Inequality and Norm Enforcement in Public Good Games

David C. Kingsley *[1]

[1]Department of Economics, University of Massachusetts Lowell

Abstract

Income inequality is often thought to undermine cooperation, yet the mechanisms through which it does so remain poorly understood. This paper studies how inequality affects coordination around and enforcement of contribution norms in an interior public good game. Unlike standard linear public good games, the interior design features multiple efficient contribution profiles, creating ambiguity over which contribution norm should be enforced. We compare an equal-endowment treatment to two unequal-endowment treatments that hold aggregate group resources constant and introduce peer punishment after a baseline phase. Peer punishment increases contributions in all treatments, but the increase is significantly smaller under inequality. In the punishment phase, unequal-endowment groups contribute and earn significantly less than equal-endowment groups. Estimated contribution norms indicate that punishment decisions in all treatments are organized around the enforcement of proportional contributions. Inequality, however, substantially weakens both the likelihood and intensity of punishment directed toward deviations from this benchmark. The results suggest that inequality undermines decentralized cooperation not by impairing coordination around a contribution norm, but by weakening the enforcement of that norm.

Keywords: Income Inequality; Cooperation; Norm Enforcement; Punishment; Public Goods
JEL Codes: C92, H41, D63

*Funding from the University of Massachusetts Lowell is gratefully acknowledged. Declarations of interest: none.

1 Introduction

Communities routinely face collective-action problems in which individuals must decide how much to sacrifice for a shared benefit. Residents may need to finance a flood barrier, maintain a neighborhood park, or contribute to local security. In each case, everyone benefits if the good is provided, but each individual would prefer others to bear the cost. The standard public goods problem is therefore one of free riding: how can groups induce individuals to contribute when private incentives point toward under-provision?

Inequality makes this problem more complex. When group members differ in their ability to contribute, cooperation requires more than simply overcoming selfishness. It also requires agreement over what cooperation means. A high-income household may view equal contributions as fair because the flood barrier or park is shared by all. A low-income household may instead view equal contributions as unfair because they impose a larger relative sacrifice on those with fewer resources. From this perspective, contributions should rise with income, or perhaps be chosen so that members are left with more equal final payoffs. These competing intuitions reflect a broader point from the distributive preference literature: people are not averse to inequality itself, but distinguish between unequal outcomes they view as fair and unfair (Konow, 2000; Cappelen et al., 2007; Almås et al., 2020). Thus, inequality can transform a public goods problem from a question of whether people cooperate into a question of which contribution norm they coordinate on and whether that norm can be effectively enforced. The issue is not only whether groups cooperate, but how the burden of cooperation is distributed.

Despite this added complexity, existing experimental work suggests that peer punishment can often sustain cooperation even when groups are unequal. In linear public good games, decentralized punishment has been shown to increase contributions in heterogeneous groups much as it does in homogeneous groups (Reuben and Riedl, 2013). At the same time, related evidence shows that heterogeneity can undermine the welfare effects of peer punishment when it creates conflicting normative expectations and counter-punishment (Nikiforakis et al., 2012). This evidence suggests that heterogeneity need not prevent decentralized enforcement of cooperative norms, but that its effectiveness depends on whether group members can identify and sanction deviations from a commonly understood benchmark.

Typical linear public good games have a boundary solution in which efficiency requires each member to contribute the entire endowment. Further, regardless of whether endowments are equal or unequal, full contribution equalizes final payoffs. The efficient benchmark is therefore unique, normatively appealing, and deviations are easy to identify (Cason and Gangadharan, 2015). As a result, standard linear public goods games limit the scope for inequality to matter through disagreement over which contribution behavior is fair.

To investigate this possibility, the experiment uses a piecewise-linear public good with an interior social optimum. When efficiency requires an interior aggregate contribution, groups can achieve efficient provision through multiple individual contribution profiles rather than a single full-contribution benchmark. With unequal endowments, these profiles divide the burden of provision differently across members: equal contributions, proportional contributions, and earnings-equalizing contributions are both efficient and normatively appealing. This design allows inequality to generate distributive ambiguity that is largely absent from standard linear public good games. The central question is whether this ambiguity limits the ability of groups to coordinate on a contribution norm, or instead weakens punishment's ability to enforce a broadly shared benchmark.

The experiment implements three treatments: one with equal endowments (*Equal*) and two with unequal endowments (*Low Inequality* and *High Inequality*), while holding aggregate group resources constant across treatments. Groups remain fixed over time and participate in a voluntary-contribution phase followed by a punishment phase. The public good generates a constant marginal return up to a known aggregate contribution level; contributions beyond that point provide no additional return. This payoff structure creates an interior efficient aggregate contribution while preserving a simple contribution environment. The equal-endowment treatment shows whether groups can coordinate and use punishment effectively in the interior setting absent inequality, while the unequal treatments identify whether inequality lowers contributions, changes the norm groups appear to follow, or weakens punishment's enforcement of a chosen norm.

