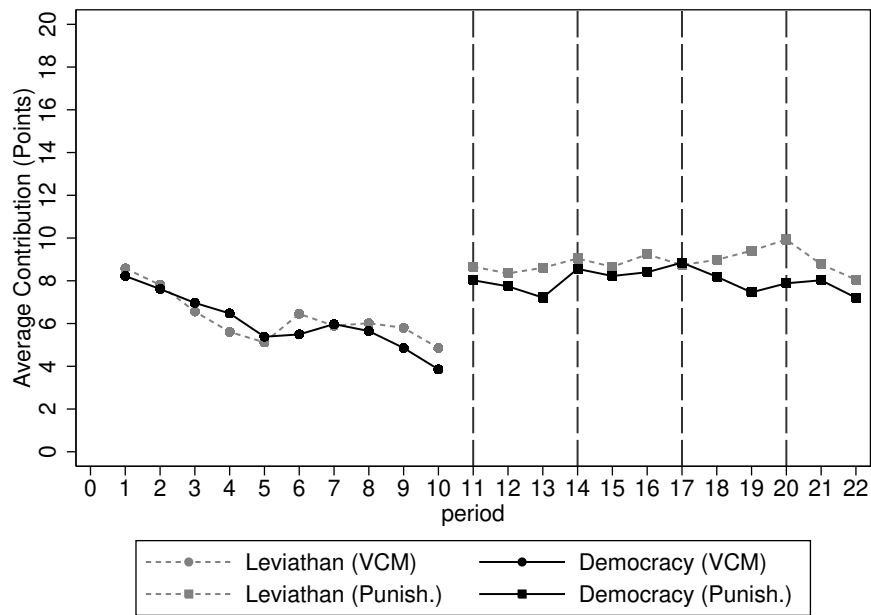
Table 5: Leadership contribution analysis

	Punishment-VCM	Manager	Others	<i>p</i> -value
				Mann-Whitney U tests ( $H_0$ : equal means)
Single election	Overall	0.870(0.370)	3.476(0.177)	0.000
	Leviathan	0.395(0.522)	3.482(0.231)	0.000
	Democracy	1.345(0.525)	3.471(0.269)	0.000
Multiple elections	Overall	2.006(0.315)	2.423(0.159)	0.061
	Leviathan	1.926(0.494)	2.881(0.231)	0.016
	Democracy	2.083(0.396)	1.985(0.216)	0.832

Notes: Standard errors clustered at group level, in parentheses. Multiple elections has 10 rounds in phase 1 and 12 rounds in phase 2; hence, for equal comparison, we only take into account the difference in contribution until round 20.

