Luck or merit: Implications for cooperation and punishment of free-riding in a public goods game

Ananish Chaudhuri[1]
Scott Claessens[2]
Michalis Drouvelis[3]
Abhijit Ramalingam[4,*]
Brock Stoddard[5]
Quentin Atkinson[6]

July 2023

Abstract

Empirical evidence suggests that humans are more accepting of inequality arising from differences in merit rather than luck. We provide causal evidence for this finding using an online linear public goods experiment with heterogeneous endowments. Such heterogeneity may result either from luck or from performance in a prior real-effort task. Over multiple rounds, participants first vote on whether to implement a costly centralised mechanism that punishes free-riding and then decide how much to contribute to the public good. Across all our treatments and compared to the poor, on average, the rich free-ride more even in the presence of punishment. The poor free-ride significantly more in a treatment where endowments are determined by luck as opposed to merit. Finally, the poor vote for stronger punishment of free-riding when endowments are luck-based rather than merit-based. This result suggests that even the poor are more accepting of merit-based endowment heterogeneity.

JEL Codes: C72; C91; C92; D91; H41

Keywords: Inequality; Public goods; Cooperation; Punishment; Voting; Online experiment

Acknowledgements: We gratefully acknowledge funding from the University of Auckland via a Faculty Research Development Fund. Ethics approval was obtained from the University of Auckland Human Participants Ethics Committee (Ref. 024639). We thank seminar participants at the University of Auckland and participants at the 2021 Western Economic Association Annual Meetings, the 2022 ESA World Meetings and the 2022 Southern Economic Association Annual Meetings for helpful comments and suggestions.

[1] Department of Economics, University of Auckland and CESifo, Munich, a.chaudhuri@auckland.ac.nz.
[2] School of Psychology, University of Auckland, scla896@aucklanduni.ac.nz.
[3] Department of Economics, University of Birmingham and CESifo, Munich, m.drouvelis@bham.ac.uk.
[4] Department of Economics, Appalachian State University, 416 Howard Street, Peacock Hall 3101A, Boone, NC 28608, USA. abhi.ramalingam@gmail.com. * Corresponding Author.
[5] Department of Economics, Appalachian State University, stoddardbv@appstate.edu.
[6] School of Psychology, University of Auckland, q.atkinson@auckland.ac.nz.

1. Introduction

As we start the third decade of the 21st century, issues of inequality loom large. Piketty (2014) argues that income and wealth inequality in developed nations are approaching levels not seen since before the Second World War. This gap will almost certainly widen further following Covid-19 (Stiglitz, 2020). However, there are individual differences in attitudes towards and tolerance of inequality depending on its source. Simply put, inequality can be the outcome of merit or luck (random chance). Almas et al. (2019) show that across a large cross-section of countries, there is higher (lower) tolerance of inequality, and even extreme inequality – measured via imputed values of the Gini coefficient – when this results from differences in merit (luck). Starmans et al. (2017) and Almas et al. (2010) provide evidence that these feelings are present even among children and adolescents. Macchia et al. (2020) report that many, particularly in developing nations, are tolerant of inequality since the relatively well-off actually feel a greater sense of well-being from their higher income rank. Indeed, as Almas et al. (2019) highlight, cross-country tolerance of inequality decreases with an increase in per capita GDP, with citizens of richer countries expressing greater aversion toward inequality.

Alesina et al. (2004) document a clear division in attitudes on two sides of the Atlantic, with Europeans being less tolerant of inequality than Americans. Alesina and Angeletos (2005) highlight how different beliefs about the fairness of social competition and the causes of inequality may influence redistributive policies. They show that based on beliefs regarding the causes of inequality and the degree of social mobility, there may exist multiple social equilibria. A society that believes individual effort (or the "signal" according to Alesina and Angeletos) determines income, will choose low taxes and low redistribution (and end up with potentially higher inequality) while a society that puts more weight on luck ("noise") in the form of accidents of birth, connections and/or corruption, may well choose higher taxes and higher redistribution resulting in lower inequality. In equilibrium perceptions of the signal to noise ratio will have implications for the choice of taxes, other redistributive policies and the size of the welfare state.

While more people may be tolerant of merit-based inequality at the aggregate level, there will be heterogeneity in perceptions and reactions to the different sources of inequality among different segments of the population. To the extent that the merit versus luck issue raises questions about earned versus unearned income and/or wealth, those who are less well-off (henceforth, for expositional ease, the "poor") may feel more aggrieved when inequality is a result of luck while the better off (the "rich") may feel more entitled to their wealth regardless of the source of inequality. Moreover, it is not clear if the poor and rich will support the same policies depending on the underlying causes of inequality. In this paper, we explore the effects of these two different sources of inequality on cooperative behaviour as well as preferences for redistribution through punishment of free-riding.

Using a paradigmatic linear public goods game (Ledyard, 1995), we look at individual level differences in contribution decisions and policy preferences when the causes of wealth dispersion differ but are also common knowledge. We ask: (1) How does the cause of inequality (luck or merit) impact cooperation? How is this different from when endowments are equal? (2) How do different causes of inequality affect cooperation in the presence of a redistributive, centralised sanctioning mechanism that punishes free-riding? And finally, (3) Do the poor and the rich vote for differing levels of sanctioning, and does this depend on the source of inequality?

It is difficult, if not impossible, to ascertain the true cause of wealth dispersion using natural data. Typically, both luck and merit are involved. A practical and cost-effective way to establish a causal relationship is to rely on controlled experiments (e.g., Falk and Heckman, 2009; Chaudhuri, 2021) where the true reason behind wealth dispersion can be exogenously manipulated. Inequality in this study is implemented by varying subjects' initial endowment of resources. We will sidestep the often-expressed concerns regarding external validity by noting that survey-based approaches also suffer from drawbacks. Surveys and experiments are complementary means of getting at the underlying truth. Indeed, as Smith (1976), in developing his induced value theory, points out, experiments can serve both as an empirical pretest prior to field testing as well as a means of interpreting field data. As the many studies cited in our work highlights, there is now a voluminous literature that explores attitudes toward and tolerance of inequality via economic experiments. Clark and D'Ambrosio (2015, Section 13.2.2) provide an overview of experimental studies of attitudes toward to inequality.

In an ambitious online experiment, we create three different types of unequal societies and allow participants to vote on policy that could reduce inequality. In our "merit" treatment, individual endowment (and therefore rich or poor status) is determined by performance in a prior real-effort task. Our "luck" treatment creates a situation where participants take part in the same real-effort task to start with, but endowments are determined purely by random chance. In the "uncertain" treatment, endowments may be determined by merit (on the basis of performance in the prior real-effort task) or by luck. Finally, behaviour in these three inequality treatments – merit, luck and uncertain – is compared to a fourth "equality" treatment where all endowments are equal.

A large literature in self-serving biases (see Babcock and Loewenstein, 1997 for an overview) suggests that the rich may feel entitled to their wealth and believe that this is the result of merit rather than luck, even where it is mostly due to luck. Similarly, the poor may feel hard done by, even where inequality is the result of merit, and may attribute this to luck alone. In our "uncertain" treatment, participants know that there is a half-chance that rich or poor status depends only on merit and a half-chance that it depends only on random chance. The "uncertain" treatment is novel and important since, in reality, the distinction of whether rich/poor status is the result of merit or luck is often unclear. This treatment then allows us to understand whether poor and rich participants perceive this third situation to be similar to luck-based or merit-based inequality.

Comparing across these different societies allows us to understand whether and how differences in the source of wealth heterogeneity lead to differences in contribution patterns. Then, we examine individual policy preferences in such settings. We look at enforcement of cooperative norms by allowing participants to establish a costly centralised scheme for punishing free-riding in the different societies. Participants in a group can vote for and pay the same small price (of course, a relatively larger proportion of endowment for the poor compared to the rich) to implement a central sanctioning scheme. The punishment scheme is infallible and represents a certain punishment on free-riding, with the level of the punishment being determined by majority vote. In any round, the part of one's endowment that is not contributed toward the public good is subject to this punishment if and when participants successfully implement the sanctioning mechanism. This stylised setting minimises potential confounds and, as discussed below, this approach is similar to those implemented in other studies exploring centralised punishment for free-riding.