The results show that punishment increases contributions in all treatments, but significantly less so when endowments are unequal. While the introduction of punishment does not significantly affect earnings when endowments are equal, it significantly reduces earnings in both unequal treatments. As a result, within the punishment phase, both contributions and earnings are significantly lower in both unequal treatments relative to the equal treatment. Importantly, our norm identification exercise suggests that punishment behavior in each treatment is consistent with the enforcement of a proportional-contribution benchmark. However, inequality weakens punishment's ability to sustain this benchmark: deviations from the proportional benchmark are less likely to be punished and, conditional on punishment, receive weaker sanctions under inequality. Overall, the evidence suggests that inequality weakens decentralized cooperation not by changing the contribution norm around which punishment is organized, but by reducing the effectiveness of decentralized norm enforcement.

The paper contributes to the literature on inequality and cooperation by separating two channels through which inequality may affect decentralized cooperation: coordination on contribution norms and enforcement of those norms. The benchmark-identification exercise indicates that unequal groups continue to organize punishment around the same proportional contribution norm observed under equality. The principal effect of inequality instead lies in weaker decentralized enforcement of that benchmark.

This distinction speaks to broader theories of self-governance and collective action, which emphasize that successful cooperation depends not only on agreement over appropriate behavior but also on the capacity to monitor and enforce shared rules. Ostrom argues that decentralized institutions are most likely to succeed when rules are clear, compliance can be monitored, and sanctions are available to enforce agreed-upon obligations (Ostrom et al., 1992; Ostrom, 2000). Coleman (1990) likewise argues that cooperative norms are sustained when individuals are willing to incur private costs to sanction violations, while more recent work links inequality to weaker social capital and cooperative outcomes (Fehr et al., 2020). Our findings suggest that inequality need not prevent groups from coordinating around a common contribution norm, but it can weaken the decentralized enforcement of that norm. In doing so, the experiment identifies one potential behavioral mechanism through which inequality may weaken the informal institutions that sustain decentralized cooperation.

2 Background

A large experimental literature shows that peer punishment can sustain cooperation in linear public good games. In the absence of punishment, contributions typically begin around half of the endowment and decline over time (Ledyard, 1995). Introducing costly peer punishment often increases contributions in these environments (Fehr and Gächter, 2000, 2002; Chaudhuri, 2011). Whether these contribution gains translate into higher group earnings depends on whether punishment is sufficiently strong, targeted primarily at free riders, and sustained long enough for the cooperation gains to outweigh the costs of sanctioning (Cinyabugama et al., 2006; Bochet et al., 2006; Egas and Riedl, 2008; Nikiforakis and Normann, 2008; Gächter et al., 2008).

Similar patterns are observed when endowments are unequal. In the absence of punishment, contributions tend to decline over time, and several studies find that inequality lowers contributions primarily because high-endowment members contribute less (Zelmer, 2003; Buckley and Croson, 2006; Hargreaves Heap et al., 2016; Suchon and Théroude, 2022). Introducing punishment tends to increase contributions in linear public good games when endowments are unequal (Reuben and Riedl, 2013; Visser and Burns, 2015; Weng and Carlsson, 2015). In particular, Reuben and Riedl (2013) show that unequal-endowment groups can sustain cooperation by enforcing either an efficiency norm or a proportional-contribution norm.

These results suggest that, even when endowments differ, groups in linear public good games can organize punishment around salient contribution norms. A contribution norm reflects a shared expectation about what behavior is appropriate in a particular situation (Coleman, 1990; Elster, 1989; Bicchieri, 2006; Young, 2008; Fehr and Schurtenberger, 2018; Vostroknutov, 2020; De Geest and Kingsley, 2021). Empirical work on punishment norms similarly emphasizes that sanctions can be used to infer the contribution benchmarks that groups enforce (Carpenter and Matthews, 2009). Viewing punishment as a mechanism for coordinating on and enforcing contribution norms clarifies why it is most effective when group members share a common benchmark for appropriate behavior.

Effective enforcement also requires that deviations from the relevant norm are easy to observe and classify. Peer punishment becomes less effective when information is imperfect or incomplete. When endowments are equal but contributions are observed with noise, punishment becomes poorly targeted (Grechenig et al., 2010; Ambrus and Greiner, 2012). Free riders are insufficiently deterred, while cooperators are punished mistakenly, causing cooperation to unravel. When endowments are unequal but not observed, De Geest and Kingsley (2021) show that groups are unable to enforce an efficient, "contribute-your-endowment" benchmark. Groups instead adapt by enforcing a lower "contribute-the-low-endowment" benchmark, which protects cooperative low-endowment members from misguided punishment but allows high-endowment members to hide behind the lower standard. In both cases, punishment becomes less effective because sanctions can no longer be directed reliably toward deviations from the cooperative benchmark. Thus, the key issue is not simply whether a norm is normatively appealing, but whether behavior can be mapped clearly into compliance or violation.