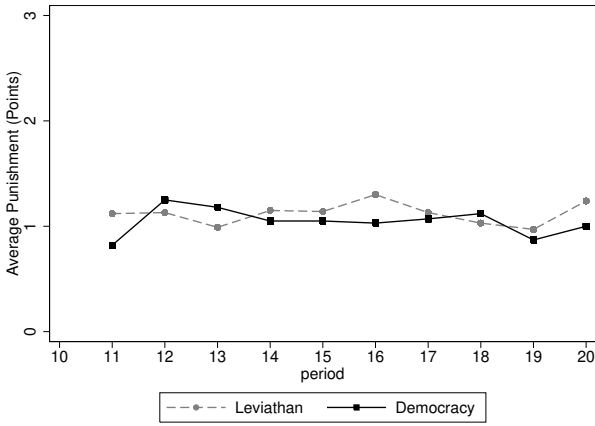


(a) Single election

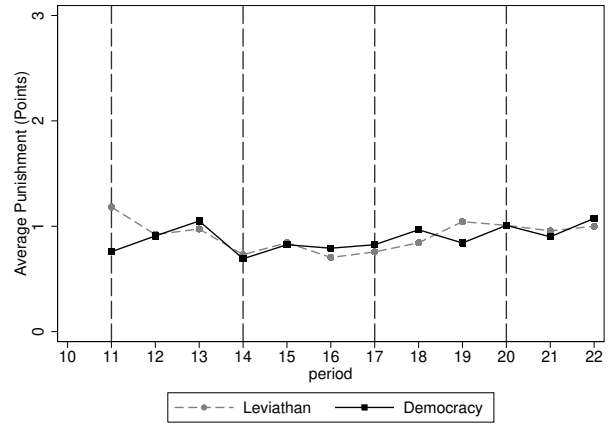


(b) Multiple elections

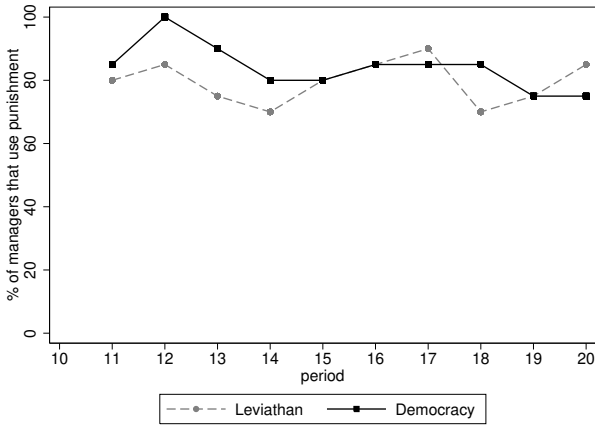
Figure 1: Average contribution dynamics



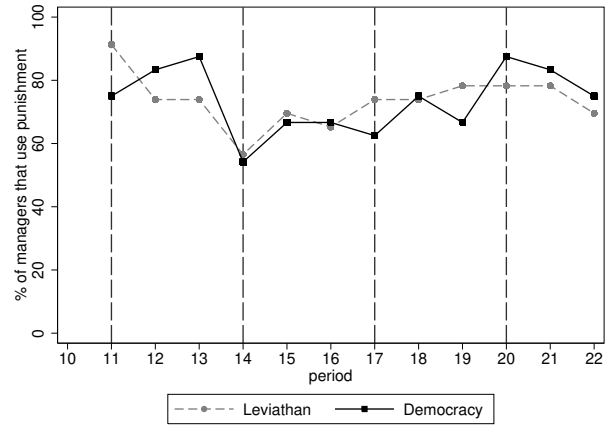
(a) Single election: punishment assigned



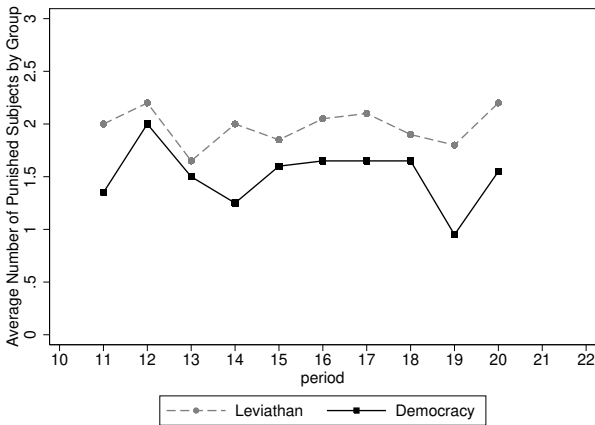
(b) Multiple elections: punishment assigned



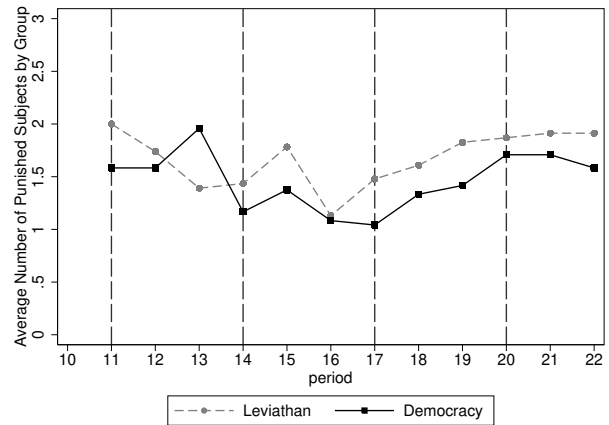
(c) Single election: manager's use of punishment