Our findings may be summarised as follows. In all treatments other than Luck, compared to the rich, the poor contribute a much larger proportion of their endowment to the public good. But the poor free ride significantly more when inequality is luck-based than when it is merit-based. Second, the poor vote for stronger punishment when inequality results from luck than from other sources. Given the proportional differences in endowment, higher punishment for free-riding results in greater redistribution. What is striking is that this tendency on the part of the poor to punish free-riding is preeminent in the luck treatment, suggesting a greater desire for redistribution among the poor when inequality is luck-based rather than merit-based. There is no such difference in the voting and contribution behaviour of the rich.

A final interesting observation is that patterns of choices made by the rich and poor, with regards to cooperation and punishment, are similar in both the merit-based and uncertain inequality treatments, suggesting both perceive the uncertain treatment as being similar to the merit-based treatment. It appears that people treat a system that has some semblance of merit as being similar to a system based on merit. In sum, we formally establish the differential causal impact of, and response to, luck-based inequality as opposed to merit-based inequality.

We proceed as follows. In Section 2, we explain our experimental design and procedures. In Section 3, we present hypotheses on the basis of existing research findings. In Section 4, we present our results and, in Section 5, we make some concluding remarks.

2. Experimental design and procedures

There were two Parts in all sessions of the experiment. In Part 1, participants completed a real-effort task, performance on which, depending on the treatment, determined a participant's endowment in Part 2. In Part 2 of the experiment, participants played a repeated public goods game that could, depending on the treatment, additionally include opportunities for punishing free-riding. Instructions to all parts and treatments of the study are available in Appendix A.

2.1 Part 1: Real effort task

In all sessions and all treatments, participants first individually and independently complete a word coding task as in Erkal et al. (2011) for three minutes. In the task, participants are presented with a sequence of randomly formed three-letter 'words', which they translate to a numeric code using a table that assigns a number to each letter of the alphabet. Participants are presented with one 'word' at a time, and are presented with the next one upon submitting a code. Each correct code submitted earns a participant a piece rate of 30 tokens. Participants are not informed of whether they submitted a correct code, and are not shown a running total of the number of attempted or accurately completed codes. They are also not informed of the total number of correct codes and earnings from Part 1 until the conclusion of the experiment. This avoids any potential wealth effects, however small.

2.2 Part 2: Public goods game

2.2.1 The base decision setting

In all treatments, the base game was a four-player linear public goods game played in fixed groups (partner matching) that was repeated for 10 rounds. In the base game, each player received an endowment of $e_i > 0$ tokens, and independently and simultaneously decides how many tokens, $c_i \in [0, e_i]$, to contribute to a group account (the public good), with the remainder being automatically allocated to his/her private account. A player earned one token for each token retained in his/her private account, and a fraction m $(0 < m < 1 < 4m)$ of the total contribution to the public good by all group members, denoted as $C = \sum_{j=1}^{4} c_j$. Our chosen parameter value of the marginal per capita return (MPCR) of m = 0.5 satisfies the conditions above. Given the restrictions on m, the base game is a social dilemma that captures the conflict between privately and socially optimal actions. Assuming self-interested players who only care about monetary payoffs, the unique Nash equilibrium of the game is zero contribution by each group member. The social optimum is 100% contribution by each member. Under finite repetition, the subgame perfect Nash equilibrium is zero contribution by all, and the social optimum remains unchanged.

2.2.2 Treatment dimensions

Our treatments vary along two dimensions. The first dimension varies the presence and source of inequality among group members. The second dimension varies whether or not groups vote to implement a punishment on all allocations to the private account. We describe each dimension and our treatments below.

Dimension 1: Inequality and its sources

In *Equality* treatments, each group member receives a per-round endowment of 50 tokens, i.e., $e_i = 50 \; \forall \; i = 1, 2, 3, 4$. This is irrespective of performance in the prior real-effort task. In the other treatments, two 'poor' group members receive a per-round endowment of 20 tokens each, and two 'rich' group members receive a per-round endowment of 80 tokens each.[7] These other treatments vary in the source of the inequality. In all cases, endowments are decided once at the beginning of Part 2 and then remain fixed for all 10 rounds in Part 2.

In *Luck* treatments, two of the four group members are randomly assigned to be poor (have 20 tokens) while the other two are rich (have 80 tokens). This assignment is independent of performance in the word coding task in Part 1. In *Merit* treatments, endowments are decided based on performance in the effort task in Part 1. All four group members are ranked according to their performance in the word coding task. The two group members with the higher (lower) performance in Part 1 receive the higher (lower) endowment.[8] All of this is common knowledge.

In the *Uncertain* treatments, there is a 50% chance that endowments within a group will be allocated according to merit (i.e., the top two performers in the group are rich and the bottom

[7] We chose an equal number of poor and rich members within unequal groups to avoid complications that might arise due to the presence of a majority and a minority within groups (see Oxoby and Spraggon, 2013), such as tacit coalition formation when voting on institutional choice.
[8] Any ties are broken randomly.

two are poor). Else, endowments are allocated according to luck (i.e. two randomly chosen participants in the group are poor and the other two are rich).

Dimension 2: Endogenous centralised punishment

In treatments without the possibility to impose punishment, groups simply interact in the public goods game described above in each round. The payoff to a group member in each round is given by

$$\pi_i = (e_i - c_i) + mC \text{ with } C = \sum_{j=1}^{4} c_j$$

where the first term is the earning from the private account and the second is the earning from the public (group) account.

In treatments where punishments are possible (denoted *Pun*), group members interact in two stages in each round. In the first stage, group members simultaneously vote to implement one of four punishment rates – 0, 30%, 60%, or 80% – in their group in that round. An implemented punishment rate reduces each group member's private-account earnings by that percentage, whether they voted for it or not. A punishment rate is chosen by a simple majority, i.e., a 'first-past-the-post' rule. Any punishment rate that receives at least two votes in a group is implemented for that round. If two rates each receive two votes, one of them is chosen randomly for implementation. Regardless of what punishment rate a participant chooses, if a positive punishment rate is implemented in a round, each group member pays a fee of 4 tokens from his/her earnings in the round. There is no fee if a zero-punishment rate is implemented.

A punishment rate of zero does not change the fully free-riding equilibrium. Neither does a punishment rate of 30%, since retaining 70% of a token in one's private account is more profitable than contributing it to the public good and receiving a return of 50% (since $m = 0.5$). Hence, a punishment of 30% is *non-deterrent*. However, punishment rates of 60% and 80%

ensure that the loss from retaining a token in the private account is greater than 50%, which would be the loss from contributing it to the group account. Thus, with a higher punishment rate, full contribution becomes the dominant strategy. Hence, punishment rates of 60% and 80% are *deterrent*.

We chose two different values for the deterrent punishment because we wanted players to decide, first, whether to implement a punishment or not (0%) and, second, whether to implement a non-deterrent (30%) or deterrent punishment (60% or 80%). Providing three choices of 0, 30, and either 60 or 80 may have made 30% focal.[9]

Note that such punishment is standard in the literature dealing with centralised and decentralised punishment. In particular, the centralised sanction in our setting is not a tax that is levied on non-contributors, and is not redistributed among any of the group members. The sanction amount is simply lost to the group, and is thus costly and potentially inefficient.[10] See discussions of centralised punishment mechanisms in, for instance, Kosfeld et al. (2009), Putterman et al. (2011) and Baldassari and Grossman (2011), and decentralised punishment in Fehr and Gächter (2000) and Chaudhuri (2011). Given that implementation depends on voting by participants, our sanctioning mechanism incorporates elements of democratic punishment as in Ambrus and Greiner (2019) and Pfattheicher et al. (2018).

That all players paid the same fee of 4 tokens implies that the poor paid a share of their endowment that was four times that of the rich (20% vs. 5%). We chose this payment scheme rather than one that exacted the same share of endowment from the poor and the rich for its simplicity and ease of explanation to participants. Moreover, the fee as a share of endowment

[9] The data show that when participants chose to implement a punishment rate, they either voted for the non-deterrent (30%) punishment or the 80% punishment rate, i.e. the larger of the two deterrent punishment rates. There were very few votes for the 60% punishment rate.

