Decentralized or Centralized Governance in Social Dilemmas? Experimental Evidence from Georgia

Mekvabishvili, Rati

10 June 2023

Decentralized or Centralized Governance in Social dilemmas? Experimental Evidence from Georgia

Rati Mekvabishvili[1]

Published in: Issues in Social Science Vol. 11, No. 1, June 30, 2023

Abstract

The vast majority of experimental studies on the effectiveness of punishments in promoting cooperation in social dilemma situation examine decentralized incentive systems where all group members can punish each other. Cross-societal experimental studies suggest that while decentralized incentives can successfully promote cooperation in one society, they fail to do so in another. So, how is social order, as a large-scale cooperation problem among strangers, maintained in such societies? Many modern societies overcome this problem through well-functioning top-down formal enforcement institutions. In the experimental setting of the public goods game, we compare a strong and weak exogenous centralized incentive system with a decentralized incentive system in the case of Georgia. Our experimental evidence suggests that in Georgia, self-governed groups are doomed to suffer from high inefficiencies under a decentralized peer-to-peer punishment incentive system. They are better off when punishment power is given to an external centralized authority that is not exposed to power abuse risks.

Keywords: centralized punishment, decentralized punishment, cooperation, public goods, welfare
JEL: C72, C92, H41

[1] Rati Mekvabishvili, Department of Theoretical Economics, Ivane Javakhishvili Tbilisi State University, Faculty of Economic and Business, address: University st. N2, 0186 Tbilisi, Georgia, E-mail: rati.mekvabishvili@eab.tsu.edu.ge.
I am deeply grateful to Prof. Simon Gächter and Prof. Christian Thöni for providing experimental data of public goods game experiment conducted in Georgia in year 2013, and Prof. Chris Starmer for facilitating this process.

1. Introduction

Self-governance is often necessary to overcome many important social dilemmas and maintain cooperation. One aspect of self-governance is informal punishment, also known as peer punishment, which operates within a decentralized system. The majority of theoretical and experimental studies on the enforcement of cooperation in public goods games have focused on peer punishment, where all group members have the ability to punish each other. A groundbreaking experimental study by Fehr and Gächter (2000) demonstrated that people are willing to punish free-riders and that peer punishment increases cooperation (Fehr and Gächter, 2000a). They also revealed that if punishment is not possible, cooperation breaks down. Another experimental study revealed that the majority of individuals prefer an environment where punishment exists (Gürerk, Irlenbusch, and Rockenbach, 2006). While the experimental literature suggests that peer punishment can effectively promote cooperation in social dilemma situations, there are instances where it fails to do so. One problematic aspect of peer punishment is the potential for some players to misuse the power of sanctioning incentives and undermine cooperation. For example, several public goods game experiments with peer punishment have documented the existence of "antisocial" punishment, where sanctions are imposed on cooperators rather than free-riders (Herrmann, Thöni, and Gächter, 2008; Nikiforakis, 2008). A seminal cross-cultural experimental study conducted by Herrmann et al. (2008) demonstrated abundant evidence of cultural differences when it comes to peer punishment (hereafter referred to as HT&G).

Mancur Olson, in a groundbreaking analysis of the free-rider problem in collective action, argued that only a "selective" incentive can motivate a rational individual to act in accordance with the group's interests (Olson, 1975). One such selective incentive in modern societies is the legal system, including courts and police, which administer punishment. However, there have been relatively few experimental studies on centralized punishment and its effectiveness in promoting cooperation. Generally, centralized punishment implies that punishment is carried out either internally by a selected central monitoring entity within the group or by an externally imposed entity that follows predefined rules.

Several experimental studies have examined endogenous centralized punishment, where one group member acts as the monitoring entity, and found it to be quite effective (Baldassarri and Grossman, 2011; O'Gorman, Henrich, and Van Vugt, 2009). However, assigning punishment power to a single member of the group has revealed cases where individuals abuse their power (Nosenzo and Sefton, 2012; Carpenter, Kariv, and Schotter, 2012). The negative impacts of this misuse of punishment power can be mitigated by introducing a rotating basis system within the group, leading to high levels of cooperation and earnings (O'Gorman, Henrich, and Van Vugt, 2009).