(d) Multiple elections: manager's use of punishment

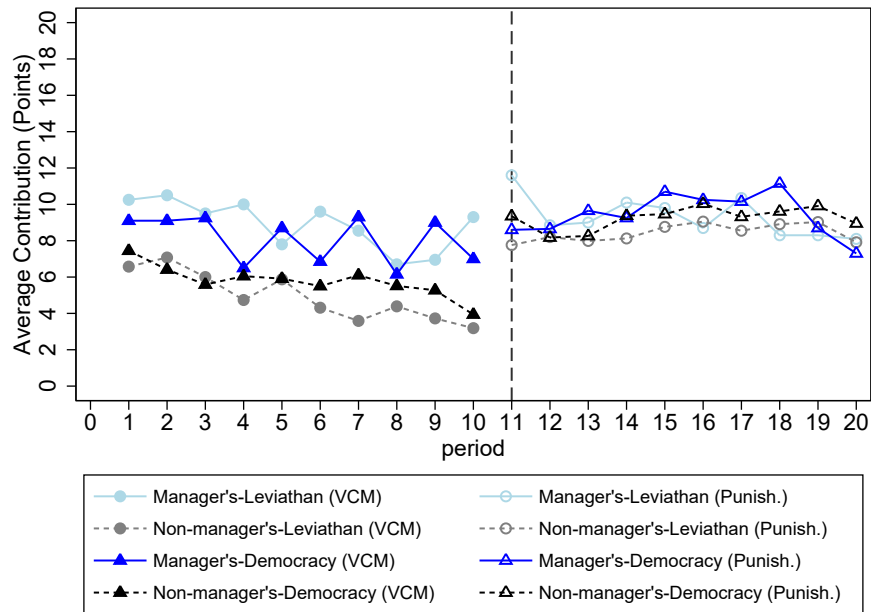


(e) Single election: subjects punished

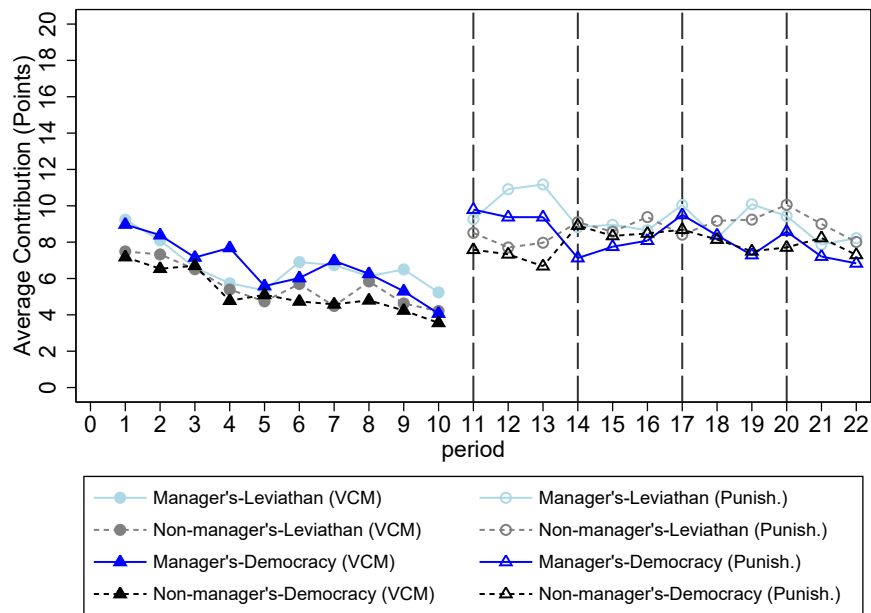


(f) Multiple elections: subjects punished

Figure 2: Punishment Behavior

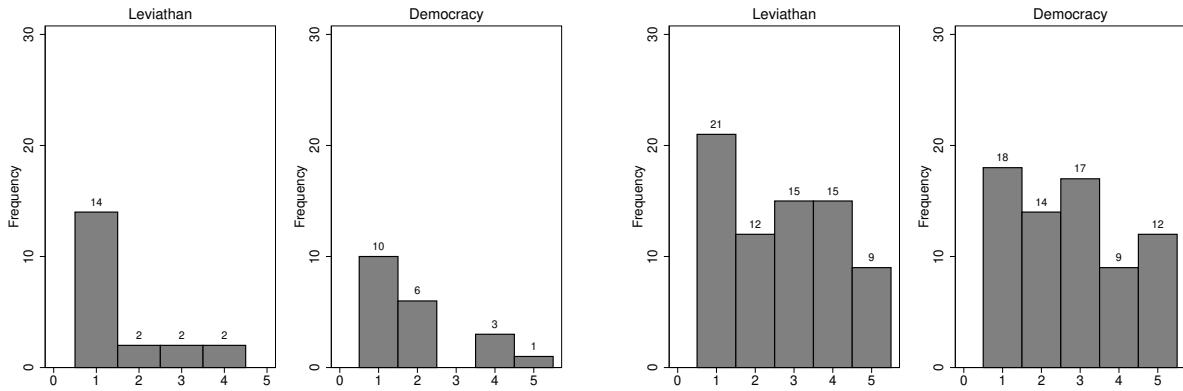


(a) Single election



(b) Multiple elections

Figure 3: Contribution's dynamics by roles



(a) Single election

(b) Multiple elections

Figure 4: Manager's contribution rank within group



# Appendix A Experimental instructions

## A.1 General instructions

Please read them carefully.

You should answer some questions to verify that you understand the instructions.

If you have any questions, raise your hand and the assistant will come to answer it.

During the session you may have to wait for a moment before the experiment continues. We thank you for your patience and cooperation. The earnings of this section will be calculated in points. The points will be converted into money with the following exchange rate:

$$1 \text{ point} = \text{US\$ } 0.02$$

This section has two phases. The following instructions explain the details of phase 1.

The instructions of phase 2 will be delivered later.

## A.2 Instructions Phase 1

Read these instructions carefully.

In this phase, all participants are randomly divided into groups of 5. That is, you will be part of a group with 4 other participants.

The members of the group will be the same throughout phase 1. No one knows the composition of the group, and no one will be informed of who participated in which group after the experiment.

### **YOUR DECISIONS**

Phase 1 includes 10 rounds. You will begin each round with 20 points. In each round you and the other 4 members of your group decide simultaneously how to assign your 20 points in two accounts:

- Group Account
- Private Account

Specifically, you will decide how many points to assign to the group account, using only integer numbers from 0 to 20. The remaining points are automatically assigned to your private account.

### **YOUR PROFITS**

Your earnings depend on the total number of points in the group account and the number of points in your private account.