The present experiment provides complete information about both contributions and endowments and therefore does not study informational ambiguity. Instead, it examines whether inequality affects groups' ability to coordinate on and enforce contribution norms when multiple efficient burden-sharing rules are available. This distinction highlights an important limitation of standard linear public good games. Because efficiency requires full contribution, the efficient benchmark remains unique and deviations are readily identified even when endowments differ. Consequently, inequality has relatively little scope to affect norm coordination or enforcement through competing efficient contribution rules.

In contrast, when the efficient outcome lies in the interior of the choice space, groups can maximize earnings under multiple contribution profiles. The distributive-preferences literature, much of it based on allocation decisions following a production stage, shows that people evaluate distributions using fairness criteria such as equality, proportionality, and accountability for differences in productive inputs (Konow, 2000; Frohlich et al., 2004; Cappelen et al., 2007). Applied to a public good environment, these criteria imply different ways of dividing the burden of cooperation across unequal endowment types (Kingsley and De Geest, 2025). The interior public good used here generates three contribution norms that are all efficient and normatively appealing: *Equal Contributions*, *Equal Proportions*, and *Equal Earnings*. Each may serve as a focal point for coordination, but each allocates the burden of provision differently across endowment types (Schelling, 1960; Young, 2008; Krupka and Weber, 2013). Under inequality, this creates distributive ambiguity because different fairness ideals support different efficient contribution rules.

Moving to an interior public good may itself weaken peer punishment. Cason and Gangadharan (2015) compare linear and nonlinear public good environments and find that punishment can still improve cooperation, but its effect is weaker and slower to emerge in nonlinear settings. They argue that nonlinear environments are more complex because socially optimal choices are harder to identify, making individually appropriate behavior less transparent and deviations harder to interpret. This makes it important to distinguish the effect of the interior environment itself from the effect of inequality within that environment. Kingsley (2016) provides suggestive evidence that endowment heterogeneity limits the effectiveness of punishment in an interior public good game: unequal groups contribute and earn significantly less than equal groups when punishment is available. However, because endowments were not observed in that design, the effect of inequality cannot be separated cleanly from the informational problem of identifying whether a given contribution is high or low relative to capacity. The present design addresses this limitation by including an equal-endowment interior baseline and by making both contributions and endowments observable in the unequal treatments. It therefore isolates whether inequality weakens punishment even when group members observe endowments and can evaluate candidate contribution norms under complete information.

Taken together, the literature suggests that decentralized punishment is most effective when groups can coordinate on a clear contribution benchmark and reliably distinguish compliance from violation. Existing studies indicate that neither endowment heterogeneity nor interior public goods necessarily prevents cooperation, but each can complicate decentralized enforcement through different mechanisms. The present study combines these strands by examining whether inequality, in an interior public good with complete information, affects coordination around contribution norms, the enforcement of those norms, or both.

3 Design

The experiment consists of 20 periods divided into two 10-period phases. Subjects are assigned to groups of five, and group composition remains fixed throughout the experiment. In phase 1, groups participate in a voluntary contribution public good game with the following payoff function:

$$\pi_i = e_i - x_i + 0.5 \min\left\{\sum_{j=1}^{n} x_j,\, 50\right\},$$

where e_i is subject i's endowment in Experimental Dollars (EDs), x_i is subject i's contribution to the public good, and $\sum_{j=1}^{n} x_j$ is total group contribution. After contribution decisions are made, each group member observes the endowment, contribution, and period earnings of every other group member. Group members are identified by fixed letters (A, B, C, D, or E) that remain the same throughout the experiment. The payoff function behaves like a linear public good until aggregate contributions reach 50 EDs, with each contributed ED returning 0.5 EDs to every group member. Once aggregate contributions reach 50 EDs, the marginal return from any additional contribution falls to zero, so contributions beyond 50 reduce the contributor's private earnings without increasing the earnings of any group member. This creates an interior efficient contribution target at 50 EDs.

In phase 2, subjects remain in the same groups for 10 additional periods, and each period consists of two stages. In the first stage, subjects make the same contribution decision as in phase 1, and stage 1 earnings are determined by the same payoff function. In the second stage, subjects observe the endowments, contributions, and stage 1 earnings of each other group member and may assign *Reduction Points*. Each subject can assign between 0 and 10 Reduction Points to each other group member. Assigning punishment is costly: each Reduction Point costs the punisher 1 ED, referred to as an *Administrative Cost*, and reduces the recipient's earnings by 3 EDs, referred to as a *Reduction Cost*. Final period earnings in phase 2 are calculated by first subtracting any Reduction Costs from Stage 1 earnings, bounded below at zero, and then subtracting the Administrative Costs of any Reduction Points assigned to others.

