CIDE

LATIN AMERICAN ECONOMIC REVIEW

Volume 34: Issue on Experimental Economics | Article 5 | 2025

Sustained Cooperation in a Public Good Game: Evidence from Guyana

Edward Cartwright* Thomas Singh†

Submission received: January 29, 2024
Final version received: October 28, 2025
Accepted: November 15, 2024
Published: February 18, 2025

Abstract

Two consistent findings from the experimental literature on public good games are that cooperation declines over time and cooperation is lower in countries with weak institutions. These findings, however, are primarily based on experiments in Europe, North America, and Asia. There is little evidence from South and Central America. In an experiment conducted in Guyana, we found consistent, indeed rising, levels of cooperation over time. The robustness of this result was checked across three different treatments and has high power (a total of 176 subjects). Our results indicate that more experimental work is needed to fully understand willingness to cooperate in public good games. Guyana has relatively weak institutions and yet cooperation remained high.

Keywords: public good, cooperation, Guyana, punishment.

*Department of Accounting, Finance and Economics, De Montfort University, Leicester, LE1 9BH, United Kingdom. E-mail: edward.cartwright@dmu.ac.uk.

†University of Guyana GREEN Institute, University of Guyana, Georgetown, Guyana.

1. Introduction

Public good experiments have a long history in economics and psychology as a means to study cooperation in small groups (Ledyard, 1995; Drouvelis, 2021; van Dijk and De Dreu, 2021). In the basic setting, members of a group are given an endowment of private money and asked how much they want to contribute to the public good. Group payoffs are maximized if group members fully contribute to the public good. Individual group members, however, have a material incentive to 'free-ride' and not contribute to the public good. We, thus, obtain a social dilemma in which the interests of the individual and group are misaligned. Basic research questions are whether groups can mutually cooperate and, if not, what mechanisms can promote cooperation (Zelmer, 2003; Chaudhuri, 2011).

Two accepted stylized facts from the public good experimental literature are that: (a) Many group members cooperate the first time they play a public good game. Specifically, average contributions are typically around a half of the endowment. This level of contribution is socially inefficient (because efficiency is maximized with full cooperation) but well above the zero contributions we would expect with 'selfish' behavior. We, thus, observe significant levels of cooperation (e.g. Fehr and Gintis, 2007; Gachter and Thoni, 2007; Gächter et al., 2017; Otten et al., 2022). (b) When group members interact repeatedly the level of contribution declines. We, thus, see declining levels of cooperation over time. Various studies have looked to analyze and explain this decline (e.g Andreoni, 1995; Croson, 2007; Fischbacher and Gächter, 2010; Burton-Chellew et al., 2017; Burton-Chellew and West, 2021). Note that stylized facts (a) and (b) determine, respectively, the initial point and direction of travel, and so are best evaluated in combination.

In this paper we report the results of a public good experiment conducted in Guyana, South America. Our baseline treatment followed a very standard protocol (Drouvelis 2021) and was part of a larger project looking at framing in public good games. In that baseline treatment we observed initial contributions consistent with stylized fact (a) and then an increase in cooperation over time. This pattern of contributions is not consistent with stylized fact (b) and very unusual to see. Intrigued by what we saw in the baseline treatment we ran two further treatments to test the robustness of our results. In a Long treatment we increased the periods of repetition from 10 to 30. In a Group treatment we limited the feedback that group members received at the end of each period to merely total group contribution (rather than also individual contributions). In both these treatments we continued to observe stable levels of cooperation over time. In short, we continued to see evidence for stylized fact (a) but saw no evidence for stylized fact (b).

Our experiment has good power to detect a decline in contributions, and so we are confident that in this experimental setting there was no decline in contributions. We, therefore, question what aspects of the experimental setting led to stable contributions. Given that our experimental protocol was standard, and one we have used many times before in the UK, the natural conjecture is that people might 'behave differently' in Guyana. But, if they behave differently in Guyana then where else to people behave differently? The problems of relying on Western, Educated, Industrialized, Rich and Democratic (WEIRD) subjects has been widely shown, including in studies of cooperation (Henrich et al., 2010a,b). We, thus, suggest the robustness of stylized fact (b) needs further analysis.

In making this claim we recognize that there are handful of cross-cultural studies of public good games. Most notably, Herrmann et al. (2008) conducted public good experiments across 16 different locations (see also Gächter et al., 2010). These were European (Athens, Bonn, Copenhagen, Dnipropetrovs'k, Istanbul, Minsk, Nottingham, St. Gallen, Zurich) and Asian (Chengdu, Muscat, Riyadh, Samara, Seoul) plus Boston and Melbourne. A focus of this work was to see whether the possibility for within group punishment increased cooperation (because free-riders faced a threat of punishment). Across all locations they find that cooperation declines over time in the absence of punishment. They also find significant cross-cultural differences in the effectiveness of punishment, due to anti-social punishment (see also Gächter and Herrmann, 2009, 2011).