Your total earnings in each round can be calculated with the following formula:

$$20 - (\text{points you assigned to the group account}) + 0.4 * (\text{sum of points assigned by all members of the group to the group account})$$

For each point assigned to the private account you get 1 point as profit. For example, if you assign 10 points to this account, your private account gain is 10 points.

Your group account earnings are equal to the total amount of points assigned to the account by all members, multiplied by 0.4.

For each point you assign to the group account, everyone in your group (including you) gets 0.4 points as profit. For example, if each member of the group, including you, allocates 10 points to the group account, the sum of points in the group account is 50 and therefore your earnings from this account are 20 points ( $20 = 0.4 * 50$ ).

Keep in mind that you receive 1 point for each point you assign to your private account. On the contrary, you, as well as the other group members, receive 0.4 points for each point assigned to the group account. Another way to interpret this is that the aggregate points of the group account are multiplied by two and divided equally for each member of the group.

Also note that you earn profits from points assigned by other members to the group account.

#### **Example**

Suppose you assign 10 points to the group account, the second and third members of your group assign 20 points to the group account, and the remaining two members assign 0 points to the group account.

Then the sum of the points to the group account is:  $10 + 20 + 20 + 0 + 0 = 50$  points.

Each member of the group receives profits of:  $0.4 * 50 = 20$  points of the group account.

Your total earnings are:  $20 - 10 + (0.4 * 50) = 10 + 20 = 30$  points.

The earnings of the second and third members of the group are:  $20 - 20 + (0.4 * 50) = 0 + 20 = 20$  points.

The gains of the fourth and fifth members of the group are:  $20 - 0 + (0.4 * 50) = 20 + 20 = 40$  points.

Please raise your hand if you have any questions, if not, click Next to continue.