[10] Prior research has noted that the efficiency implications of punishments are ambiguous; whether punishments raise efficiency or not depends crucially on the cost-benefit ratio of said punishments. See Chaudhuri (2011); Egas and Riedl (2008) and Nikiforakis and Normann (2008).

20% for poor and 5% for rich) mirrors the 4:1 imbalance in the ratio of endowments between the rich and the poor.

In the second stage, group members are informed of the punishment rate implemented in the first stage before they make contribution decisions in the round. The punishment is automatically and universally applied to all group members, regardless of how they voted. The payoff to a group member in each round is given by the following:[11]

$$\pi_i^{Pun} = \begin{cases} (e_i - c_i) + mC & \text{if } X = 0 \\ (1 - X)(e_i - c_i) + mC - 4 & \text{if } X \in \{0.3, 0.6, 0.8\} \end{cases},$$

where X is the implemented punishment rate.

Treatments

We have a 2 × 4 design, leading to a total of 8 treatments. Table 1 summarises our 8 treatments and lists the number of participants (and independent groups) in each.

Table 1. Experimental design: summary of treatments and number of participants

		Source of inequality			
Vote for punishment?	No inequality	Luck	Merit	Uncertain	Total
No	*Equality* 64 (16)	*Luck* 76 (19)	*Merit* 68 (17)	*Uncertain* 72 (18)	280 (70)
Yes	*Equality – Pun* 48 (12)	*Luck – Pun* 68 (17)	*Merit – Pun* 48 (12)	*Uncertain – Pun* 48 (12)	212 (53)
Total	112 (28)	144 (36)	116 (29)	120 (30)	492 (123)

Figures in parentheses are the numbers of independent groups of four members.

[11] There is no possibility of negative earnings and the minimum possible is zero. To see this, consider the case of a poor participant who votes for a deterrent 80% punishment at a cost of 4 tokens, which is implemented. Then this participant allocates all of his/her 20 tokens to the private account. With the 80% punishment rate this participant is fined 16 tokens leaving him/her with 4 tokens. Once the 4-token fee is deducted this participant ends up with zero.

2.3 Procedures

The experiments for this study were carried out via the on-line platform Prolific (https://www.prolific.co). To minimise location, timing and cultural differences among participants, we restricted participation to adults (18+ years old) based in the United Kingdom. The average age of participants was 34.8 (min = 18, max = 82) with 69.5% of participants being female. Participants log into the Prolific site where they first read a description of the study design, which includes the projected length and the expected payment from taking part. Those who consented to participate were directed by Prolific to Heroku, a cloud application platform that ran our oTree program (Chen et al., 2016).

Participants were told there would be two Parts to the experiment, but were only presented with instructions for Part 1 to start with. We ran the *No-Punishment* and *Punishment* treatments in separate sessions to keep the first part of the instructions (projected length and earnings) identical. Within the *No-Punishment* or *Punishment treatments*, the four treatments (along the inequality dimension) were run concurrently. Upon completion of Part 1, subjects read the instructions for their assigned treatment with or without a description of punishment, depending on the session, worked through a few illustrative examples and control questions[12], and played a hypothetical practice round (*with no interaction with others*). They then arrived at a waiting room. As soon as a group of four was assembled in the waiting room, they were assigned their endowments according to the treatment, and then began the public goods game (Part 2). In inequality treatments, participants were told the source of inequality in their groups

[12] In case participants answered a question incorrectly, they were presented with an explanation of the required calculation and were asked to answer the question again. They could not proceed until they had answered each question correctly.

in the instructions, and were reminded of this on the decision screen in each round. Participants remained anonymous at all times.

Participants were informed of their performance and earnings in Part 1 only at the end of the experiment. In all treatments, at the end of each round in Part 2, participants were informed of total contributions to the public good in their group in the round, and their own earnings in the round. They were not informed of individual contributions or earnings of other group members. In the *Punishment* treatments, group members were informed of the outcome of the vote (i.e., the implemented punishment rate) in each round before they made contribution decisions. They were not informed of the votes of other group members at any point.

At the end of the experiment, participants answered a few demographic questions (see Appendix A).[13] Finally, participants rated the fairness (on a 5-point Likert scale: 1 = Very unfair, to 5 = Very fair) of the mechanism used to determine the distribution of endowments in their group in Part 2.[14]

Participants were paid their earnings from Part 1 and accumulated earnings from all 10 rounds in Part 2. Token earnings were converted to cash using an exchange rate of 250 tokens to GBP 1. Participants in a *No-Punishment* (*Punishment*) treatment took around 25 (40) minutes on average to complete the experiment, and earned an average of GBP 8 (11), including a GBP 5 fixed payment for completing the experiment. The experiments reported in the study were approved by the University of Auckland Human Participants Ethics Committee (Ref. 024639).

[13] In addition, they answered a series of questions designed to measure their 'Social Dominance Orientation' (Sidanius and Pratto, 1999) and 'Right Wing Authoritarianism' (Altemeyer, 1981).

[14] We could have asked about perceived fairness at the outset immediately after the participants learn how endowments had been decided. However, this has the potential to prime participants and lead to experimenter demand effects affecting behaviour. On balance, we felt that asking this question later rather than earlier made more sense.

2.4 Drop-outs and sample size

Participants could exit the experiment at any time by closing their internet browser. Once they quit, they could not re-join the experiment. In Part 2, this presents a problem for group members who do not quit the study. To deal with this issue, we implemented a time limit for decisions in a group. If a group member did not enter a decision in a round (either because they quit or due to inattention) within 45 seconds, a vote for zero punishment was entered for that member (in the Punishment treatments) and his/her entire endowment was contributed to the group account in that round. If that group member did not enter a decision on time for two consecutive rounds, that member and his/her entire group were taken to the end of the study. That member (and those who quit at any stage) was not paid at all, while the other group members were paid for the rounds they completed. All participants were informed of this drop-out procedure at the beginning of the experiment, and were shown the time remaining to submit decisions on every decision screen.

We aimed to recruit 400 participants in each condition – *No Punishment* and *Punishment*. We ended up with 407 in *No Punishment* and 400 in *Punishment*. Within each of these conditions, we create four different types of societies as noted earlier: societies with equal (50 token) endowments for all for group members and three unequal (two with 20 tokens and two with 80 endowment) societies based on luck, merit or uncertainty.

We had a relatively large drop-out rate in both treatments leaving us (as noted in Table 1) with 280 participants (70 groups) in the no punishment condition and 212 (53 groups) in the punishment conditions. The number of participants in individual treatments is given in Table 1. This drop-out rate was unanticipated but the experiments described in this study are more elaborate and longer lasting compared to typical online experiments. We daresay that ours is one of the first studies to undertake such an elaborate online design. This was necessitated given the fact that running lab experiments with human participants were rendered impossible at the time due to Covid-19 related restrictions over much of the world.

However, the drop-outs we experienced are not unusual by the standards of online experiments. Arechar et al. (2018) provide an excellent overview of the methodological and practical challenges facing those undertaking such complex interactive experiments in an online environment. They conduct an online public goods experiment with decentralised punishment. In their experiment, 18% of all participants dropped out leading to only 53% of all groups that began the experiment completing it. Our experiments are longer lasting than theirs since we have the additional word coding task at the outset.

A detailed ex-ante and ex-post power analysis and an analysis of drop-outs during the online sessions are discussed in Appendices C and D, respectively.

3. Hypotheses

In the following, our hypotheses refer to both the public goods game without and with punishment. We look at punishment and contribution behaviour, which are closely intertwined. Prior results suggest that in the presence of endowment heterogeneity, and the absence of punishment, the rich contribute a smaller proportion of their endowment to the public good compared to the poor (e.g., Anderson et al., 2008; Cherry et al., 2005; Buckley and Croson, 2006; Hargreaves Heap et al., 2016; Zelmer, 2003).

In comparison, the literature looking at how punishments affect cooperation in unequal groups in the presence of a centralised sanction mechanism is less voluminous. Kingsley (2016) suggests that the effectiveness of peer punishment is weakened by endowment heterogeneity. Waichman (2020) finds peer punishment successfully increases efficiency when endowment heterogeneity exists, but only when the rich also have stronger capabilities.