An experimental study by Stagnaro et al. (2017) observed a positive impact of different strengths of exogenous centralized punishment mechanisms in sustaining a high level of cooperation in public goods games (PGG). Qin and Wang (2013) also found that an exogenous centralized incentive improves cooperation in the PGG experiment (Qin and Wang, 2013). Andreoni and Gee (2012) illustrated that, under reasonable conditions, individuals prefer to be governed by a delegated punishment mechanism rather than relying on peer-to-peer punishments (Andreoni and Gee, 2012). A recent experimental study by Engel et al. (2021) also demonstrated a positive effect of external centralized punishment mechanisms (Engel, Riedl, and Weber, 2021). However, it is difficult to conclude the relative effectiveness of centralized incentive systems compared to decentralized incentive systems, particularly considering the influence of cross-societal variations.

In our paper, we aim to examine the performance of a decentralized incentive system in a social dilemma situation in the case of Georgia, comparing it to the participant pools of 16 different countries studied by HT&G. Additionally, within the same public goods game (PGG) domain setting, we compare the impact of the decentralized incentive system on cooperation and

welfare in Georgia with the exogenous centralized incentive system of varying strength studied recently by Mekvabishvili (forthcoming). The study reveals that even a low level of exogenous top-down incentive can ensure a relatively high level of cooperation and welfare. The paper has two objectives: first, to assess the effectiveness of the decentralized incentive system in a social dilemma situation in case of Georgia, and second, to compare the performance of the decentralized incentive system with the exogenous centralized incentive system in the same environment of Georgian society.

The remainder of the paper is organized as follows: Section 2 presents the experimental design; Section 3 provides the experimental results; Section 4 concludes.

2. Experimental design

2.1 Method

In our comparative experiment, we utilized the PGG data of Mekvabishvili (2021) and an experiment conducted in Georgia by research team of Prof. Simon Gächter in 2013.[2] The study by Mekvabishvili (forthcoming) examined rule-based exogenous centralized punishment with a high and low probability of inspection and sanctioning in the PGG, which was conducted in Georgia. The experiment conducted in 2013, on the other hand, examined peer punishment in the same experimental setting of PGG domain. Therefore, both experiments were conducted using the same PGG experimental design settings, allowing us to perform a comparability analysis. Furthermore, given the identical experimental design and setting, we compare the results of the PGG with peer punishment conducted in Georgia with the experimental findings from the seminal cross-cultural study by HT&G.

2.2 A public goods game with decentralized punishment

The laboratory experiment was conducted in Georgia using Z-Tree software (Fischbacher, 2007). A total of 100 subjects participated from Tbilisi State University. The participants engaged in a two-stage public goods game (PGG), consisting of 10 periods each. In the first stage, participants played a standard linear PGG based on the voluntary contributions mechanism (VCM). Each participant was grouped with three other players, and simultaneously decided how much of their 20-point endowment to keep or invest into the group account, denoted as g_i, in each period, where $0 \leq g_i \leq 20$. The rest (20 - g_i) of the endowment points were kept by the player. In addition to the points that player i kept, hey received a fixed return equal to 40% of the group's total contribution to the group account. The monetary payoff for each subject i in the group is given by:

$$(1)\ \pi_i^1 = 20 - g_i + 0.4 * \sum_{j=1}^{4} g_i$$

Once the first stage was completed, participants received new instructions for the second stage. The game began only when all group members had completed the control questions. The groups remained constant throughout the experiment (Partner protocol). In the second stage, participants were provided with information about the contributions made by each of their group

[2] The author assisted the research team of Prof. Simon Gächter in conducting the lab experiment in Georgia at that time. The experimental data of Georgia lab experiment of year 2013 has never published before.

members. In each period, they were given the opportunity to simultaneously punish each other. To administer punishment, group member i had to assign punishment points to group member j, p_{ij}, i ≠ j, where $0 \leq p_{ij} \leq 10$. Thus, they were allowed to assign up to 10 punishment points to each of their peers. The allocation of a punishment point p_{ij} by player i to player j reduces the payoff of player i by one point and that of player j by 3 points. If player i receives p_{ji} punishment points from the other group members and assigns p_{ij} punishment points to member j, the final payoff of subject i, π_i, is:

$$(2)\ \pi_i = \pi_i^1 - (3 * \sum_{j=1}^{4} p_{ji} + \sum_{j=1}^{4} p_{ij})$$

At the end of each period, subjects were reminded of their income, the total punishment points they assigned, the associated cost, and the punishment points they received from the group. The experiment session lasted approximately 40-50 minutes. Participants were paid in cash after the experiment session and earned an average of 11.4 GEL (equivalent to 6.7 USD at that time).