### **FEEDBACK**

At the end of each period you will be informed about the number of points you have allocated to your private account and group account. You will also learn the total amount allocated to the group account in your group. Below this, you will see your total earnings, in points, for the round. You will also be shown the decisions by you and the decisions of all the other members of your group anonymously (letters are randomly assigned), from the current and previous periods.

### **A.3 Instructions Phase 2 (“Leviathan”, name not show on the instructions)**

Read these instructions carefully.

This phase includes 12 rounds in total. Similar to phase 1, you continue interacting with the same 4 participants in your group and in each round you make a decision about the allocation of 20 points to a group account. The remaining points are automatically assigned to the private account.

Your earnings are determined in the same way as in phase 1 of the experiment, that is, for each point assigned to your private account you get 1 point as profit. Your group account earnings are equal to the sum of all the points assigned by the 5 members, multiplied by 0.4.

However, in this phase, before beginning to make your decisions, ONE participant of your group will be chosen as administrator. An administrator will be elected every 3 periods, that is, during the second phase there will be several instances of the administrator’s selection, according

to the instructions described below.

### **HOW IS THE ADMINISTRATOR SELECTED?**

Before starting phase 2, you will be able to observe the allocation decisions in each round, as well as the average contribution, of each group member in phase 1.

The computer will choose an administrator based on a distribution that assigns a 75% probability that the group's largest contributor will be chosen. If the highest phase 1 contributor is not chosen (this happens with 25% chance), there is an equal chance of any other group member being chosen to play the role of administrator.

You will see who was chosen before starting phase 2. Note that the administrator will be chosen every three rounds.

The election in the later rounds will be made based on the average contribution of phase 1, and also the information of the contributions of the 3 intermediate rounds. Specifically, in rounds 4, 7 and 10, you will see the average of phase 1 contributions to the group account and the total contribution of all group members in the 3 preceding rounds; the administrator will be selected based on this information.

### **YOUR DECISIONS IN THIS PHASE**

There are two stages in each round. In the first stage, you make your decision to assign points to the group account and then observe the decisions of the other members of the group, arranged in a random manner, along with your earnings.

In the second stage, 2 points are automatically deducted from each member of the group for the administrator's account. Therefore, each administrator has 10 points in each round. The administrator has the opportunity to use the points in this account to reduce the earnings of the participants that they choose within the group, except for themselves. The administrator can not reduce their own profits.

Suppose you are selected as an administrator. After the first stage of each round, you will observe the amount assigned to the group account by each member. At the same time, you will receive 2 points from each member of the group including you.

You will then choose the members of the group (one or several) and how many of the 10 available points you want to use to reduce the earnings of the selected members. For each point you wish to assign, the earnings of the selected participants are reduced according to the following table:

Points ASSIGNED	0	1	2	3	4	5	6	7	8	9	10
Points REDUCED	0	1	2	4	6	9	12	16	20	25	30

The administrator is free to leave the gains of all the members of the group intact, without changes, assigning zero (0) points or leaving empty the corresponding cell.

After the administrator makes their decisions, the remaining points of the administrator's account, which are not allocated to reduce the profits of others, are returned to all the members of the group equally. The points assigned by the administrator to any group member are deducted from the account and therefore, are not recoverable by the group members.

The profits in each round correspond to the result of the administrator's decisions and the decisions of each group member; however, if these are negative, you will assume the cost of the points assigned to you by the administrator (not their equivalent in reduced points) or the cost of your negative gain (loss); whichever is less in absolute terms.

Therefore, the gains in each round are calculated as follows:

**Total profits**

**If earnings  $\geq 0$ :**

= 20 - (points assigned to the group account) + 0.4 \* (sum of points assigned by all members of the group to the group account) - 2 - (reduced points in your earnings for the assigned points of the administrator, if selected) + (1/5) \* (remaining points not used by the administrator)

**If earnings  $< 0$ :**

= lower value between: points assigned to you by the administrator (if selected), or negative gain (loss)

Note that your earnings in this phase can be negative if the points reduced by the administrator are greater than your earnings in the first stage.

You can, however, avoid these losses through your decisions!

### Example

Suppose you are the administrator of your group. In a round, you assign 20 points to the group's account, the second and third members of your group assign 10 points to the group's account, and the remaining members assign 5 points each to the group's account.