Nockur et al. (2021) look at the impact of peer punishment, democratic punishment (where a group member is punished only if a majority of the 4-person group agrees) and centralised punishment in the presence of endowment heterogeneity, where such heterogeneity is always generated randomly thereby corresponding to our "luck" treatment. However, the centralised punishment in their study was neither deterrent nor non-deterrent, leaving participants indifferent between contributing to the public account and free-riding.[15] Nockur et al. find that the presence of endowment heterogeneity does not make much of a difference. All three punishment systems increase contributions over a no punishment benchmark although peer punishment leads to higher contribution than centralised punishment. However, when it came to earnings, the centralised punishment mechanism did at least as well as the control treatment with no punishment while earnings were lower with peer and democratic punishment. (This may be an artefact of the relatively low 1:2 cost benefit ratio of peer/democratic punishments in this study.)

Of more immediate relevance to us, Nockur et al. find that the amount invested in punishment was not significantly different when comparing those with equal endowments to those with unequal endowments or between the advantaged (rich) and the disadvantaged (poor) individuals. In our setting, we expect that a non-deterrent punishment of either 0% or 30% will not make a significant difference and, as a result, the poor will still contribute higher amounts to the public good compared to the rich. However, the presence of deterrent punishment is expected to have an effect on cooperative behaviour especially for the rich participants. Specifically, when the punishment is either 60% or 80%, cooperation rates are expected to increase and be similar between the rich and poor, potentially reaching the social optimum of full contributions.

[15] The MPCR from the public account is 0.4. But one token placed in one's private account is penalised by a 0.6 token fine. Assuming pure self-interest, this should render participants indifferent between contributing a token to the public account and placing it in the private account.

As noted earlier in our introductory remarks, punishment in this setting has a strong redistributive effect. To see this, consider the following example. Suppose there is no punishment for free-riding. The two poor participants contribute 20 tokens (their entire token endowment each) while the two rich participants match this 20 token contribution.[16] This implies that each rich participant is left with 60 tokens in his/her private account. There are 80 tokens in the public account which generates 40 tokens for each participant. The poor end up with 40 tokens each while the rich end up with 100 tokens each. This leads to an income inequality ratio of 2.5:1 in favour of the rich as opposed to the 4:1 ratio that existed prior to the game.

But now suppose this society decides to implement an 80% punishment rate on tokens not contributed to the group account. Now if each rich participant contributes 20 tokens and retains 60 tokens in his/her private account, each will be looking at a punishment of 48 tokens leaving each with 12 tokens out of 60. This implies that the poor participants will end the round with 40 tokens each while the rich end with 52 tokens, resulting in a much lower income inequality ratio of 1.3:1 in favour of the rich. Punishment then reduces the extent of payoff inequity between the rich and the poor and, more importantly, should induce the rich to contribute more in later rounds. This provides the poor strong incentives to vote for higher punishment rates since doing so results in a higher payoff for the poor and reduced inequity in payoffs. This leads to our first hypothesis.

Hypothesis 1: *Across all inequality treatments, the poor vote for higher punishment than do the rich.*

[16] Prior evidence as in Sugden (1984) or Croson (2007) suggests that the rich follow a reciprocity rule that approximately matches contributions by the poor. Typically, this implies that the rich may end up contributing more than the poor in absolute terms but much less in relative terms.

Our second hypothesis relies on some of the same arguments as the first. As mentioned above, previous studies documented increased predilection for greater free-riding on the part of the rich in the presence of endowment heterogeneity and in the absence of punishment. Low punishment rates (0% and 30%) are non-deterrent in that they provide no incentives to change contribution behaviour. However, deterrent punishment rates (60% and 80%) make it rational for all players to contribute their entire endowment to the public good.

Hypothesis 2: *(a) With zero or non-deterrent punishment, the rich engage in greater free-riding and contribute a lower proportion of their endowment to the public good than do the poor. (b) With a deterrent punishment, the poor and rich contribute similar proportions of their endowments to the public good. We expect these amounts to be close to the social optimum.*

Prior research suggests that when participants earn their endowments, they feel more attached to it (e.g., Hoffman et al., 1994; Cherry et al., 2002), and are, therefore, less inclined to engage in pro-social behaviour. This suggests that the rich, who earned a higher endowment in *Merit*, will free-ride more than those who have not earned it.

Hypothesis 3: *The rich contribute a lower proportion of their endowment to the public good when inequality results from Merit than from Luck.*

Finally, the literature on self-serving biases (Babcock and Loewenstein, 1997) or misattribution of cause (Weber et al., 2001) suggest that when it comes to the *Uncertain* treatment, where the source of endowment heterogeneity is unclear, both the rich and the poor will interpret this in a way that is favourable to their own self-image. In particular, the rich would like to believe they earned their higher endowment while the poor would like to attribute their lower endowment to luck. This leads to our final hypothesis. We note that we did not specifically ask participants whether they perceived their endowments as resulting from merit or luck in the uncertain treatment. Instead, we draw indirect inferences by looking at behavioural cues; whether behaviour of the rich and/or the poor in the uncertain treatment mirror those in the merit or the luck treatment.

Hypothesis 4: *Voting behaviour and contribution of the poor are similar in the Luck and Uncertain treatments; for the rich, behaviour is similar in the Merit and Uncertain treatments.*

4. Results

4.1 Testing Hypothesis 1: do the poor vote for higher punishment?

In what follows, we use Wilcoxon ranksum (RS) tests when making comparisons between treatments, and signed rank (SR) tests when making comparisons within treatments. Since group composition is fixed, each group is independent and therefore, we take a group as the unit of observation. For each group, we first calculate the average punishment rate voted for in the group in each round, and then average this over all 10 rounds. This provides us one independent observation per group. Means and standard deviations are calculated at the group level using this data. The number of observations in each treatment is the number of groups in that treatment. For each test, the number of observations is the number of groups in the two treatments being compared. We report p-values from two-sided tests.

Table 2 presents average punishment rates voted for by the rich and the poor in the different Punishment treatments. The table also presents p-values from SR tests for differences between punishment rates voted for by the rich and poor within each treatment. In all three inequality treatments, the poor vote for higher punishment on average than do the rich. This is consistent with a desire to ensure more equitable earnings among group members. However, this difference is statistically significant only in *Luck-Pun*. That is, there is no significant difference in average voting behaviour between the poor and the rich in *Merit-Pun* and *Uncertain-Pun*.

Table 2. Summary statistics of individual votes for punishment rates

	Obs.	Endowment 20 (Poor)	50	80 (Rich)	SR p-value
Equality-Pun	12	-	32.73 (15.87)	-	-
Luck-Pun	*17*	*41.88* *(19.49)*	-	*24.76* *(13.86)*	*0.0056*
Merit-Pun	12	39.17 (25.19)	-	28.17 (19.94)	0.2393
Uncertain-Pun	12	33.63 (21.99)	-	29.33 (17.59)	0.4556

Figures in parentheses are standard deviations.

In Table 3, we corroborate this finding about differences in voting by running individual-level probit regressions for votes. In each regression, the dependent variable is 1 if the vote of the player in a round is equal to the vote specified at the top of the column (i.e., a vote for a punishment rate of 0%, 30%, 60% or 80%) and = 0 otherwise. The regressors include three dummies for the three inequality treatments with *Equality* being the reference category, and then we interact the three dummies with a dummy for poor participants (with endowment of 20) to look for differential voting patterns between the poor and the rich. Table 3 presents marginal effects estimated after the regressions.