Table 1: *Measures of GDP (PPP $1,000), rule of law, corruption, government effectiveness, representation and rights in Guyana and countries studies in Herrmann et al. (2008).*

	GDP/capita		Rule of law		Corruption		Effectiveness		Representation		Rights	
	2008	2019	2008	2019	2008	2019	2008	2019	2008	2019	2008	2019
Guyana	3.3	6.1	-0.58	-0.45	-0.56	-0.14	-0.18	-0.40	0.53	0.6	0.52	0.58
Australia	49.2	56.4	1.76	1.69	2.03	1.79	1.78	1.54	0.85	0.84	0.9	0.86
Belarus	6.6	6.4	-1.07	-0.86	-0.56	0.01	-1.17	-0.17	0.29	0.27	0.48	0.48
China	3.4	9.8	-0.48	-0.26	-0.53	-0.31	0.17	0.54	0	0	0.34	0.35
Denmark	64.5	61.7	1.95	1.83	2.38	2.12	2.24	1.87	0.89	0.9	0.99	0.96
Germany	46.4	48.0	1.73	1.58	1.75	1.87	1.51	1.50	0.99	0.93	0.93	0.94
Greece	32.2	19.8	0.85	0.16	0.13	0.01	0.58	0.31	0.85	0.85	0.79	0.79
Oman	25.1	19.9	0.59	0.53	0.51	0.42	0.39	0.21	0.35	0.35	0.42	0.4
Russia	12.5	11.3	-1.01	-0.77	-1.13	-0.82	-0.42	0.02	0.39	0.35	0.43	0.39
Saudi Arabia	23.4	28.0	1.80	1.86	2.11	1.95	2.03	1.92	0	0	0.25	0.29
South Korea	21.4	33.4	0.86	1.16	0.43	0.71	0.99	1.33	0.82	0.81	0.73	0.76
Switzerland	74.7	85.5	0.02	-0.35	0.12	-0.33	0.29	-0.02	0.83	0.83	0.9	0.92
Turkey	10.8	9.5	-0.70	-0.74	-0.86	-0.80	-0.79	-0.33	0.64	0.45	0.53	0.37
Ukraine	3.9	3.1	1.70	1.57	1.67	1.75	1.63	1.45	0.53	0.54	0.54	0.52
UK	47.9	43.4	1.63	1.42	1.44	1.18	1.61	1.45	0.82	0.81	0.78	0.73
USA	48.5	62.8	0.07	0.14	-0.01	0.25	-0.11	0.26	0.81	0.65	0.79	0.72

A key finding of Herrmann et al. (2008) and Gächter et al. (2010) was that anti-social punishment is higher (and, thus, cooperation lower) in societies with weak norms of civic cooperation and a weak rule of law. Our results go against this finding. In Table 1 we provide some data for Guyana and the 15 countries studied by Herrmann et al. (2008). We include an estimate of GDP per capita from the IMF, rule of law, corruption and government effectiveness from the World Bank, and representation and rights from the Global State of Democracy Indices. For completeness we provide data from both 2008 and 2019 to capture any changes over time. You can see in Table 1 that Guyana scores relatively low for all of the measures considered. Yet we observe relatively high levels of cooperation and no anti-social punishment. The results of Herrmann et al. (2008) and Gächter et al. (2010) were essentially driven by low levels of cooperation in Russia, Belarus, Ukraine, Saudi Arabia, Oman, Greece and Turkey. Our results suggest that Guyana is 'different'.

In judging this difference, it is notable that little experimental work on cooperation has been run in South America. Spadaro et al. (2022) performed a comprehensive meta-analysis of estimates of cooperation based on a wide array of games including public good games. Their overall conclusion is a lack of evidence of cross-cultural difference in cooperation. This finding is more in line with our results. Interestingly, they identify only 46 experimental data points for Latin America, of which 30 are for Colombia. This compares to 954 data points for the USA, 211 for Germany, and 163 for Great Britain. Interestingly, Lopez et al. (2012) conducted a public good experiment on the Caribbean coast of Colombia and obtained similar results to us. In particular, they found sustained cooperation with no drop off in contributions over time. Clearly, therefore, there is a need for more studies of cooperation in South America to inform our overall understanding of cooperation in small groups.

We proceed as follows. In Section 2 we introduce the public good game. In Section 3 we explain our experimental design. In Section 4 we provide our results. In Section 5 we conclude. The experiment instructions are contained in supplementary material.

2. Public good game

We introduce a public good game (without punishment). There are four players in a group, each endowed with 20 units of private tokens. Players independently and simultaneously have the opportunity to contribute tokens into a group project. Any tokens contributed to the group project are doubled in value and then split equally amongst the four group members. This results in a marginal per-capita return (MPCR) form the public good of 0.5. Let c_i denote the contribution of player i for $i = 1, .., 4$ and let $C = c_1 + \ldots + c_4$ denote total contributions. The payoff of player i can be written:

$$U_i = 20 - c_i + 0.5C. \tag{1}$$

The public good game is widely studied in the literature (Zelmer, 2003; Chaudhuri, 2011; Drouvelis, 2021).

We distinguish between a public good game with *individual feedback* and *group feedback*. In the former case each player is told, at the end of the game, the individual contributions of all other players in the group. In the case of group feedback players are merely told the total contribution to the public good. The comparison between individual and group feedback has been studied in several experiments (e.g. Sell and Wilson, 1991; Croson, 2001; Pereda et al., 2019). An overall finding from this work is that contributions are (weakly) higher with individual than group feedback (Bigoni and Suetens, 2012).[1] This finding is important because individual feedback, as we now describe, is the natural starting point in a public good game with punishment. The comparison between a public good game and public good game with punishment needs, therefore, to control for the consequence of individual level feedback.