In the first stage, the gains are obtained in the following way:

The sum of the points to the group's account is:  $20 + 10 + 10 + 5 + 5 = 50$  points.

Each member of the group receives earnings of:  $0.4 * 50 = 20$  points, from the group account.

Your gross earnings are:  $20 - 20 + (0.4 * 50) = 0 + 20 = 20$  points.

The earnings of the second and third members of the group are:  $20 - 10 + (0.4 * 50) = 10 + 20 = 30$  points.

The gains of the fourth and fifth members of the group are:  $20 - 5 + (0.4 * 50) = 15 + 20 = 35$  points.

In the second stage, you receive as administrator 2 points from each member of the group, 10 points in total.

You decide to assign 5 points to reduce the profits, only from the fifth member of the group.

The remaining 5 points of the account will be returned to all the members of the group, that is, 1 point for each one.

Your net earnings are:  $20(\text{from the first stage}) - 2 + 1 = 19$  points.

The net earnings of the second and third members are:  $30 - 2 + 1 = 29$  points.

The net earnings of the fourth member are:  $35 - 2 + 1 = 34$  points.

The net earnings of the fifth member of the group are:  $35 - 2 - 9 + 1 = 25$  points.

Note that the 5 points assigned by the administrator reduce by 9 points the earnings of the selected participant, according to the conversion table indicated in the instructions.

Please raise your hand if you have any questions, if not, click Next to continue.

## **A.4 Instructions Phase 2 (“Democracy”, name not show on the instructions)**

Read these instructions carefully.

This phase includes 12 rounds in total. Similar to phase 1, you continue interacting with the same 4 participants in your group and in each round you make a decision about the allocation of 20 points to a group account. The remaining points are automatically assigned to the private account.

Your earnings are determined in the same way as in phase 1 of the experiment, that is, for each point assigned to your private account you get 1 point as profit. Your group account earnings are equal to the sum of all the points assigned by the 5 members, multiplied by 0.4.

However, in this phase, before beginning to make your decisions, ONE participant of your group will be chosen as administrator. An administrator will be elected every 3 periods, that is, during the second phase there will be several instances of the administrator’s selection, according to the instructions described below.

### **HOW IS THE ADMINISTRATOR SELECTED?**

Before starting phase 2, you will be able to observe the allocation decisions, in each round, of each group member in phase 1.

Then you will have the opportunity to vote for any member of your group (except you) to be elected as administrator for the next three rounds. Whomever receives the most votes will be elected administrator. In case of a tie, the administrator will be selected at random from the tied members.

You will see who was elected as administrator before starting phase 2. Note that the administrator will be chosen every three rounds by vote, just as in round 1.

Before voting in later rounds, each group member will again see the average contributions from phase 1, and the information of the 3 intermediate rounds. Specifically, in rounds 4, 7 and 10, you will see the average of phase 1 contributions to the group account and the total contribution of all group members in the 3 preceding rounds. Thus, you will be able to vote for the manager of your preference.

### **YOUR DECISIONS IN THIS PHASE**

There are two stages in each round. In the first stage, you make your decision to assign points to the group account and then observe the decisions of the other members of the group, arranged in a random manner, along with your earnings.

In the second stage, 2 points are automatically deducted from each member of the group for the administrator's account. Therefore, each administrator has 10 points in each round.

The administrator has the opportunity to use the points in this account to reduce the earnings of the participants that they choose within the group, except for themselves. The administrator can not reduce their own profits.

Suppose you are selected as an administrator. After the first stage of each round, you will observe the amount assigned to the group account by each member. At the same time, you will receive 2 points from each member of the group including you.

You will then choose the members of the group (one or several) and how many of the 10 available points you want to use to reduce the earnings of the selected members.

For each point you wish to assign, the earnings of the selected participants are reduced according to the following table:

Points ASSIGNED	0	1	2	3	4	5	6	7	8	9	10
Points REDUCED	0	1	2	4	6	9	12	16	20	25	30

The administrator is free to leave the gains of all the members of the group intact, without changes, assigning zero (0) points or leaving empty the corresponding cell.

After the administrator makes their decisions, the remaining points of the administrator's account, which are not allocated to reduce the profits of others, are returned to all the members of the group equally. The points assigned by the administrator to any group member are deducted from the account and therefore, are not recoverable by the group members.