Given this payoff structure, aggregate contributions of exactly 50 EDs maximize total group earnings in every treatment. Contributing less than 50 leaves potential gains unrealized, while contributing more than 50 reduces private earnings without generating additional public returns. There are, however, multiple ways to allocate this efficient contribution total across group members. We focus on three candidate contribution benchmarks: Equal Contributions, Equal Proportions, and Equal Earnings.

Across treatments, total group endowment is held fixed at 100 EDs, but the distribution of endowments varies. In *Equal*, all five group members receive an endowment of 20 EDs. In *Low Inequality*, one group member receives 10 EDs, three group members receive 20 EDs, and one group member receives 30 EDs. In *High Inequality*, two group members receive 10 EDs, one group member receives 20 EDs, and two group members receive 30 EDs. Throughout the paper, subjects with 10 EDs are referred to as *Low* types, subjects with 20 EDs as *Middle* types, and subjects with 30 EDs as *High* types. Contributions are chosen in five-ED increments from zero up to the subject's endowment: $x_i \in \{0, 5, 10\}$ for Low types, $x_i \in \{0, 5, 10, 15, 20\}$ for Middle types, and $x_i \in \{0, 5, 10, 15, 20, 25, 30\}$ for High types.

Table 1 reports the individual contribution and resulting earnings implied by each candidate benchmark. Each benchmark requires aggregate contributions of 50 EDs and is therefore efficient, but the benchmarks allocate the burden of provision differently across types. Under Equal Contributions, all types contribute 10 EDs, regardless of one's endowment. Under Equal Proportions, each type contributes half of their endowment, so Low, Middle, and High types contribute 5, 10, and 15 EDs, respectively. Under Equal Earnings, High types contribute 20 EDs, Middle types contribute 10 EDs, and Low types contribute 0 EDs, so that all types earn 35 EDs.

	Equal Contributions		Equal Proportions		Equal Earnings	
	x_i	π_i	x_i	π_i	x_i	π_i
Equal	10	35	10	35	10	35
Low type	10	25	5	30	0	35
Middle type	10	35	10	35	10	35
High type	10	45	15	40	20	35

Table 1: Candidate contribution benchmarks and implied earnings. The *Equal* row applies to each group member in the equal-endowment treatment; Low, Middle, and High rows apply to endowment types in the unequal treatments.

Because endowments and contributions are observable, group members can evaluate behavior relative to each candidate benchmark. The key difference across treatments is whether these benchmarks identify a unique focal contribution. In *Equal*, all three benchmarks prescribe the same contribution for every participant. In the unequal treatments, the benchmarks diverge for Low and High types: this affects one Low type and one High type in *Low Inequality*, and two Low types and two High types in *High Inequality*. Thus, moving from Equal to Low Inequality to High Inequality increases endowment dispersion and makes the efficient contribution target increasingly open to competing contribution benchmarks. This creates scope for inequality to affect not only contribution levels, but also the benchmark around which punishment is organized and the strength with which deviations from that benchmark are enforced. The analysis below examines these possibilities by studying contributions, earnings, punishment decisions, and responses to punishment.

4 Results

Data were collected at the Centre for Decision Research and Experimental Economics (CeDEx) at the University of Nottingham in November and December 2024. Three experimental sessions with 25 participants each were completed for each treatment. The experiment was programmed and conducted using z-Tree (Fischbacher, 2007). Experimental earnings were converted at a rate of 30 EDs per £1. Participants earned an average of approximately 542 EDs over the 20 periods, equivalent to £18.07 (approximately $22.85 at late-2024 exchange rates). Table 2 summarizes the collected demographic data across treatments.

	Age	Gender	Semesters	Economics
Equal	22.01	0.45	3.58	1.01
	(3.72)	(0.49)	(2.33)	(2.76)
Low Inequality	21.8	0.45	3.84	1.07
	(3.03)	(0.49)	(2.20)	(2.90)
High Inequality	21.32	0.47	3.42	1.06
	(2.41)	(0.49)	(2.62)	(2.23)

Table 2: Means of collected demographics with standard deviations in parentheses. Age is measured in years, Gender equals 1 if Male, Semesters is the number of university semesters completed, and Economics is the number of economics courses taken. Missing observations: Gender = 5, Semesters = 3, Economics = 6.

We begin by summarizing the main patterns in the data. The introduction of peer punishment increases contributions in all treatments, but the increase is largest in *Equal* and significantly smaller under inequality. Earnings follow a different pattern. Average earnings do not change significantly after punishment is introduced in *Equal*, but significantly decline under inequality.