Table 3. Probit regressions of individual votes: Marginal effects

Individual Probit regressions	(1) Vote = 0	(2) Vote = 30	(3) Vote = 60	(4) Vote = 80
Luck-Pun	0.109	0.002	-0.041	-0.074
	(0.089)	(0.063)	(0.030)	(0.076)
Merit-Pun	0.121	-0.083	-0.035	0.007
	(0.117)	(0.070)	(0.035)	(0.084)
Uncertain-Pun	0.042	0.006	-0.009	-0.039
	(0.102)	(0.079)	(0.034)	(0.086)
Luck × (End = 20)	-0.164^{**}	-0.078	0.037	0.246^{***}
	(0.209)	(0.046)	(0.029)	(0.087)
Merit × (End = 20)	-0.123	≈0.000	-0.040	0.163
	(0.095)	(0.107)	(0.034)	(0.109)
Uncertain × (End = 20)	-0.054	0.004	-0.007	0.066
	(0.094)	(0.074)	(0.033)	(0.067)
Observations	2120	2120	2120	2120

Excluded treatment: Equality. Excluded endowment: 50. The endowment of 80 is captured by the treatment dummy in the inequality treatments. Marginal effects from probit regressions. Std. errors clustered on group in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

It is clear from Table 3 that there is a strong difference in the voting pattern for the poor across the different treatments. In the *Luck-Pun* treatment, the poor voted disproportionately more for the highest 80% punishment and disproportionately against having no punishment at all. The marginal effects suggest that compared to the Equal-Pun treatment (where endowments are equal), the poor in Luck-Pun are 17% (= 0.246 – 0.074) more likely to vote for the 80% fine and 6% (= -0.164 + 0.109) less likely to vote for no punishment. Both differences are significant at the 5% or better. This is despite the fact that costs for the poor, if punishment is implemented, are 20% of their token endowment. The voting patterns are not significantly different in the other treatments. We take this as partial corroboration of Hypothesis 1 that the poor will vote for higher punishment. We find this to be true only in the *Luck-Pun* treatment.

Result 1: *The poor vote for higher punishment than do the rich when inequality derives from Luck, but there are no differences in the voting patterns between the poor and the rich in the Merit and Uncertain treatments.*

4.2 Testing Hypothesis 2: differences in disaggregated contributions

In this section, we talk about both punishments and contributions in conjunction since contributions depend crucially on the presence or absence of punishment and, if present, on its magnitude. First presented are average group contributions to the public good and earnings with and without punishment. Total group contributions/earnings averaged across all four members in each round and over all 10 rounds for each group yields one independent observation per group. Group earnings in the *Punishment* treatments account for the cost of punishment. Note that since total resources available to a group are the same (200 tokens) in all groups, total (absolute) group contributions and earnings are directly comparable across treatments. These averages are shown in Table 4.[17]

[17] All Ranksum tests comparing average group round earnings within and across *No-Pun* and *Pun* treatments are insignificant ($p \geq 0.20$ in all cases). While aggregate earnings do not differ across treatments, the distribution of earnings do differ (see Table 8).

Table 4. Average per-round group contributions and earnings

	No-Punishment			*Punishment*		
	Obs.	Contributions	Earnings	Obs.	Contributions	Earnings
Equality	16	113.88 (35.50) [56.9%]	314.80 (35.33)	12	142.08 (32.17) [71%]	323.81 (36.33)
Luck	19	108.93 (38.39) [54.5%]	309.71 (38.32)	17	127.14 (32.58) [63.6%]	309.38 (34.14)
Merit	17	99.18 (29.21) [49.6%]	299.95 (29.13)	12	131.78 (32.31) [65.9%]	314.82 (36.51)
Uncertain	18	109.05 (23.71) [54.5%]	309.86 (23.69)	12	137.91 (40.08) [68.9%]	321.33 (39.27)

Figures in parentheses are standard deviations. Figures in brackets are percentage contributions.

In the absence of punishment opportunities, average group contributions (over all 10 rounds) are around 50% of group endowment in *Merit*, and around 55% of endowment in the other three treatments.[18] Average group contributions are higher in the presence of punishment opportunities across the board; contributions are around 65% of endowment in *Luck-Pun* and *Merit-Pun*, and around 70% in *Equality-Pun* and *Uncertain-Pun*. Pairwise tests show that punishment leads to significantly higher average group contributions in *Equality* (RS p = 0.04), and when the source of inequality is *Merit* (RS p = 0.01) or *Uncertain* (RS p = 0.02). However, with *Luck*, punishment does not significantly raise group contributions (RS p = 0.16).

In order to understand why punishment failed to increase contributions in *Luck-Pun*, we look at disaggregated behaviour in Table 5. The Table shows percentage (of endowment)

[18] Our results suggest that in the absence of punishments, average contributions are not significantly different between groups with equal or unequal endowments. While a number of prior studies such as Anderson et al. (2008), Cherry et al. (2005), Hargreaves Heap et al. (2016) and Ramalingam and Stoddard (2023) find significantly lower group contributions in unequally endowed groups compared to equally endowed groups, Nockur et al. (2021) also do not find significant differences between equal and unequal groups. We do not elaborate on or explore this further since there is a substantial literature in this area and this is not the focus of the current study.

contributions by group members in *Equality* and the poor and rich in the inequality treatments. This table shows that the presence of punishments leads to increased contributions by both the rich and the poor, at least at 6% level, in the *Merit* and the *Uncertain* treatments, but not in the *Luck* treatment. The relative ineffectiveness of punishment in *Luck-Pun* is driven by a lack of response by both the rich and the poor. In *Merit-Pun* and *Uncertain-Pun*, both respond by increasing their contributions.

Table 5. Comparing percentage contributions of rich and poor in the presence/absence of punishments

Treatments and endowments	*No-Punishment*	*Punishment*	RS p-values
Equality	56.94 (17.75)	71.04 (16.08)	0.04
Poor in *Luck*	69.71 (18.94)	70.82 (16.26)	0.95
Rich in *Luck*	50.65 (21.91)	61.75 (18.53)	0.12
Poor in *Merit*	69.71 (17.29)	81.63 (12.38)	0.06
Rich in *Merit*	44.56 (17.81)	61.95 (19.68)	0.03
Poor in *Uncertain*	70.21 (13.10)	83.50 (12.60)	0.01
Rich in *Uncertain*	50.60 (15.11)	65.32 (23.33)	0.03

Figures in parentheses are standard deviations.

Table 5 reveals that, in the absence of punishments, the rich contribute more in absolute amounts, but proportionally less than the poor (SR p = 0.01 in all cases). This corroborates prior findings (e.g., Buckley and Croson, 2006; Hargreaves Heap et al., 2016). But, as Table 5 shows, this proclivity to free-ride in relative terms on the part of the rich persists even in the presence of punishments. Wilcoxon signed rank tests (SR) support these findings. The only case where this is not true is in *Luck-Pun* (SR p = 0.09), where the percentage contribution of the poor is the lowest of the three inequality treatments, approximately 71% in *Luck-Pun* as opposed to 82% in *Merit-Pun* and 84% in *Uncertain-Pun*.[19]

Table 5 suggests that the presence of punishments *per se* does not eliminate the gap in percentage contributions between the poor and the rich. However, in Tables 6A and 6B we look at whether the greater free-riding on the part of the rich is still true in the presence of deterrent versus non-deterrent punishment rates. The tables also present p-values from SR tests comparing contributions of the poor and the rich within each treatment. Not surprisingly, Table 6A shows that the rich still free-ride at higher levels than the poor, as a proportion of their endowment, when there is no punishment or the punishment is non-deterrent (0% or 30%), except in *Luck-Pun*.[20] However, as we can see from Table 6B, when the punishment is deterrent there are no statistically significant differences between the percentage contributions of the poor and the rich in any of the three treatments. Moreover, across the board, percentage contributions are 'high' and close to optimal levels.

Table 6A. Average percentage contributions in the presence of punishment: Non-deterrent punishment rates

		Endowment			
	Obs.	20 (Poor)	50	80 (Rich)	SR p-value
Equality-Pun	12	-	65.01 (14.35)	-	-
Luck-Pun	17	62.20 (20.28)	-	52.08 (17.51)	0.1298
Merit-Pun	12	77.51 (17.32)	-	52.13 (19.43)	0.0047
Uncertain-Pun	11	82.12 (14.20)	-	55.91 (23.47)	0.0099

Figures in parentheses are standard deviations. One group in *Uncertain-Pun* never voted for non-deterrent or zero punishment.

[19] The % contribution of the poor is at least weakly significantly lower in *Luck-Pun* than in *Merit-Pun* and *Uncertain-Pun* (RS p = 0.07 & 0.03, respectively).