2.3 A public goods game with centralized punishment

The experiment by Mekvabishvili (forthcoming) was conducted in Georgia using the LIONESS software platform for interactive online experiments (Arechar, Gächter, and Molleman, 2018). A total of 121 students from various majors at Tbilisi State University participated in the treatment involving centralized punishment. The experiment maintained constant group members throughout (Partner protocol). Participants engaged in a two-stage public goods game (PGG) with ten periods each. In the first stage, participants played a standard linear PGG with an exogenously imposed centralized inspection and punishment probability mechanism. In the second stage, the centralized punishment mechanism was removed, and participants played a standard linear public goods game. In the first stage, the payoff was determined by the following equation:

$$(3)\ \ \pi_i^1 - 2 * (20 - g_i) * P(A|B) * P(B)$$

where $\pi_i^1 = 20 - g_i + 0.375 \sum_{j=1}^{4} g_i$, *P(A)* was the probability that a penalty would be imposed, given the probability *P(B)* that the contribution would be inspected. Mekvabishvili (forthcoming) conducted two treatments in which the probability of punishing free riders was varied. In the first treatment, the probability of punishment was high at 90% (referred to as strong centralized punishment or SCP), while in the second treatment, it was low at 10% (referred to as weak centralized punishment or WCP). Therefore, the levels of probability in the centralized punishment mechanisms represented the strength of formal enforcement institutions. The penalty imposed on free riders was twice the amount the player kept for themselves. In both treatments, participants had to answer control questions to ensure comprehension of the game; otherwise, they could not proceed. The SCP treatment lasted 30 to 40 minutes, and participants earned an average of 20.7 GEL (approximately 6.3 USD at that time), while the WCP treatment also lasted 30 to 40 minutes, and participants earned an average of 21.2 GEL (approximately 6.6 USD at that time). Table 1 provides a summary of the experimental design information for both experiments.

Table 1: Experimental design information

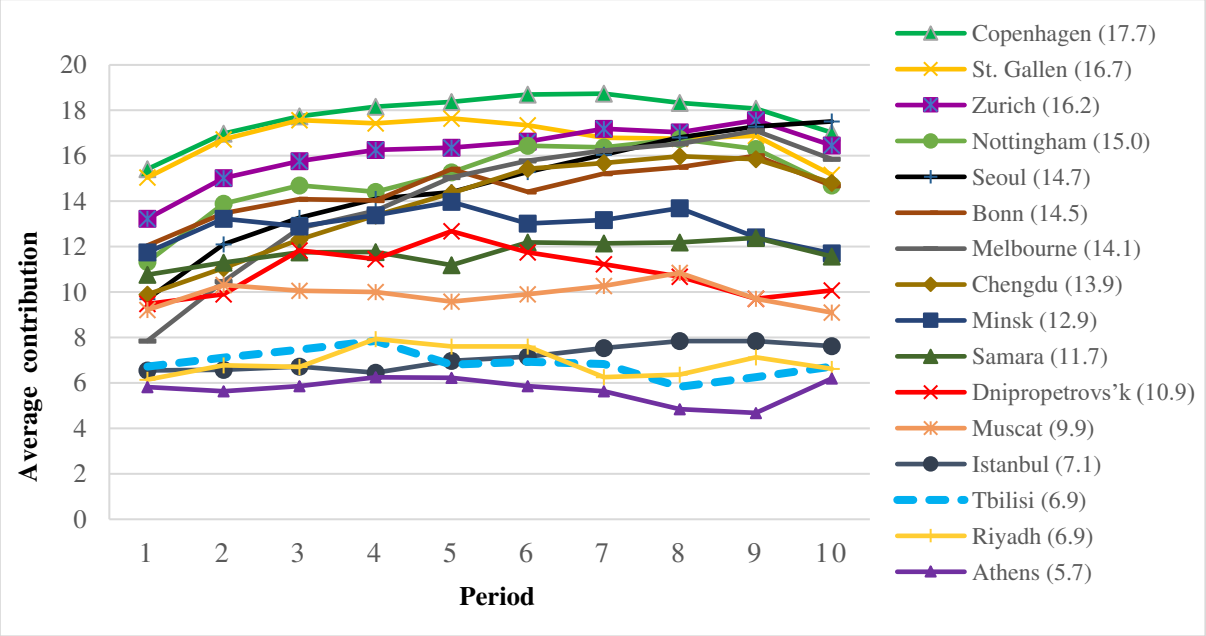
Experimental game	Experiment conduct method	Matching protocol	Order of treatment		Number of group members	Number of subjects
			Stage 1 (periods 1-10)	Stage 2 (periods 11-20)		
PGG with decentralized peer punishment	laboratory	Partner	VCM	DPP	4	100
PGG with strong centralized punishment	online	Partner	SCP	VCM	4	64
PGG with weak centralized punishment	online	Partner	WCP	VCM	4	65