A public good game with punishment is a two stage game. In the first stage players play a public good game. Each player is then subsequently informed of the individual contributions of the other players. The four players then simultaneously and independently have the option to punish others. Specifically, player i is asked to assign punishment points to each player $j \neq i$ in the group. Let $p_{ij} \in [0, 7]$ denote the number of punishment points assigned. Each punishment point reduces the payoff of i by 1 token and the payoff of player j by 3 tokens. Let $P_i = \sum_j p_{ij}$ denote the number of punishment points assigned by player i and let $R_i = \sum_j p_{ji}$ denote the number of punishment points assigned to player i. The final payoff of player i can be written:

$$U_i = 20 - c_i + 0.5C - P_i - 3R_i. \tag{2}$$

The public good game with punishment has been widely studied, although with variants on the punishment mechanism (e.g. Fehr and Gächter, 2000, 2002; Nikiforakis and Normann, 2008; Nikiforakis, 2010). Note that a public good game with punishment requires individual feedback on contributions.

Given that the MPCR is less than 1, the unique Nash equilibrium of the public good game is for every player to contribute 0. By contrast, given than the MPCR is greater than $1/4$, the total group payoff is maximized when every player contributes 20. We obtain, therefore, a social dilemma in which the interests of the individual and group clash. Punishment does not change this trade-off. In particular, in the second stage of a public good provision game with punishment there is no material incentive to punish because it lowers own payoff. The unique Nash equilibrium is, therefore, for players to contribute 0 in the first stage and never punish in the second stage. Note that these arguments naturally extend to a finitely repeated sequence of the public good game.

[1] Studies have also looked at the impact of providing feedback on earnings as opposed to contributions (Nikiforakis, 2010). Throughout, we focus on feedback on contributions.

3. Experiment Design

In our Baseline treatment subjects played a public good with individual feedback for 10 periods and then a public good game with punishment for a further 10 periods. Their was fixed matching to groups across the 20 periods. We remark that this treatment was conducted as part of a four treatment study on framing that will be reported separately. We observed 'surprisingly' sustained cooperation in this treatment (without punishment) and so ran two further treatments to test the robustness of this finding. Our three treatments are summarized in Table 2.

In terms of our Baseline experiment design it is worth to note that the initial instructions informed subjects there would be two parts to the experiment with the first part lasting 10 periods (and instructions for part 2 to follow after the 10th period). Subjects were, thus, informed they would play a public good game for 10 periods. Full instructions are available in the Appendix.

In our Long treatment subjects played a public good game with individual feedback for 30 periods. In this treatment, subjects were informed that they would play the public good game for 30 periods in fixed groups.[2] Our motivation for the Long treatment is to test whether cooperation could be sustained beyond 10 periods over a longer time horizon. As pointed out by a reviewer, the Long treatment also has the advantage of removing the need to tell subjects there will be two parts to the experiment. It, thus, also allows us to test whether cooperation could be sustained when subjects are not expecting a second part to the experiment (which may influence their behavior in the first part of the experiment). Gachter et al. (2008) provide evidence that longer play yields similar levels of cooperation in both a public good game without punishment and with punishment. In the setting without punishment they observe contributions falling from an average of around 9.5-11.5 in period 1 to around 4-5 in period 10 (for their baseline treatment) and period 50 (in their long treatment). There is, thus, a notable drop in contributions over the course of their experiment but the overall drop is similar in the short and long treatments.

In our Group treatment subjects played a public good with group feedback for 10 periods and then a public good game with punishment for a further 10 periods. We remind that in a public good game with punishment there is individual level feedback by default. Thus, a 'cleaner' evaluation of the impact of punishment is provided by having, as in our Baseline and Long treatment, individual level feedback in the setting without (and with) punishment. Many experimental studies, however, use a public good game with group feedback (e.g. Gachter et al., 2008). Given that individual feedback has been shown to increase contributions we felt it was important to check whether, or to what extent, sustained cooperation is driven by individual feedback. One plausible mechanism whereby individual information can influence behavior is awareness of heterogeneity in contributions. To give a stylized example, suppose that a player contributes 10 and others in the group contribute 20, 4 and 0. With group level feedback the player can merely infer that the average contribution of others was 8 and, given this is less than the 10 they contributed, they may decrease their own contribution. If, by contrast, the player can see that someone else in the group contributed 20 it may positively influence their own contribution. Individual level feedback can, thus, generate very different contribution dynamics when compared to group level feedback (Pereda et al., 2019). We wanted to test if contributions can be sustained with both individual and group level feedback.

The experiment took place in a lab at the University of Guyana in 2014 with subjects recruited from the student population. The experiment was widely advertised on campus and the majority of subjects were social science majors. The experiment used z-tree (Fischbacher, 2007). The experiment was incentivized with total tokens earned converted into Guyanese Dollars at a rate of G$10 per token. This resulted in average earnings of around G$6000 which, at the time, was equivalent to around US$30. This compares to a minimum wage of

[2]Evidence suggests that a 'restart' can lead to an increase in contributions (Chaudhuri, 2018) and so it is important subjects knew in advance the game would last 30 periods.

around G$40,000 per month in Guyana. Subjects were paid privately at the end of the experimental session. A total of 176 subjects took part in the experiment. Given that our Long and Group treatments were motivated by initial findings from the Baseline treatment we remark that the Long and Group sessions took place within a couple of weeks of the Baseline sessions. Moreover, it was not possible for a subject to take part in more than one treatment (and so those in the Baseline were excluded from the Long and Group sessions).