[20] The % contribution of the poor with no or non-deterrent punishment is significantly lower in *Luck-Pun* than in *Merit-Pun* and *Uncertain-Pun* (RS p = 0.04 & 0.01, respectively).

Table 6B. Average percentage contributions in the presence of punishment: Deterrent punishment rates

	Obs.	Endowment			SR p-value
		20 (Poor)	50	80 (Rich)	
Equality-Pun	8	-	84.72	-	-
			(14.15)		
Luck-Pun	13	87.54	-	86.31	0.7267
		(12.83)		(14.15)	
Merit-Pun	8	87.07	-	84.56	0.8886
		(11.52)		(16.11)	
Uncertain-Pun	7	91.65	-	85.79	0.4990
		(11.96)		(19.23)	

Figures in parentheses are standard deviations. Four groups in each of *Equality-Pun*, *Luck-Pun*, and *Merit-Pun*, as well as five groups in *Uncertain-Pun*, never voted for deterrent punishment.

Result 2: *(a) In the presence of non-deterrent punishment, compared to the rich, the poor contribute significantly higher proportions of their endowment in Merit-Pun and Uncertain-Pun, but their contributions are not significantly different in Luck-Pun. (b) In the presence of deterrent punishment, the poor and rich contribute similar proportions in all inequality treatments.*

In Table 7, we present individual-level panel random effects regressions of individual contributions in a round on a time trend, one-round lagged percentage contributions of other group members, a dummy for a deterrent punishment rate, a dummy for rich group members, and an interaction between the two dummies. For expositional ease, we present regressions for each punishment treatment separately, and report standard errors clustered on independent groups. The full set of results using the entire data is available in Appendix B.[21]

[21] In Appendix B of the Supplementary Material, we present regressions with all treatments combined. See Table B.1.

Table 7. Panel regressions of individual contributions: the effects of punishment

	Punishment treatments			
	Equality	*Luck*	*Merit*	*Uncertain*
Round	-0.502	-1.183	-1.235*	-0.465
	(0.808)	(0.721)	(0.720)	(0.552)
Lagged percentage cont. of others in group	-0.043	0.025	0.115	0.146***
	(0.078)	(0.068)	(0.101)	(0.051)
Deterrent punishment dummy	25.43***	31.67***	13.65***	8.745
	(7.129)	(5.454)	(4.868)	(8.587)
End = 80 dummy	-	-10.96*	-26.98***	-28.13***
		(6.345)	(6.955)	(8.186)
Det. Punishment × End = 80	-	4.870	20.50***	27.40**
		(7.942)	(6.203)	(12.39)
Constant	68.60***	65.93***	77.37***	73.46***
	(8.178)	(7.374)	(8.061)	(8.159)
Observations	432	612	432	432

Panel RE regressions. Std. errors clustered on group in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

This table provides additional support for the differences in *Luck-Pun* compared to *Merit-Pun* and *Uncertain-Pun*. It is clear that the rich contribute a smaller percentage of their endowments to the public good in all three inequality treatments, except in *Luck-Pun* where the End=80 dummy is only marginally significant. This suggests that the poor and rich contribute similar proportions of their endowments in *Luck-Pun*. Further, the presence of deterrent punishment raises contributions of the rich in *Merit-Pun* and *Uncertain-Pun*, but not in *Luck-Pun* – the interaction term is only significantly positive in the former two treatments. The regressions show that deterrent punishments are effective in raising (percent) contributions in all treatments except *Uncertain-Pun*, where lagged contributions of other group members play a bigger role than in the other treatments. This suggests that conditional cooperation (Sugden, 1984; Fischbacher et al., 2001) may play a more important role in this treatment than in the other treatments.

The fact that the past contributions of others are not significant in the other three treatments (Equality, Luck and Merit) may have to do with the following. Own contributions are typically positively correlated with lagged contribution of others. In the absence of punishment, this dependence can be explained via conditional cooperation. With punishment, evidence shows that those who contribute less than others (or the average) tend to attract greater punishment, in which case it makes sense to pay attention to what others are doing. But in our study, the punishment is *absolute* and independent of what others are doing. So, there is less need to condition one's own contribution on that of others. But in the *Uncertain-Pun* treatment, since no one is certain of the circumstances surrounding endowments, it makes sense to look to others for cues as to what might be 'appropriate' behaviour.

Collectively, the evidence here suggests that in the absence of punishment or with a non-deterrent punishment, the rich do free-ride more than the poor except when inequality obtains from luck. But the reason that there are no differences in the case of *Luck-Pun* is not that the rich increase their percentage contribution to match those of the poor, but that the poor reduce their proportional contributions to match those of the rich. In the other treatments, the poor contribute proportionally higher amounts.

4.3 Testing Hypotheses 3 and 4: comparing contributions of the rich and poor under the different sources of inequality

Consistent with Hypothesis 3, in the absence of punishment, average percentage contributions of the rich are lower under *Merit* than in *Luck* (44.56 vs. 50.65; Table 5). However, this difference is not statistically significant (RS $p > 0.10$). There is hardly any difference in the percentage contributions of the rich across treatments in the presence of non-deterrent or deterrent punishment (Tables 6A and 6B) and combined (Table 5). Thus, we do not find support for the hypothesis that the rich free-ride more when inequality stems from merit.

Result 3: *The proportional contributions of the rich do not differ significantly across the different sources of inequality.*

Hypothesis 4 argues that the poor will interpret the *Luck* and *Uncertain* treatments in similar ways (arising from random chance) while the rich will do the same for the *Merit* and *Uncertain* treatments (caused by merit). This implies that we would expect the poor to vote for higher punishment rates in *Luck-Pun* and *Uncertain-Pun* than in *Merit-Pun*. We also expect the poor to engage in greater free-riding in those two treatments. Similar arguments should apply to the rich. We would expect greater free-riding on the part of the rich in *Merit-Pun* and *Uncertain-Pun*.

There is little support for this hypothesis. The evidence presented in Tables 2 and 3 above show that the poor voted for higher punishment rates only in the *Luck-Pun* treatment. Their voting behaviour is indistinguishable in *Merit-Pun* and *Uncertain-Pun*. Similarly, as made clear by Table 5, the degree of free-riding by the rich is not significantly different based on the source of inequality; the rich contribute a smaller proportion of their endowment to the public account regardless of the source of that endowment. This only changes in the presence of deterrent punishment. Then the contributions between rich and poor are not significantly different and they are uniformly high, though not necessarily at the social optimum (consistent with Hypothesis 3).

In *Luck-Pun*, there is relatively higher free-riding on the part of the poor compared to the poor in other inequality treatments. More importantly, the poor disproportionately vote for deterrent punishment in that treatment.[22] Such behaviour has implications for the evolution of

[22] While the poor were more likely to vote for deterrent punishment than the rich were in *Luck-Pun*, it did not necessarily lead to more deterrent punishment being implemented. The percentage of rounds deterrent punishment

inequality over time. Table 8 presents average final (accumulated) earnings of the poor and the rich in all inequality treatments. The Table also shows the relative inequality in final earnings across the different treatments without and with punishments. This is done by calculating the ratio of final accumulated earnings of the rich to those of the poor within each group, giving one observation per group. Then we calculate the average of the ratios across all groups.

Table 8 shows that it is indeed the case that when punishment is available, relative inequality diminishes. This reduction in inequality is brought about by two factors. First, the redistributive nature of the public goods game itself. Second, the presence of punishment induces both the rich and the poor to increase their contributions to the public good (see Table 4), resulting in a further reduction in the inequity. Recall that at the outset each rich player has 80 tokens and each poor player has 20 tokens, with an inequality factor of 4. As Table 8, shows the relative inequalities decrease to a factor of around 1.6 to 1.7 in the absence of punishment and further to around 1.3 to 1.4 with punishment.