3. Results

3.1 Decentralized peer-to-peer punishment treatment

We conducted a comparison of cooperation levels in the PGG experiment with the decentralized peer punishment treatment (DPP) and the cross-cultural experiment conducted by HT&G. Figure 1 illustrates the comparison of average contributions under peer punishment conditions across a pool of 16 different countries. As shown in Figure 1, cooperation levels varied significantly among the participant pools. In our experiment conducted in Tbilisi, the cooperation level (indicated by the blue dotted line) was one of the lowest, with an average contribution of 34.3% of the endowment (6.9 points). We observed that peer punishment was highly ineffective, as cooperation remained consistently low throughout all ten periods.

Figure 1: Average contributions under peer punishment condition: comparison to the cross-cultural study



Source: Mekvabishvili, R. (2021). "Centralized Punishment in Public Good Experiments". Dataset, Zenodo, DOI: https://doi.org/10.5281/zenodo.5033369 and HT&G experiment data on https://datadryad.org/stash/dataset/doi:10.5061/dryad.87301

Result 1: *In the decentralized peer-to-peer punishment (DPP) treatment, the average contributions remained consistently low over time, indicating a lack of cooperation. This finding placed our participant pool among the least cooperative groups in the study.*

In the DPP condition, participants were frequently subjected to punishment, with 55.1% of opportunities resulting in punishment. On average, a punishment of 2.76 points out of ten was assigned. Following the approach of HT&G, we define cooperative behavior as a subject's contribution equal to or above the average contribution of the group, and non-cooperative behavior as a contribution below the group average. Prosocial punishment refers to punishment directed towards non-cooperative individuals, while antisocial punishment refers to punishment targeted at cooperative individuals. Our findings indicate that non-cooperative behavior was punished in 39.2% of cases, while even cooperative behavior received significant punishment in 30.7% of cases. The average punishment points received for non-cooperative and cooperative behavior were 3.93 and 3.67, respectively. Thus, our results provide considerable evidence of both antisocial punishment and its relatively harsh nature.

We examined participant behavior within the group over the course of all ten periods. In many cases, subjects exhibited significant variation in their cooperative and punishment behaviors, frequently transitioning from cooperative to non-cooperative behavior and assigning punishments to both non-cooperators and cooperators. Overall, individuals who behaved cooperatively were punished in 51.2% of opportunities, while those who behaved non-cooperatively were punished in 48.8% of opportunities. The average punishment points assigned were 2.91 and 2.59, respectively. The presence of punishment did not exhibit a strong disciplinary effect on non-cooperative behavior in terms of encouraging low contributors to increase their contributions. Instead, we observed a "battle" scenario where group members responded to punishment with punishment, assigning deduction points to each other.