We then examine whether these differences reflect the contribution benchmark that groups appear to enforce. The punishment data are most consistent with the proportional-contribution benchmark in each treatment, suggesting that subjects largely target deviations from a common proportional contribution norm. However, enforcement of this norm differs sharply across treatments. Negative deviations from the proportional benchmark are punished less strongly under inequality, on both the extensive and intensive margins. There is also suggestive evidence that punished deviators are less responsive to sanctions under inequality. Together, these results suggest that inequality does not primarily change which benchmark is enforced; rather, it weakens groups' ability to sustain and enforce that benchmark through peer punishment.

4.1 Contributions

We first examine whether the introduction of peer punishment increases contributions and whether this effect differs when endowments are unequal.[1] Because endowments differ across treatments, the dependent variable is individual contribution measured as a proportion of the subject's endowment. We estimate an individual random-effects specification with standard errors clustered at the group level. The model includes treatment indicators, an indicator for the punishment phase (Part 2), and treatment-phase interactions. To account for within-phase dynamics and adjustment to prior behavior, we include separate linear time trends for the VCM and punishment phases, along with the subject's own lagged proportional contribution and the lagged average proportional contribution of the other group members.

Table 3: Proportional Contributions

	Proportion Contributed
Low Inequality	0.0115
	(0.0142)
High Inequality	0.0141
	(0.0121)
Punishment phase (Part 2)	0.1458***
	(0.0152)
Low Inequality × Part 2	-0.0517***
	(0.0186)
High Inequality × Part 2	-0.0557***
	(0.0171)
VCM period trend	-0.0048***
	(0.0017)
Punishment period trend	-0.0049***
	(0.0016)
Lagged own prop. cont.	0.2703***
	(0.0295)
Lagged mean others' prop. cont.	0.0973**
	(0.0471)
Constant	0.1767***
	(0.0215)
Observations	4,275
Subjects	225
Wald χ^2	1397.47***

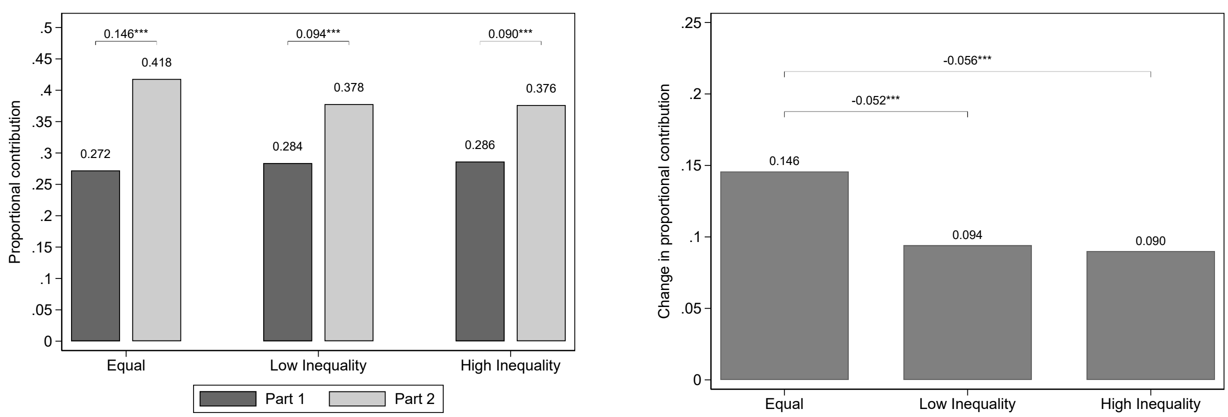
Random-effects GLS estimates with standard errors clustered at the group level in parentheses. *Equal* is the omitted treatment. Part 2 indicates the punishment phase. VCM period trend is coded 1–10 within Part 1; punishment period trend is coded analogously within Part 2. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

[1] Figure A.1 in the Appendix reports descriptive statistics and time series for group level contributions and Table A.1 reports non-parametric analysis.

Table 3 reports the main contribution regression.[2] The regression confirms that proportional contributions do not differ significantly across treatments during the VCM phase. The time trends indicate modest declines in contributions within phases, while the lagged contribution variables show persistence in behavior: subjects contribute more when they and their group members contributed more in the previous period. The key treatment effects are summarized in Figure 1.

Panel A of Figure 1 reports predicted proportional contributions by treatment and phase, while Panel B reports the estimated effect of introducing punishment within each treatment. Peer punishment increases proportional contributions in all three treatments: by 0.146 in *Equal*, 0.094 in *Low Inequality*, and 0.090 in *High Inequality* (all $p < 0.01$). However, the increase is significantly larger in *Equal* than in either inequality treatment. Relative to *Equal*, the punishment effect is smaller by 0.052 in *Low Inequality* ($p < 0.01$) and by 0.056 in *High Inequality* ($p < 0.01$).