Table 2: *Experimental treatments and number of subjects.*

Treatment	Design	Subjects	Groups
Baseline	10 periods without punishment and individual feedback. 10 periods with punishment	56	14
Long game	30 periods without punishment and individual feedback	60	15
Group information	10 periods without punishment and group feedback. 10 periods with punishment	60	15

4. Results

In Figure 1 we plot average contributions by treatment over the 20 or 30 periods of play. In Table 3 we detail the analogous average contribution in periods 1, 10, 11, 20, 21 and 30 as well as average contributions over 5 and 10 period blocks. In Period 1 we see that average contributions are around 9-10.5. The level of average contribution we observe in period 1 is consistent with prior experimental evidence (e.g. Herrmann et al., 2008). There are no statistically significant differences in contributions across the three treatments in period 1 ($p = 0.52$ Kruskal-Wallis test).

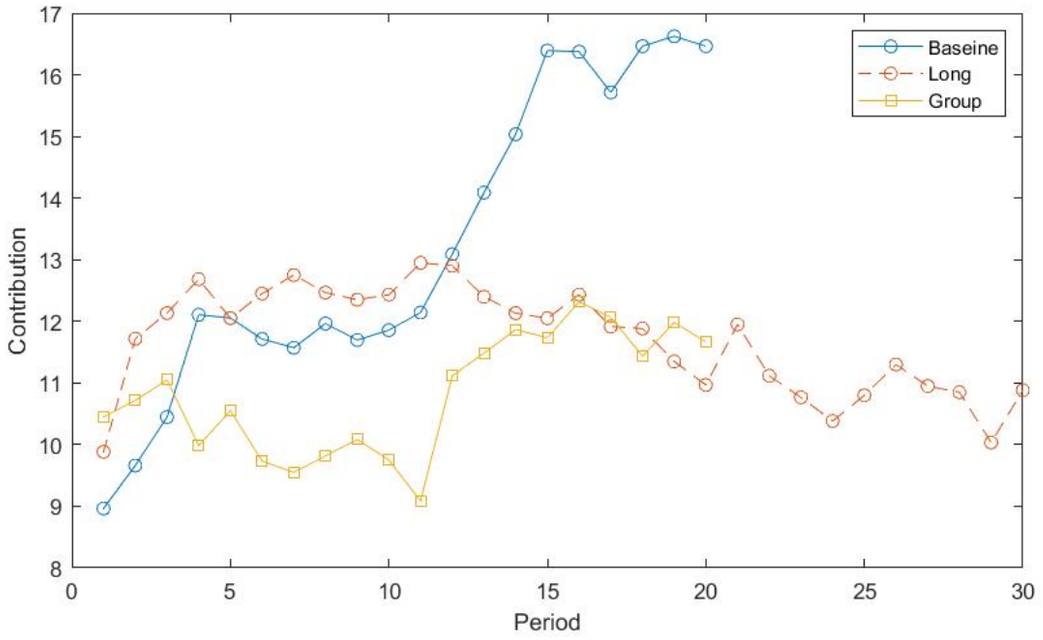
In the Baseline treatment we see that average contributions increase between period 1 and period 10 from 8.96 to 11.86. Of the 14 groups, contributions increase in 9 groups, remain the same in 2 groups and decrease in 3 groups. The difference in contributions between periods 1 and 10 is not statistically significant ($p = 0.073$ Wilcoxon rank sum test, two-sided, using group as the unit of observation). We do not, therefore, observe the conventional decline in average contributions typically observed in public good experiments.

As discussed in the preceding section, a motivation for the Long treatment was to test the robustness of our finding that contributions did not decline in the baseline treatment. You can see in Figure 1 that average contributions in the Long treatment follow a very similar pattern in the first 10 periods to those in the Baseline treatment. Specifically, average contributions rise from 9.88 to 12.43. Of the 15 groups contributions increase in 10 groups and fall in 5 groups. Again, the difference in contributions between periods 1 and 10 is not statistically significant ($p = 0.26$ Wilcoxon Test). We can, thus, state with some confidence that there is no evidence of contributions declining over the first 10 periods.

Between periods 11 and 30 contributions decline in the Long treatment. Average contributions in period 30 are still, however, above those in period 1, 10.88 compared to 9.88. The difference in contributions between period 1 and 30 is not statistically significant ($p = 0.48$ Wilcoxon Test). We now state our first result. We remark that this result, and those to follow, should be interpreted as a finding from our experiment in Guyana. We will discuss the implications of our work for a more general understanding of cross-cultural behavior in public goods in the concluding discussion.