Table 8. Mean (ratio of) final earnings in the inequality treatments

	No-Punishment				*Punishment*				RS
	Obs.	End=20	End=80	Ratio	Obs.	End=20	End=80	Ratio	p-value
Luck	19	607.16	941.37	*1.679*	17	666.44	880.44	*1.373*	0.07
		(175.16)	(37.69)	*(0.500)*		(141.14)	(60.85)	*(0.289)*	
Merit	17	558.41	941.35	*1.782*	12	672.08	902.00	*1.407*	0.01
		(142.43)	(33.91)	*(0.416)*		(144.87)	(84.96)	*(0.382)*	
Uncertain	18	606.86	942.44	*1.607*	12	697.46	909.21	*1.422*	0.03
		(120.96)	(25.62)	*(0.305)*		(177.30)	(55.65)	*(0.548)*	

Figures in parentheses are standard deviations. Ratio is the ratio of average final earnings of rich group members to average final earnings of poor group members in a group. Averages of group-level ratios are reported.

The last column of Table 8 presents p-values from RS tests comparing the ratios between the two treatments (*No Punishment* and *Punishment*) within each source of inequality.

was implemented by treatment: *Equality-Pun* - 27.5%; *Luck-Pun* - 29.4%; *Merit-Pun* - 30.8%; *Uncertain-Pun* - 29.2%.

The opportunity to impose punishment significantly reduces inequality in final earnings between the poor and the rich when the source of inequality is *Merit* or *Uncertain*, but only marginally so when inequality is a result of *Luck.* Once again, outcomes are different when inequality derives from luck as opposed to the other two sources. Group contributions in *Luck* do not significantly increase in the presence of punishment, unlike the other two treatments (Table 4). So, the lack of a significant increase in group contributions limits the extent of that redistribution. The end result is that the relative earnings of the poor do not increase (due to punishment) as much in the *Luck* treatment as in the other inequality treatments.

The preponderance of evidence suggests that both the rich and the poor treated *Merit* and *Uncertain* as being analogous. This was expected for the rich, but goes against our ex-ante conjecture that the poor will treat *Uncertain* as being analogous to *Luck*. We find partial support for Hypothesis 4.

Result 4: *The behaviour of both the poor and rich are similar in Merit and Uncertain.*

Much prior research suggests that luck-based inequality is less acceptable than merit-based inequality (e.g., Starmans et al. 2017; Almas et al., 2020). The observed behaviour of the poor in our experiment is in line with an aversion to luck-based inequality. Our results suggest that the poor perceive luck-based inequality as more 'unfair' and, more importantly, are more willing to act to reduce such unfair inequality.

We find that the poor in *Luck-Pun* contribute a similar percentage of endowment to the public good as do the rich; punishment raises the contributions of the rich but *not* those of the poor. This is not observed in the other inequality treatments. Moreover, the fact that the poor choose deterrent punishment rates more often in *Luck-Pun* than in the other inequality treatments suggests a stronger desire to equalise earnings across the poor and the rich when inequality results from luck than from other sources.

As mentioned in Section 2.3, all participants were asked to report their perceived level of fairness (on a 5-point Likert type scale: Scale: 1 = Very unfair, to 5 = Very fair) of the source of their endowments. While an admittedly crude proxy for a true measure of fairness perceptions[23], in Appendix E we examine if our self-reported fairness measure potentially correlates with observed behaviour in the prior public goods game in the different treatments. While not conclusive, the fairness-ratings findings suggest that observed behaviour in *Luck-Pun* stems from a perception of unfairness of the (source of) inequality compared to *Merit-Pun*, especially by the poor.

5. Conclusion

We report data from an online repeated public goods experiment that varies the source of inequality and punishment opportunities. With our experimental manipulations, we provide causal evidence that people are sensitive to the sources of inequality, especially luck-based inequality. When people are randomly allocated to groups with luck-based inequality, they free-ride more and vote for harsher punishment systems compared to people randomly allocated to equal groups or groups with merit-based inequality. These results provide causal support for previous correlational evidence showing that people care about the sources of inequality (Almas et al. 2019, Starmans et al. 2017).

Our results point to two main insights into behaviour in unequal groups. First, the behaviour of the poor suggests they respond more forcefully when inequality is the result of luck than the result of merit. Their proportional contribution is lower in *Luck-Pun* than in both

[23] Our participants reported their perceived fairness at the end of the experiment. Hence, their perceptions were undoubtedly coloured by their experience in the public goods game. Respondents in earlier studies, who find that luck is a less acceptable than merit as a source of inequality, were made aware of the different sources of inequality and then they ranked the sources. This was not the case in our experiment. Participants were only aware of how endowments were determined in their own treatment, and were not aware of the other endowment scenarios. Since they were unable to rank scenarios in terms of fairness, they had to enter a rating with little context. These features make our measure a proxy at best.

Merit-Pun and *Uncertain-Pun*. In fact, *Luck-Pun* is the only treatment where there is no significant difference between the proportional contributions of the rich and the poor. This lack of difference in *Luck-Pun* does not arise from the rich free-riding less but the poor free-riding more. The poor also voted for higher punishment in *Luck-Pun*. This is expected given the redistributive nature of punishments. What is striking is that the poor did not do so across the board, i.e., in *Merit-Pun* or in *Uncertain-Pun*; the intent to redistribute is higher under *Luck*.

A second insight is that participants appear to view the *Uncertain* treatment as being on par with the *Merit* treatment. In retrospect, it is possible that self-serving bias and/or misattribution of cause are certainly factors in decision making, but not in the way we initially conceived. Misattribution of cause seems to work primarily via convincing participants that the *Uncertain* treatment, which has an element of merit to it, is effectively similar to *Merit*. Wilson (2003, p. 300) writes that it is possible to conflate real and apparent merit by arguing as follows:

> *As long as A_1 is believed to work harder than A_2 and/or is perceived as having more talent …whether or not she actually does, A_1 deserves to enjoy a higher level of well-being than A_2. It would be unjust if A_2 were to obtain as much as or more than A_1.*

While Wilson argues against this as being a defensible ethical standard, it is conceivable that a similar standard may have operated in the minds of our study participants. Wilson (p. 302) goes on to argue that "*Beliefs regarding performance are adjusted in order to satisfy the preference that good things happen to people judged to be good overall.*"

It is clear that most people have intrinsic notions of justice and at least by some standards, merit is considered more "just" than luck. It appears that the *Uncertain* treatment, that appears to encompass an element of merit, is also perceived to be more just than luck. See Konow (1996, 2000) for further arguments along similar lines.

Finally, Cappelen et al. (2007) argue for three distinct fairness ideals: *egalitarianism*, that argues for equalising resources; *libertarianism*, that argues for allocating resources on the basis of merit and *liberal egalitarianism*, that falls in between the two. Cappelen et al. find that 35% of their participants adhere to the libertarian ideal. For such participants, some of whom may be poor, it is clear that a merit-based system is preferred and given the arguments above regarding conflation of *Merit* and *Uncertain*, it may not be surprising that behaviour is similar in the two treatments.

Recent years have seen a sustained debate on the question of inequality; particularly since the publication of Piketty (2014). As Piketty points out, over the last three decades the industrialised West experienced levels of inequality not seen since the early years of the 20^{th} century. There is also evidence that people consider inequality based on merit more acceptable than inequality due to luck. We provide causal evidence that behaviour in our setting is consistent with the above. Our findings have implications for our perceptions about the causes of inequality and therefore for public policy, particularly in the presence or absence of social mobility and access to means that may help reduce inequality. As the rich do not vary behaviour across settings with different sources of inequality, the poor would need a stronger voice to implement change (such as social mobility and redistribution). However, our results suggest that, even with a stronger voice, poor would only implement change in settings where luck is clearly identified as the source of inequality. How do the poor respond in other settings when the source of inequality and the strength of their voice are varied? This question is left for future research.

References

Alesina, A. and Angeletos, G-M. (2005). Fairness and redistribution. American Economic Review, 95(4), 960-980.

Alesina, A., Di Tella, R. and MacCulloch, R. (2004). Inequality and happiness: Are Europeans and Americans different? Journal of Public Economics, 88(9-10), 2009-2042.

Almås, I., Cappelen, A.W., and Tungodden, B. (2020). Cutthroat capitalism versus cuddly socialism: Are Americans more meritocratic and efficiency-seeking than Scandinavians? Journal of Political Economy, 128 (5), 1753-1788.