The weaker punishment effect under inequality translates into clear treatment differences during the punishment phase. Predicted proportional contributions in Part 2 are 0.418 in *Equal*, compared with 0.378 in *Low Inequality* and 0.376 in *High Inequality*. Both inequality treatments are significantly lower than *Equal* in the punishment phase (*Low Inequality–Equal*: -0.040, $p < 0.01$; *High Inequality–Equal*: -0.042, $p < 0.01$).[3]



(a) Predicted prop. cont. by treatment and part

(b) Ave. marginal effect of punishment

Figure 1: Effects of punishment on proportional contributions. Panel A reports predicted proportional contributions by treatment and part; brackets indicate the within-treatment effect of introducing punishment. Panel B reports the average marginal effect of punishment by treatment; brackets indicate differences in treatment effects relative to *Equal*.

These results show that peer punishment improves cooperation in all treatments, but substantially less so when endowments are unequal. Inequality does not eliminate the cooperative response to punishment; rather, it dampens the size of that response.

[2]Table A.2 reports alternative specifications accounting for demographic controls and subject fixed effects. The results are qualitatively unchanged. Tables A.3 and A.4 present nonparametric analysis of the proportional contributions by group and by type.

[3]All post-estimation comparisons, including Part 1 predicted levels, are reported in Table A.5.

4.2 Earnings

We next examine whether the weaker contribution response under inequality translates into lower earnings.[4] The dependent variable is individual period earnings, net of any punishment costs paid or received. We estimate an individual random-effects specification with standard errors clustered at the group level. The model includes treatment indicators, an indicator for the punishment phase (Part 2), and treatment-phase interactions. We include separate linear time trends for the VCM and punishment phases to account for changes in earnings within each part of the experiment.

Table 4: Earnings

	Earnings
Low Inequality	0.467
	(0.554)
High Inequality	0.577
	(0.549)
Punishment phase (Part 2)	-2.130
	(1.556)
Low Inequality × Part 2	-3.979**
	(1.833)
High Inequality × Part 2	-4.620***
	(1.550)
VCM period trend	-0.346***
	(0.054)
Punishment period trend	0.260**
	(0.123)
Constant	29.498***
	(0.522)
Observations	4,500
Subjects	225
Wald χ^2	75.33***

Random-effects GLS estimates with standard errors clustered at the group level in parentheses. *Equal* is the omitted treatment. Part 2 indicates the punishment phase. VCM period trend is coded 1–10 within Part 1; punishment period trend is coded analogously within Part 2. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

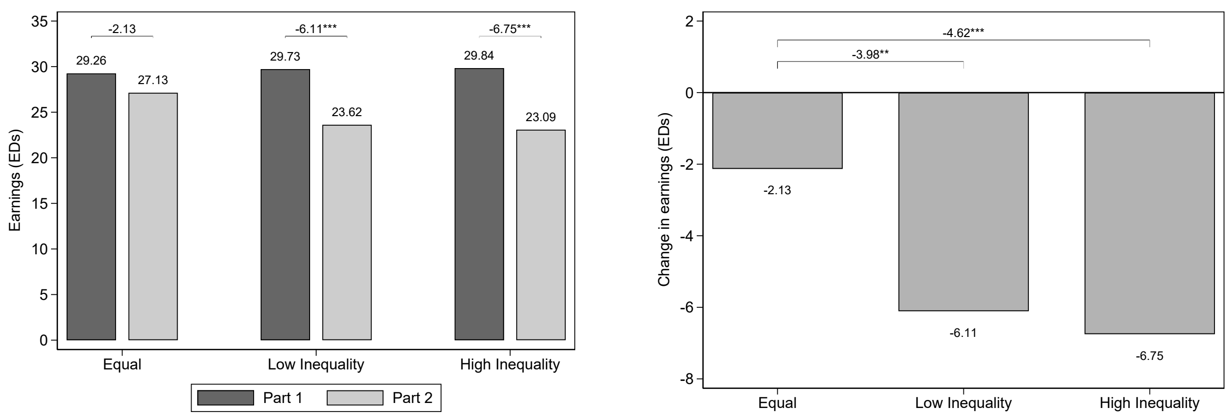
Table 4 reports the main earnings regression.[5] The regression confirms that earnings do not differ significantly across treatments during the VCM phase. Earnings decline over time during the VCM phase but increase over time during the punishment phase. Figure 2 summarizes the main treatment effects.

Panel A of Figure 2 reports predicted earnings by treatment and phase, while Panel B reports the estimated effect of introducing punishment within each treatment. Earnings do not increase with the introduction of punishment in any treatment. The estimated change in earnings is -2.13

[4]Figure A.2 in the Appendix reports descriptive statistics and time series for group level earnings and Table A.6 reports non-parametric analysis.

[5]Table A.7 reports alternative specifications accounting for demographic controls and subject fixed effects. The results are qualitatively unchanged.