Result 1. *In a public good game with individual information (without punishment) we observe average contributions remaining stable over time at or above the level observed in period 1.*

Figure 1: *Average contributions across the three treatments.*



Our main motivation for the Group treatment was to explore whether Result 1 was driven by the availability of information on invidivudal contributions. As you can see in Figure 1 average contributions decline in the Group treatment over the first 10 periods, from 10.45 to 9.75. Of the 15 groups, contributions increase in 6 and decrease in 9 over the 10 periods. Again, the the difference in contributions between periods 1 and 10 is not statistically significant ($p = 0.56$ Wilcoxon Test). We also see in Figure 1 that average contributions are lower in the Group treatment than either the Baseline of Long treatment by period 10. This is consistent with prior results (Bigoni and Suetens, 2012). The difference is, however, not statistically signification ($p = 0.22$ Wilcoxon Test comparing the Group treatment with the Baseline and Long treatments combined). While, therefore, there are hints of an impact of individual versus group information it seems unlikely that information is driving Result 1.

Result 2. *In a public good game (without punishment) we observe that contributions are lower, but not significantly so, with group information compared to individual information.*

In exploring why contributions remain relatively stable across the three treatments we summarize in Figure 2 the distribution of contributions in the first 10 periods. We detail both the distribution, subfigures (a) and (b). We also detail the cumulative distribution in period 1 and period 10, subfigures (c) and (d). We see in Figures 2 (c) and (d) a slight increase in the proportion of contributions above 10 in period 10 compared to period 1, consistent with Result 1. Overall, however, the distribution of contributions is similar across treatments and periods. A particularly noteworthy aspect of the distribution is the lack of 'free-riders' contributing 0 or a low amount. Previous studies suggest that around a quarter of individuals can be classified as free riders (e.g. Fischbacher et al., 2001; Thöni and Volk, 2018; Weber et al., 2018). Although Herrmann and Thöni (2009) find lower levels in Russia (2-10%). If the conventional decline in contributions is driven by conditional cooperation and a reaction to free-riding Fischbacher and Gächter (2010), then the absence of free-riding may help explain why subjects were able to sustain cooperation. Only in the Group treatment did we see any evidence of free-riding and even here it is less than 20% of subjects in period 10.

Result 3. *In a public good game (without punishment) we observe very few subjects contributing zero (or a*

Table 3: *Average contribution by treatment across periods*

Period(s)	Baseline	Long	Group
1	8.96	9.88	10.45
1-5	10.65	11.69	10.55
5-10	11.76	12.49	9.79
10	11.86	12.43	9.75
1-10	11.21	12.09	10.17
11	12.14	12.95	9.08
11-15	14.15	12.49	11.06
16-20	16.33	11.71	11.89
20	16.46	10.97	11.67
11-20	15.24	12.10	11.475
21		11.95	
21-25		11	
26-30		10.8	
30		10.88	
21-30		10.9	

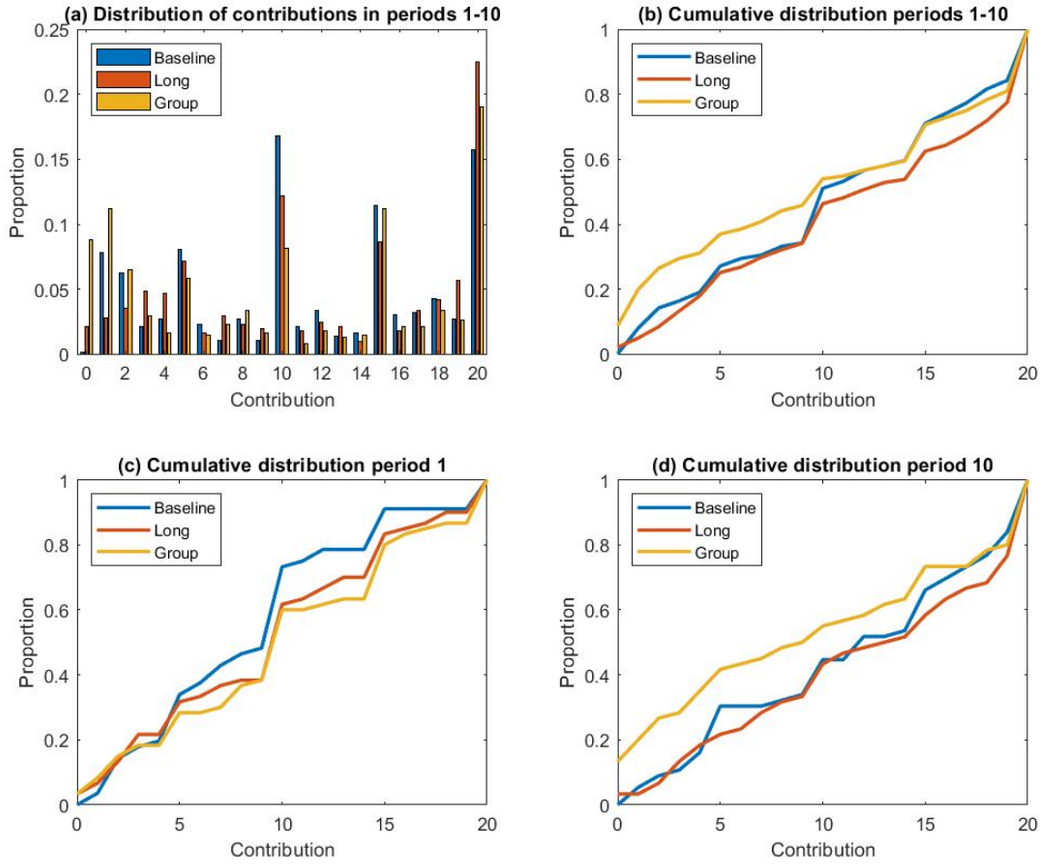
low amount).