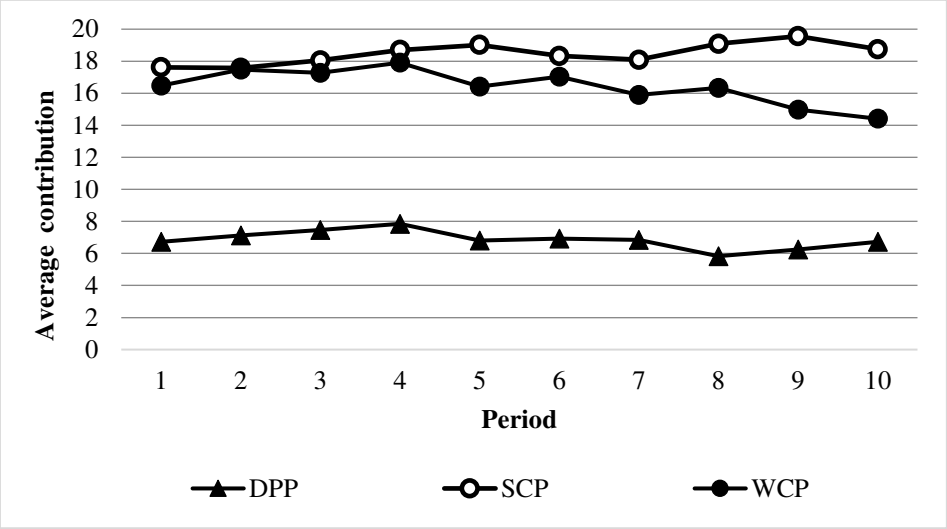
Result 2: *In the decentralized peer-to-peer punishment (DPP) treatment, we found significant evidence of antisocial punishment. Additionally, the punishment did not demonstrate a strong disciplinary effect on non-cooperative behavior.*

3.2 Strong and weak centralized punishment

To compare cooperation and welfare levels between decentralized peer punishment and centralized punishment, we analyzed data from periods 1 to 10 of the first stages of the SCP and WCP treatments, as well as data from periods 11 to 20 of the second stage of the DPP treatment. Figure 2 depicts the average contributions in both SCP and WCP treatments, starting at around 82% and 88% of endowments, and remaining consistently high across all periods (92% and 82%, respectively). However, in the WCP treatment, average contributions exhibit a declining pattern from period 4. The difference between these treatments is statistically significant (Mann-Whitney test, two-sided, $p = 0.0000$).

On the other hand, average contributions in the DPP treatment start at around 34% of the endowment and remain within this low range throughout the game. These contributions are significantly different from the average contributions in the WCP and SCP treatments (Mann-Whitney test, two-sided, $p = 0.000$ and $p = 0.000$, respectively). This evidence suggests that punishment is more effective, even under the condition of a weak centralized punishment system, than under the condition of a decentralized punishment system.

Figure 2: Average contribution



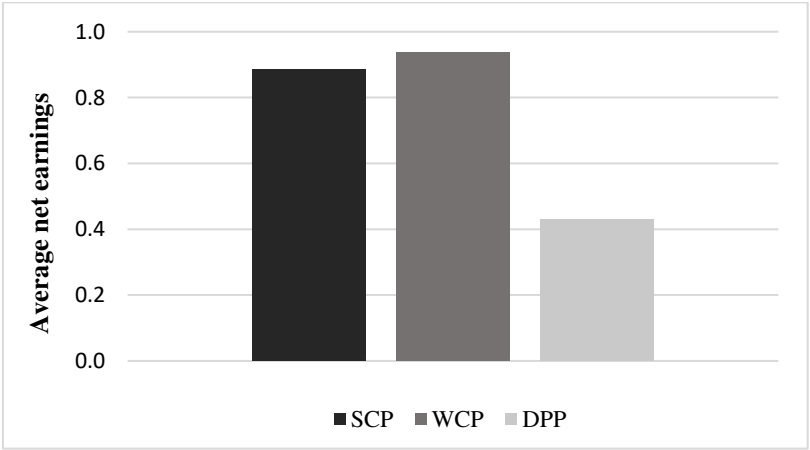
Result 3: *The punishment in the decentralized peer punishment (DPP) treatment has a weaker disciplinary effect on the average level of cooperation compared to the centralized punishment treatments.*

3.3 Welfare

Next, we investigate whether the significant differences in cooperation levels between treatments with centralized punishment and decentralized punishment also lead to substantial disparities in welfare. Welfare is measured by the average individual net earnings. In the case of the DPP treatment, net earnings are calculated after deducting the costs of executed punishment and received punishment, while in the SCP and WCP treatments, net earnings are calculated after deducting received punishment only.