EDs in *Equal* ($p = 0.17$), compared with -6.11 EDs in *Low Inequality* ($p < 0.01$) and -6.75 EDs in *High Inequality* ($p < 0.01$). The earnings effect of punishment is significantly more negative in both inequality treatments relative to *Equal* (*Low Inequality*: -3.98, $p = 0.03$; *High Inequality*: -4.62, $p < 0.01$). As a result, the predicted earnings in Part 2 are 27.13 EDs in *Equal*, compared with 23.62 EDs in *Low Inequality* and 23.09 EDs in *High Inequality*. The difference between *High Inequality* and *Equal* is statistically significant (-4.04, $p = 0.010$), while the difference between *Low Inequality* and *Equal* is marginally significant (-3.51, $p = 0.054$).[6]



(a) Predicted earnings by treatment and part **(b)** Average marginal effect of punishment on earnings

Figure 2: Effects of punishment on earnings. Panel A reports predicted earnings by treatment and part; brackets indicate the within-treatment effect of introducing punishment. Panel B reports the average marginal effect of punishment by treatment; brackets indicate differences in treatment effects relative to *Equal*.

The earnings results show that the weaker contribution response under inequality has meaningful welfare consequences. Although punishment increases contributions in all treatments, it does not improve earnings. Instead, introducing punishment significantly lowers earnings in the inequality treatments, suggesting that the contribution gains from punishment are not large enough to offset the costs imposed through sanctions. Thus, inequality not only dampens the cooperative effect of punishment, but also worsens its earnings consequences. The remaining analysis examines why punishment is less effective under inequality by studying the contribution norms groups appear to enforce, the strength of punishment for deviations from those norms, and the behavioral response to sanctions.

4.3 Punishment, Norms, and Enforcement

The contribution and earnings results show that peer punishment is less effective under inequality, but they do not by themselves explain why. We therefore examine whether inequality changes the sanctioning environment along four dimensions. First, we identify the contribution benchmark that best predicts punishment decisions. This allows us to assess how inequality affects the ability of groups to coordinate on a contribution benchmark. Second, we examine whether inequality affects

[6] All post-estimation comparisons, including Part 1 predicted levels, are reported in Table A.8.

the overall amount of punishment deployed. Third, we determine how intensely deviations from estimated contribution norms are punished across treatments to understand how inequality affects norm enforcement. Finally, we estimate how deviators respond to being punished.

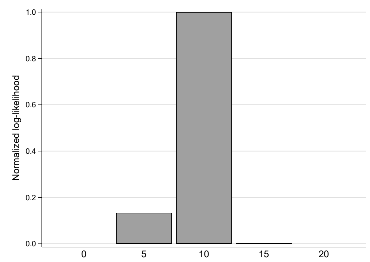
4.3.1 Punishment Benchmark

Here we determine which contribution benchmark best predicts punishment decisions. This exercise asks whether inequality changes the norm that subjects appear to enforce, or whether the same benchmark remains relevant across treatments. That is, if inequality impairs the ability of groups to agree on what constitutes fair behavior and establish contribution norms, decentralized enforcement will be ineffective. The unit of observation is each punishment opportunity by each potential punisher: for each group member i, each potential target j, and each period t in the punishment phase, we observe whether i assigns any punishment to j. For each candidate benchmark b, we define an indicator equal to one if the target's contribution falls below that benchmark, $x_{jt} < b$, and zero otherwise. We then estimate random-effects probit models of the probability that punishment is assigned, with punisher-level random effects and standard errors clustered at the group level.

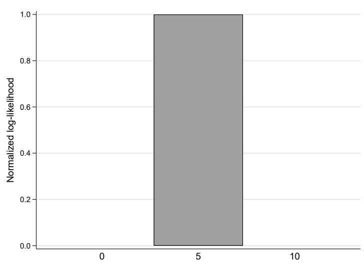
This approach follows norm-identification exercises in the punishment literature, which infer enforced norms by asking which candidate contribution rule best predicts sanctioning behavior (Carpenter and Matthews, 2009; Reuben and Riedl, 2013). Here, we select benchmarks using the extensive margin of punishment rather than a model that combines the probability and intensity of punishment. This choice reflects the purpose of the benchmark exercise: to identify whether a target's contribution is treated as a norm violation. The decision to punish provides a clear signal that a target's contribution is viewed as falling below an enforceable benchmark. We therefore use the probit benchmark model to identify the relevant contribution benchmark, and then examine punishment deployed, punishment intensity, and behavioral responses separately below.

For each treatment and target-endowment type, we estimate the probit model across candidate benchmarks and select the benchmark that maximizes the log likelihood. Because contributions are chosen in five-ED increments, candidate benchmarks are restricted to feasible contribution levels: $b \in \{0, 5, 10\}$ for Low types, $b \in \{0, 5, 10, 15, 20\}$ for Middle types, and $b \in \{0, 5, 10, 15, 20, 25, 30\}$ for High types. The model controls for the punisher's own proportional contribution, total group contributions, period, and the punisher's lagged punishment received.