We briefly look at the impact of punishment. In the Baseline treatment we observe, see Figure 1 and Table 3 and the comparison between periods 1-10 and 11-20, that contributions are higher in the periods with punishment than without. This holds whether we look at the respective averages across 10 periods (11.2 compared to 15.2, $p = 0.02$ Wilcoxon test), compare period 10 with period 20 ($p = 0.001$) or the average over periods 11-20 with period 10 ($p = 0.041$). In the Group treatment, by contrast, while contributions are higher in the periods with punishment than without the difference is not statistically significant. Again, this holds whether we compare averages across 10 periods (10.2 compared to 11.5, $p = 0.62$ Wilcoxon test), compare period 10 with period 20 ($p = 0.42$) or the average over periods 11-20 with period 10 ($p = 0.38$). Moreover, by period 20, contributions are significantly higher in the Baseline than Group treatment (16.5 compared to 11.7, $p = 0.043$ Wilcoxon test).

The impact of punishment appears, therefore, to be influenced by information. Recall that the second 10 periods of the Group treatment are strategically equivalent to those of the Baseline treatment because individual information is provided in both cases. In Figure 1 it is notable that average contributions *fall* in period 11 in the Group treatment. A closer look at behavior in individual groups suggests, however, that this is unlikely to be critical in explaining the differences between the Baseline and Group treatment. Specifically, contributions fell between periods 10 and 11 in 5 out of the 14 groups in the Baseline treatment and 8 out of 15 groups in the Group treatment and so there is no clear difference between the two treatments. Moreover contributions were higher in period 20 than 10 in 11 out of 14 groups in the Baseline treatment and 11 out of 15 groups in the Group treatment and so there is, again, no clear difference. We do find, however, that in those groups where contributions increase between periods 20 and 10 the difference is larger in the Baseline treatment compared to the Group treatment (25.9 vs 14.1, $p = 0.02$ Wilcoxon test). This then contributes to punishment having an overall greater impact in the Baseline treatment.

To further explore the lower contributions in the Group treatment we look at evidence for anti-social punishment. In Figure 3 we detail the average punishment points a player received as a function of their contribution relative to that of the group average. We see that those who contribute less than the group average receive more punishment point than those contributing more. There is, therefore, no evidence of anti-social punishment, particularly in the Baseline treatment. There is also no evidence of a difference between

Figure 2: *(a) Distribution and (b) cumulative distribution of contributions across the three treatments in the first 10 periods as well as (c) cumulative distribution in period 1 and (d) period 10.*



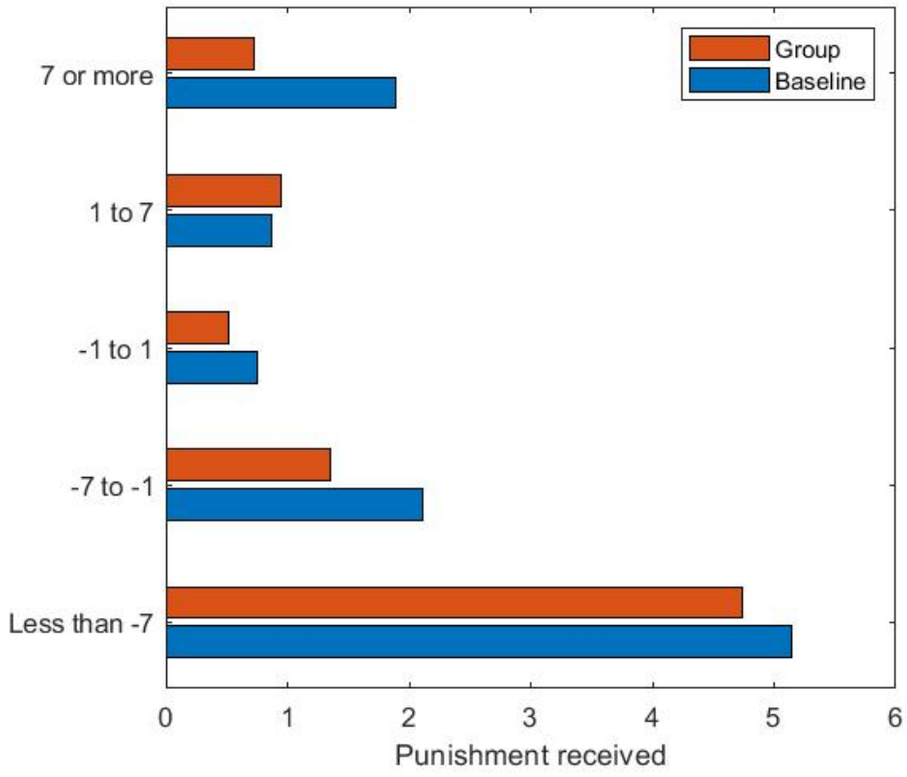
the Group and Baseline treatments. It remains, therefore, a puzzle why punishment did not have the positive impact on contributions in the Group treatment as it did in the Baseline treatment. This could be explored further in future work.