The profits in each round correspond to the result of the administrator's decisions and the decisions of each group member; however, if these are negative, you will assume the cost of the points assigned to you by the administrator (not their equivalent in reduced points) or the cost of your negative gain (loss); whichever is less in absolute terms.



Therefore, the gains in each round are calculated as follows:

**Total profits**

**If earnings  $\geq 0$ :**

= 20 - (points assigned to the group account) + 0.4 \* (sum of points assigned by all members of the group to the group account) - 2 - (reduced points in your earnings for the assigned points of the administrator, if selected) + (1/5) \* (remaining points not used by the administrator)

**If earnings  $< 0$ :**

= lower value between: points assigned to you by the administrator (if selected), or negative gain (loss)

Note that your earnings in this phase can be negative if the points reduced by the administrator are greater than your earnings in the first stage.

You can, however, avoid these losses through your decisions!

**\*THE REST: EXAMPLE, TEST AND QUESTIONNAIRES, ARE SIMILAR TO BOTH TREATMENTS.**

## Appendix B Additional tables and figures

Table 6: Profits: regression results

Dep. Variable=Profits (points)	Single election		Multiple elections	
	FEgt	FEgtc	FEgt	FEgtc
Democracy vs. Leviathan (P*D)	0.2581 (0.9280)	0.1371 (0.6650)	-0.7698 (0.7968)	-0.5026 (0.5967)
Punishment	0.5749 (0.6496)	0.3035 (0.4205)	1.9827*** (0.7317)	1.4269** (0.6154)
Other member's Av. Contribution (t-1)		0.4340*** (0.0583)		0.3821*** (0.0507)
Punishment Received (t-1)		-0.7982 (0.5322)		-0.3644 (0.5786)
Punishment in the group (t-1)		-0.3576** (0.1516)		-0.4183** (0.1626)
Other controls	No	No	No	No
Individual FE	Yes	Yes	Yes	Yes
Group FE	Yes	Yes	Yes	Yes
Trend within phase	Yes	Yes	Yes	Yes
R.squared	0.3184	0.3791	0.2587	0.2846
Observations	4000	3800	5170	4935

Notes: Dummy variable for the Democracy treatment (D) excluded since it is time invariant; hence it has a null coefficient for a Fixed Effect (FE) estimation.

Standard errors clustered at group level in parentheses.

\* Significant at the 10 percent level.

\*\* Significant at the 5 percent level.

\*\*\* Significant at the 1 percent level.

Table 7: Determinants of contributions conditional on manager's contribution rank

Dep. Variable=Contributions (points)	Single election		Multiple election	
	FEgt	FEgtc	FEgt	FEgtc
Democracy vs Leviathan (MR=1)	0.8025 (1.2698)	0.6179 (0.8296)	-2.0502*** (0.7454)	-2.0678*** (0.7179)
Democracy vs Leviathan (MR=2)	0.1458 (1.0894)	0.2301 (0.7190)	-1.3687 (0.8309)	-1.3388* (0.7066)
Democracy vs Leviathan (MR=3)			-0.7094 (0.6707)	-0.2789 (0.5395)
Democracy vs Leviathan (MR=4)	-1.1609 (1.2462)	-0.9963 (0.6657)	0.8803 (1.0414)	1.1309 (0.8631)
Democracy vs Leviathan (MR=5)	-1.8475*** (0.6082)	-1.1788*** (0.3879)	-0.2437 (0.8745)	0.3809 (0.6699)
Punishment (P)		2.0573*** (0.3936)	3.9136*** (0.5606)	3.1281*** (0.4518)
Other member's Av.Contribution(t-1)		0.3896*** (0.0496)		0.3623*** (0.0407)
Punishment received (t-1)		0.4501 (0.4228)		0.1312 (0.3306)
Punishment in the group (t-1)		-0.1571 (0.1289)		-0.0960 (0.1121)
Other controls	No	No	No	No
Individual FE	Yes	Yes	Yes	Yes
Group FE	Yes	Yes	Yes	Yes
Trend within phase	Yes	Yes	Yes	Yes
R.squared	0.4255	0.4753	0.4391	0.4764
Observations	4000	3800	5170	4935

Notes: Dummy variable for the Democracy treatment (D) excluded since it is time invariant; hence it has a null coefficient for a Fixed Effect (FE) estimation.