Almås, I., Cappelen, A.W., Sørensen, E., and Tungodden, B. (2010). Fairness and the development of Inequality Acceptance. Science, 328 (5982), 1176-1178.

Almås, I., Cappelen, A., Sørensen, E., and Tunggoden, B. (2019, November 25-26). Fairness across the world: Preferences and beliefs [Keynote presentation]. Australia New Zealand Workshop in Experimental Economics, Monash University, Clayton, Victoria, Australia.

Altemeyer, B. (1981). Right-wing Authoritarianism. Winnipeg, Manitoba. University of Manitoba Press.

Ambrus, A., and Greiner, B. (2019). Individual, dictator, and democratic punishment in public good games with perfect and imperfect observability. Journal of Public Economics, 178, 1–23.

Anderson, L.R., Mellor, J.M., and Milyo, J. (2008). Inequality and public good provision: An experimental analysis. Journal of Socio-Economics, 37(3), 1010-1028.

Arechar, A.A., Gächter, S., and Molleman, L. (2018). Conducting interactive experiments online. Experimental Economics, 21(1), 99-131.

Babcock, L., and Loewenstein, G. (1997). Explaining Bargaining Impasse: The Role of Self-Serving Biases. Journal of Economic Perspectives, 11(1), 109-126.

Baldassarri, D., and Grossman, G. (2011). Centralized sanctioning and legitimate authority promote cooperation in humans. Proceedings of the National Academy of Sciences, 108 (27), 11023–11027.

Buckley, E., and Croson, R. (2006). Income and wealth heterogeneity in the voluntary provision of linear public goods. Journal of Public Economics, 90, 935-955.

Cappelen, A.W., Hole, A.D., Sorensen, E. O., and Tungodden, B. (1997). The Pluralism of Fairness Ideals: An Experimental Approach. American Economic Review, 97(3), 818-827.

Chaudhuri, A. (2011). Sustaining cooperation in laboratory public goods experiments: a selective survey of the literature. Experimental Economics, 14(1), 47-83.

Chaudhuri, A. (2021). Behavioural Economics and Experiments. London and New York: Routledge.

Chen, D.L., Schonger, M., and Wickens, C. (2016). oTree – An open-source platform for laboratory, online and field experiments. Journal of Behavioral and Experimental Finance, 9, 88-97.

Cherry, T.L., Frykblom, P., and Shogren, J.F. (2002). Hardnose the dictator. American Economic Review, 92(4), 1218-1221.

Cherry, T.L., Kroll, S., and Shogren, J.F. (2005). The effect of endowment heterogeneity and origin on public good contributions: evidence from the lab. Journal of Economic Behavior and Organization, 57 (3), 356-365.

Claessens, S., Fischer, K., Chaudhuri, A., Sibley, C., and Atkinson. Q. (2020). Dual Evolutionary Foundations of Political Ideology. Nature Human Behaviour, 4, 336–345.

Clark, A. and D'Ambrosio, C. (2015). Attitudes to income inequality: Experimental and Survey Evidence. Chapter 13, pages 1147-1208 in Atkinson, A. and Bourguignon, F. (Eds.) Handbook of Income Distribution, Volume 2A. Amsterdam: Elsevier.

Croson, R. (2007). Theories of Commitment, Altruism and Reciprocity: Evidence From Linear Public Goods Games. Economic Inquiry, 45(2), 199-216.

Egas, M., and A. Riedl, (2008). The economics of altruistic punishment and the maintenance of cooperation. Proceedings of the Royal Society B, 275(1637), 871-878.

Erkal, N., Gangadharan, L., and Nikiforakis, N. (2011). Relative earnings and giving in a real-effort experiment. American Economic Review, 101(7), 3330-3348.

Falk, A., and Heckman, J.J. (2009). Lab Experiments are a Major Source of Knowledge in the Social Sciences. Science, 326(5952), 535-538.

Fehr, E., and Gächter, S. (2000). Cooperation and Punishment in Public Goods Experiments. American Economic Review, 90(4), 980-994.

Hargreaves Heap, S., Ramalingam, A., and Stoddard, B.V. (2016). Endowment inequality in public goods games: A re-examination. Economics Letters, September 2016, 4-7.

Hoffman, E., McCabe, K., Shachat, K., and Smith, V. (1994). Preferences, property rights, and anonymity in bargaining games. Games and Economic Behavior, 7(3), 346-380.

Kingsley, D. (2016). Endowment heterogeneity and peer punishment in a public good experiment: Cooperation and normative conflict. Journal of Behavioral and Experimental Economics, 60, 49-61.

Konow, J. (1996). A positive theory of economic fairness. Journal of Economic Behavior & Organization, 31 (1), 13-35.

Konow, J. (2000). Fair Shares: Accountability and Cognitive Dissonance in Allocation Decisions. American Economic Review, 90(4), 1072-1091.

Kosfeld, M., Okada, A., and Riedl, A. (2009). Institution Formation in Public Goods Games. American Economic Review, 99(4), 1335-1355.

Ledyard, J. (1995). Public goods: A survey of experimental results. Chapter 2, pages 111-194 in Kagel, J. & Roth, A. (Eds.), Handbook of Experimental Economics. Princeton, NJ: Princeton University Press.

Macchia, L., Plagnol, A.C., and Powdthavee, N. (2020). Buying happiness in an unequal world: Rank of income more strongly predicts well-being in more unequal countries. Personality and Social Psychology Bulletin, 46(5), 769-780.

Nikiforakis, N., and Normann, H-T. (2008). A comparative statics analysis of punishment in public-good experiments. Experimental Economics, 11, 358-269.

Nockur, L., Pfatteicher, S., and Keller, J. (2021). Different punishment systems in a public goods game with asymmetric endowments. Journal of Experimental Psychology, 93, 104096.

Norton, M. I., and Ariely, D. (2011). Building a better America—One wealth quintile at a time. Perspectives on Psychological Science, 6(1), 9-12.

Oxoby, R.J., and Spraggon, J. (2013). A Clear and Present Minority: Heterogeneity in the Source of Endowments and the Provision of Public Goods. Economic Inquiry, 51(4), 2071-2082.

Pfattheicher, S., Bohm, R., and Kesberg, R. (2018). The advantage of democratic peer punishment in sustaining cooperation within groups. Journal of Behavioral Decision Making, 31(4), 562–571.

Piketty, T. (2014). Capital in the 21st Century. Cambridge. MA: Harvard University Press.

Putterman, L., Tyran, J-R., and Kamei, K. (2011). Public goods and voting on formal sanction schemes. Journal of Public Economics, 95(9-10), 1231-1222.

Ramalingam, A., and Stoddard, B. (2023). Does reducing inequality increase cooperation? Working Paper. Boone, NC: Appalachian State University

Sidanius, J., and Pratto, F. (1999). Social Dominance: An Intergroup Theory of Social Hierarchy and Oppression. Cambridge: Cambridge University Press.

Smith, V.L. (1976). Experimental Economics: Induced Value Theory. American Economic Review (American Economic Association Papers and Proceedings), 66(2), 274-279.

Starmans, C., Sheskin, M., and Bloom, P. (2017). Why people prefer unequal societies. Nature Human Behaviour, 1(4), Article 0082.

Stiglitz, J. (2020). Conquering the Great Divide. International Monetary Fund Finance and Development. Fall 2020. https://www.imf.org/external/pubs/ft/fandd/2020/09/COVID19-and-global-inequality-joseph-stiglitz.htm.

Sugden, R. (1984). Reciprocity: The supply of public goods through voluntary contributions. Economic Journal, 94(376), 772-787.

Waichman, I. (2020). Linking wealth and punishment effectiveness: Punishment and cooperation under congruent heterogeneities. Economic Inquiry, 58(1), 86-103.

Weber, R., Camerer, C. Rottenstreich, Y., and Knez, M. (2001). The Illusion of Leadership: Misattribution of Cause in Coordination Games. Organization Science, 12 (5), 582-598.

Wilson, C. (2003). The Role of a Merit Principle in Distributive Justice. The Journal of Ethics, 7(3), 277-314.

Zelmer, J. (2003). Linear Public Goods Experiments: A Meta-Analysis. Experimental Economics, 6, 299–310.