Result 4. *In a public good game with punishment we observe no evidence of anti-social punishment. Cooperation is higher with punishment than without in the Baseline treatment (where there is individual feedback in periods 1-10) but not significantly so in the Group treatment (where there is group feedback in periods 1-10).*

5. Concluding discussion

In this paper we have reported the results of a public good experiment conducted in Guyana, South America. Across three treatments we have shown that cooperation levels remain stable (and, if anything, increase over time). This finding goes against the stylized fact in the experimental economics literature of decreasing contributions over time. Obviously, one experimental study can obtain results that (by chance) go against convention, without us needing to revisit that convention. In our case, however, we note that our experiment yielded many results otherwise consistent with the literature, most notably: weakly higher contributions with

Figure 3: *Average punishment received as a function of contribution relative to the group average.*



punishment than without, weakly higher contributions with individual information than group information, and a consistent pattern of punishment. Moreover, existing evidence relies heavily on WEIRD subjects with almost no studies in South America. Lopez et al. (2012) who conduct a study in Colombia also find, like us, sustained cooperation over time. We suggest, therefore, that our study is documenting behavior that seems to be more than just a 'random outlier'.

We remind that Herrmann et al. (2008) find that norms of civic cooperation and the weakness of the rule of law are significant predictors of anti-social punishment, and therefore low levels of cooperation. As we detailed in Table 1, however, Guyana scores relatively low for all of the measures considered by Herrmann et al. (2008) (for which we could obtain comparable data). Yet we observe relatively high levels of cooperation and no anti-social punishment. Moreover, in the study of Herrmann et al. (2008) they find a drop in contributions over time in all subject pools. This drop in contributions is statistically significant at the 1% level in 12 of the subject pools (see their supplementary material); the exceptions being Ukraine, Greece, Saudi Arabia and Oman. In our experiment we observe increasing contributions over time. Lopez et al. (2012) also document high levels of cooperation in Colombia. The notion that weak institutions are correlated with anti-social punishment and lower levels of cooperation may, therefore, need further investigation.

The importance of taking into account cultural differences has been made many times before. Henrich (2000), for instance, uses the example of the Machiguenga in Peru and behavior in an ultimatum game to illustrate the role of culture. While subjects in Los Angeles split money 50-50 and rejected low offers, those in Peru kept more money for themselves and did not reject low offers. This difference in behavior appears to be driven by differences in cultural norms of fairness. In particular, it is not that Machiguenga proposers were less generous but more that there was no a-priori expectation that a 50-50 split was fair. Henrich et al. (2001)

report the results of experiments with 15 hunter-gatherer, nomadic herding and other small-scale societies. A key conclusion of their work is the role that everyday economic and social interactions play in influencing cooperative behavior. Societies, for instance, where sharing is normal in everyday life, like the Lamelara in Indonesia who hunt whales in large groups, are more likely to share in economic experiments. Norms of civic cooperation and the rule of law clearly feed into a person's everyday experience but are only one part of that. Indeed, Guyana has a history of strong social institutions and high levels of social capital, and at least up until a market system was introduced in 1989, a culture of cooperation among citizens.

Importantly, the public good game is one of the most used experiments to study cooperation; a wide range of theories of cooperation are based, in large part, on observed behavior in public good games (Drouvelis, 2021). The vast majority of those experiments are, however, based on WEIRD subjects. Herrmann et al. (2008) show that behavior in other subject pools may differ in terms of anti-social punishment. We provide evidence of difference in terms of changes in cooperation over time. Echoing the point of Henrich et al. (2001) there is, therefore, a need for more public good experiments with non WEIRD subjects and, crucially, theories of behavior that are able to capture the behavior observed in a cross cultural context.

References

Andreoni, James (1995), "Cooperation in public-goods experiments: kindness or confusion?" *The American Economic Review*, 891–904.

Bigoni, Maria and Sigrid Suetens (2012), "Feedback and dynamics in public good experiments." *Journal of Economic Behavior & Organization*, 82, 86–95.

Burton-Chellew, Maxwell N, Claire El Mouden, and Stuart A West (2017), "Social learning and the demise of costly cooperation in humans." *Proceedings of the Royal Society B: Biological Sciences*, 284, 20170067.

Burton-Chellew, Maxwell N and Stuart A West (2021), "Payoff-based learning best explains the rate of decline in cooperation across 237 public-goods games." *Nature Human Behaviour*, 5, 1330–1338.

Chaudhuri, Ananish (2011), "Sustaining cooperation in laboratory public goods experiments: a selective survey of the literature." *Experimental economics*, 14, 47–83.

Chaudhuri, Ananish (2018), "Belief heterogeneity and the restart effect in a public goods game." *Games*, 9, 96.

Croson, Rachel TA (2001), "Feedback in voluntary contribution mechanisms: an experiment in team production." In *Research in Experimental Economicss*, 85–97, Emerald Group Publishing Limited.

Croson, Rachel TA (2007), "Theories of commitment, altruism and reciprocity: Evidence from linear public goods games." *Economic Inquiry*, 45, 199–216.

Drouvelis, Michalis (2021), *Social preferences: An introduction to behavioural economics and experimental research*.