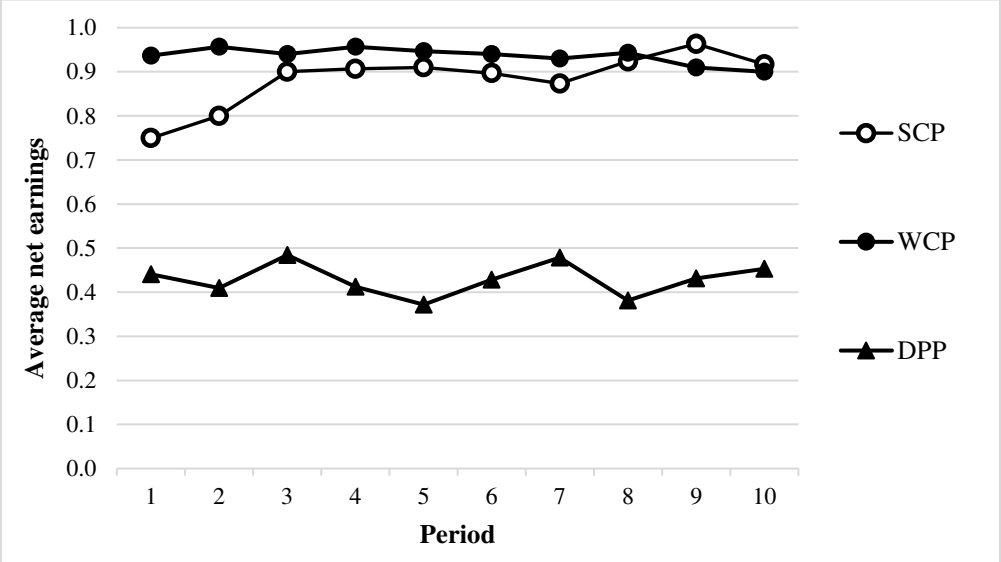
In all three treatments, the group welfare-maximizing level of contribution is when all four members of the group make a full contribution of 20 points. If all players choose this amount, their earnings would be 30 points in the SCP and WCP treatments, and 32 points in the DPP treatment. Figure 3 illustrates the net earnings (in experimental money units) averaged over all 10 rounds in each treatment as a percentage of the potential welfare-maximizing level. The figure clearly demonstrates that in periods 1 to 10, average net earnings are more than twice as high in the SCP and WCP treatments compared to periods 11-20 of the DPP treatment.

Figure 3: Average net earnings



Next, in Figure 4, we provide further insights by presenting the average individual net earnings over time (in experimental money units). The net earnings in the centralized punishment treatments exhibit a high and consistent level throughout all periods. However, in the initial periods, the earning levels in the SCP treatment are lower than those in the WCP treatment, and this difference is statistically significant (Mann-Whitney test, two-sided, p=0.0126). This evidence suggests that welfare can still be maintained at relatively high levels even with a low probability of sanctioning free riders.

Figure 4: Average individual net earnings over time across treatments



Relative to the centralized punishment treatments (SCP and WCP), the average individual net earnings are significantly lower in the DPP treatment (Mann-Whitney test, two-sided, p = 0.000 and p = 0.000, respectively). This finding suggests that the centralized punishment mechanism results in greater improvements in welfare compared to a decentralized punishment system.