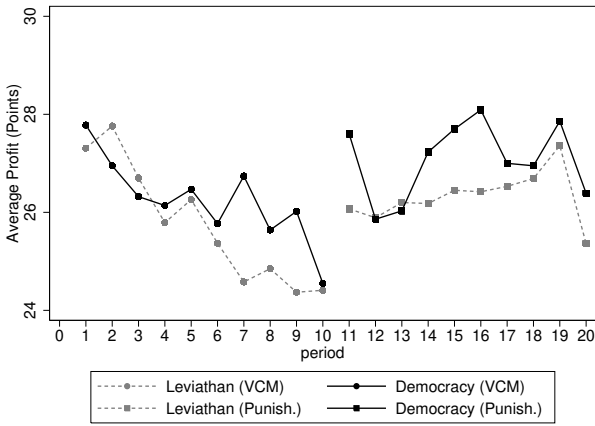
MR: Manager's rank of cumulative contribution in the first phase.

Standard errors clustered at group level in parentheses.

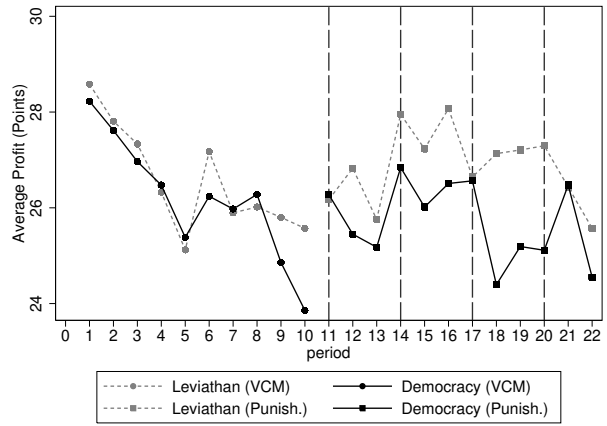
\* Significant at the 10 percent level.

\*\* Significant at the 5 percent level.

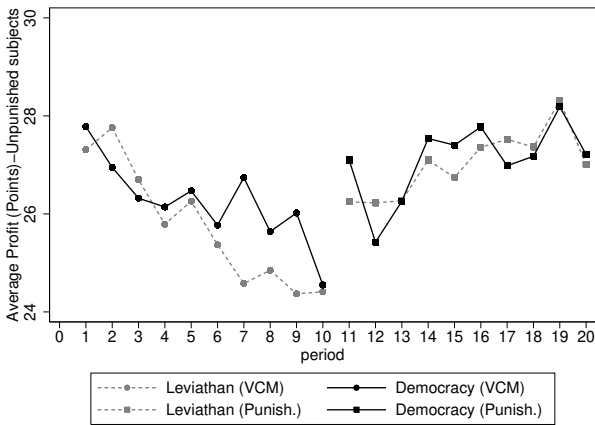
\*\*\* Significant at the 1 percent level.



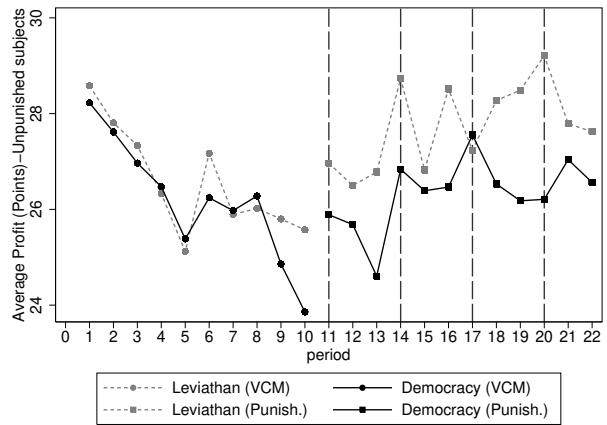
(a) Single election



(b) Multiple elections



(c) Single election: Non-free-riders



(d) Multiple elections: Non-free-riders

Figure 5: Profit's dynamics