Fehr, Ernst and Simon Gächter (2000), "Cooperation and punishment in public goods experiments." *American Economic Review*, 90, 980–994.

Fehr, Ernst and Simon Gächter (2002), "Altruistic punishment in humans." *Nature*, 415, 137–140.

Fehr, Ernst and Herbert Gintis (2007), "Human motivation and social cooperation: Experimental and analytical foundations." *Annu. Rev. Sociol.*, 33, 43–64.

Fischbacher, Urs (2007), "z-Tree: Zurich toolbox for ready-made economic experiments." *Experimental economics*, 10, 171–178.

Fischbacher, Urs and Simon Gächter (2010), "Social preferences, beliefs, and the dynamics of free riding in public goods experiments." *American economic review*, 100, 541–556.

Fischbacher, Urs, Simon Gächter, and Ernst Fehr (2001), "Are people conditionally cooperative? Evidence from a public goods experiment." *Economics letters*, 71, 397–404.

Gächter, Simon and Benedikt Herrmann (2009), "Reciprocity, culture and human cooperation: previous insights and a new cross-cultural experiment." *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364, 791–806.

Gächter, Simon and Benedikt Herrmann (2011), "The limits of self-governance when cooperators get punished: Experimental evidence from urban and rural Russia." *European economic review*, 55, 193–210.

Gächter, Simon, Benedikt Herrmann, and Christian Thöni (2010), "Culture and cooperation." *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365, 2651–2661.

Gächter, Simon, Friederike Mengel, Elias Tsakas, and Alexander Vostroknutov (2017), "Growth and inequality in public good provision." *Journal of Public Economics*, 150, 1–13.

Gachter, Simon, Elke Renner, and Martin Sefton (2008), "The long-run benefits of punishment." *Science*, 322, 1510–1510.

Gachter, Simon and Christian Thoni (2007), "in Voluntary Cooperation: Insights from Experimental Economics." *Rationality and commitment*, 175.

Henrich, Joseph (2000), "Does culture matter in economic behavior? Ultimatum game bargaining among the Machiguenga of the Peruvian Amazon." *American Economic Review*, 90, 973–979.

Henrich, Joseph, Robert Boyd, Samuel Bowles, Colin Camerer, Ernst Fehr, Herbert Gintis, Richard McElreath, et al. (2001), "Cooperation, reciprocity and punishment in fifteen small-scale societies." *American Economic Review*, 91, 73–78.

Henrich, Joseph, Steven J Heine, and Ara Norenzayan (2010a), "Most people are not WEIRD." *Nature*, 466, 29–29.

Henrich, Joseph, Steven J Heine, and Ara Norenzayan (2010b), "The weirdest people in the world?" *Behavioral and brain sciences*, 33, 61–83.

Herrmann, Benedikt and Christian Thöni (2009), "Measuring conditional cooperation: a replication study in Russia." *Experimental Economics*, 12, 87–92.

Herrmann, Benedikt, Christian Thoni, and Simon Gachter (2008), "Antisocial punishment across societies." *Science*, 319, 1362–1367.

Ledyard, John O (1995), "Public Goods: A Survey of Experimental Research." *The handbook of experimental economics*, 111.

Lopez, Maria Claudia, James J Murphy, John M Spraggon, and John K Stranlund (2012), "Comparing the effectiveness of regulation and pro-social emotions to enhance cooperation: experimental evidence from fishing communities in Colombia." *Economic Inquiry*, 50, 131–142.

Nikiforakis, Nikos (2010), "Feedback, punishment and cooperation in public good experiments." *Games and Economic Behavior*, 68, 689–702.

Nikiforakis, Nikos and Hans-Theo Normann (2008), "A comparative statics analysis of punishment in public-good experiments." *Experimental Economics*, 11, 358–369.

Otten, Kasper, Ulrich J Frey, Vincent Buskens, Wojtek Przepiorka, and Naomi Ellemers (2022), "Human cooperation in changing groups in a large-scale public goods game." *Nature Communications*, 13, 6399.

Pereda, María, Ignacio Tamarit, Alberto Antonioni, Jose A Cuesta, Penélope Hernández, and Angel Sánchez (2019), "Large scale and information effects on cooperation in public good games." *Scientific reports*, 9, 15023.

Sell, Jane and Rick K Wilson (1991), "Levels of information and contributions to public goods." *Social Forces*, 70, 107–124.

Spadaro, Giuliana, Caroline Graf, Shuxian Jin, Sakura Arai, Yukako Inoue, Eleanor Lieberman, Maria Isabela Rinderu, Mingliang Yuan, Caspar J Van Lissa, and Daniel Balliet (2022), "Cross-cultural variation in cooperation: A meta-analysis." *Journal of Personality and Social Psychology*.

Thöni, Christian and Stefan Volk (2018), "Conditional cooperation: Review and refinement." *Economics Letters*, 171, 37–40.

van Dijk, Eric and Carsten KW De Dreu (2021), "Experimental games and social decision making." *Annual Review of Psychology*, 72, 415–438.

Weber, Till O, Ori Weisel, and Simon Gächter (2018), "Dispositional free riders do not free ride on punishment." *Nature communications*, 9, 2390.

Zelmer, Jennifer (2003), "Linear public goods experiments: A meta-analysis." *Experimental Economics*, 6, 299–310.