Result 4: *Welfare, as measured by average individual net earnings, is higher in the SCP and WCP treatments compared to the DPP treatment. The exogenous centralized punishing mechanism leads to an improvement in welfare compared to the decentralized peer punishment mechanism.*

4. Conclusion

Our experimental evidence reveals that in the context of a decentralized (peer-to-peer) punishment system, antisocial punishment is widespread in Georgia, undermining cooperation and welfare. Instead of serving as a disciplinary mechanism for free riders, peer-to-peer punishment became a weapon that fueled conflicts among group members. In our study, we found that the decentralized punishment system in Georgia was less effective in increasing contributions and welfare compared to a centralized punishment system. Our findings suggest that centralized rule-based punishment systems can successfully promote cooperation in social dilemmas. However, it is important to note that our study focused solely on the credibility of a centralized incentive system aimed at deterring free-riding behavior, while excluding the risk of misuse of punishment power. Even in the case of weak and less credible centralized punishment incentives, both cooperation and welfare were higher compared to conditions with peer-to-peer punishment. Therefore, our comparative analysis suggests that self-governed groups facing

inefficiency due to a peer-to-peer punishment mechanism may benefit from the allocation of punishment power to external centralized authority, as long as there is no risk of power abuse. While we believe that our findings shed some light on the implications of punishment mechanisms in fostering cooperation within a cultural context, it would be desirable for future research to investigate the relative effectiveness of exogenous centralized punishment incentives under the threat of power abuse.

References:

1. Andreoni, James, and Laura K Gee. 2012. "Gun for hire: Delegated enforcement and peer punishment in public goods provision." *Journal of Public Economics* no. 96 (11-12):1036-1046.
2. Arechar, Antonio A, Simon Gächter, and Lucas Molleman. 2018. "Conducting interactive experiments online." *Experimental Economics* no. 21 (1):99-131.
3. Baldassarri, Delia, and Guy Grossman. 2011. "Centralized sanctioning and legitimate authority promote cooperation in humans." *Proceedings of the National Academy of Sciences* no. 108 (27):11023-11027.
4. Carpenter, Jeffrey, Shachar Kariv, and Andrew Schotter. 2012. "Network architecture, cooperation and punishment in public good experiments." *Review of Economic Design* no. 16 (2):93-118.
5. Engl, Florian, Arno Riedl, and Roberto Weber. 2021. "Spillover effects of institutions on cooperative behavior, preferences, and beliefs." *American Economic Journal: Microeconomics* no. 13 (4):261-99.
6. Fehr, Ernst, and Simon Gachter. 2000a. "Cooperation and punishment in public goods experiments." *American Economic Review* no. 90 (4):980-994.
7. Fischbacher, Urs. 2007. "z-Tree: Zurich toolbox for ready-made economic experiments." *Experimental Economics* no. 10 (2):171-178.
8. Gurerk, Ozgur, Bernd Irlenbusch, and Bettina Rockenbach. 2006. "The competitive advantage of sanctioning institutions." *Science* no. 312 (5770):108-111.
9. Herrmann, Benedikt , Christian Thöni, and Simon Gachter. 2008. "Antisocial Punishment Across Societies." *Science* no. 319:1362-1367.
10. Herrmann, Benedikt; Thöni, Christian; Gächter, Simon (2017), Data from: Antisocial punishment across societies, Dryad, Dataset, https://doi.org/10.5061/dryad.87301
11. Mekvabishvili, R. (fothecoming), in *Journal of Behavioral Economics for Policy,* "Weak and Strong Formal Institutions in Resolving Social Dilemmas: Are They Double-Edged Swords?"
12. Mekvabishvili, R. (2021). "Centralized Punishment in Public Good Experiments". Dataset, Zenodo, DOI: https://doi.org/10.5281/zenodo.5033369
13. Nikiforakis, Nikos. 2008. "Punishment and counter-punishment in public good games: Can we really govern ourselves?: Journal of Public Economics no. 92 (1-2):91-112.
14. Nosenzo, Daniele, and Martin Sefton. 2012. Promoting cooperation: the distribution of reward and punishment power. CeDEx Discussion Paper Series.
15. O'Gorman, Rick, Joseph Henrich, and Mark Van Vugt. 2009. "Constraining free riding in public goods games: designated solitary punishers can sustain human cooperation." *Proceedings of the Royal Society B: Biological Sciences* no. 276 (1655):323-329.
16. Olson, Mancur. 1975. *The Logic of Collective Action*. Vol. CXXIV. London, England: Harvard University Press.
17. Qin, Xiangdong, and Siyu Wang. 2013. "Using an exogenous mechanism to examine efficient probabilistic punishment." *Journal of Economic Psychology* no. 39:1-10