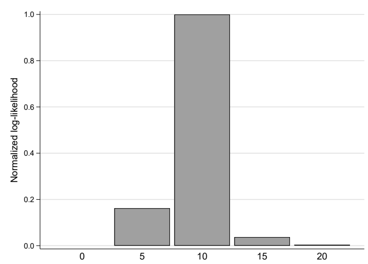
Figure 3 reports the normalized log likelihoods from the benchmark-selection exercise. The best-fitting model in each panel is normalized to 1. The other bars indicate how well the alternative benchmarks fit the data relative to the best fitting benchmark. More sharply peaked panels indicate that punishment decisions are more tightly anchored to a specific benchmark. Across treatments and target-endowment types, the best-fitting benchmark corresponds to contributing one-half of one's endowment. In *Equal*, where all subjects have an endowment of 20, the winning benchmark is 10. In both unequal treatments, the winning benchmark is 5 for Low types, 10 for Middle types, and 15 for High types. Thus, the benchmark that best predicts the decision to punish



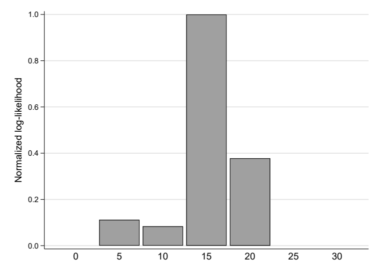
(a) *Equal*



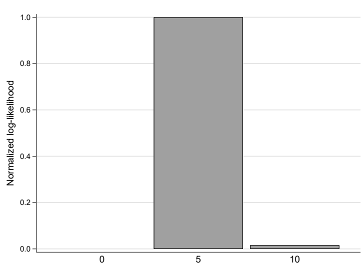
(b) *Low Inequality*, *Low*



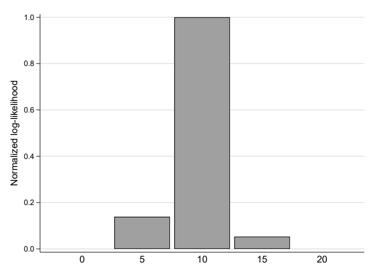
(c) *Low Inequality*, *Middle*



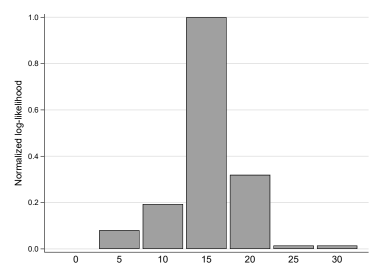
(d) *Low Inequality*, *High*



(e) *High Inequality*, *Low*



(f) *High Inequality*, *Middle*



(g) *High Inequality*, *High*

Figure 3: Benchmark fit from extensive-margin punishment models. Each panel reports normalized log-likelihood values from random-effects probit models estimated separately by treatment and target-endowment type, with values normalized relative to the best-fitting benchmark within the panel. The x-axis gives the candidate contribution used to define below-benchmark contributions. Higher bars indicate benchmarks that better predict whether punishment is assigned.

is the proportional-contribution rule.[7] Although inequality creates multiple efficient and normatively appealing contribution rules, punishment is best explained by the proportional-contribution benchmark in both unequal treatments. Thus, inequality does not appear to weaken cooperation primarily by preventing groups from coordinating on a common contribution norm. The following analyses examine why enforcement of this common benchmark produces weaker contribution and earnings effects under inequality.

4.3.2 Punishment Use

The benchmark analysis identifies the contribution norm that best organizes punishment behavior, but it does not show whether inequality changes the amount of punishment used. This distinction is important for interpreting the contribution and earnings results. If unequal groups use less punishment overall, then the smaller increase in contributions under inequality may reflect weaker sanctioning pressure. Alternatively, if unequal groups deploy more punishment, then the lower earnings observed under inequality may reflect the direct costs of punishment. We therefore next examine whether the total amount of punishment differs across treatments.

Table 5: Punishment Imposed

	Punishment Cost Imposed
Low Inequality	0.584
	(0.389)
High Inequality	0.597*
	(0.320)
Punishment period trend	-0.071**
	(0.030)
Constant	1.602***
	(0.331)
Observations	2,250
Subjects	225
Wald χ^2	9.25**

Random-effects GLS estimates with standard errors clustered at the group level in parentheses. The sample is restricted to the punishment phase, periods 11–20. *Equal* is the omitted treatment. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

To examine this, we measure punishment imposed as the total cost a subject pays to sanction other group members in a period. We estimate a random-effects model of punishment imposed on treatment indicators and a linear period trend within the punishment phase, with standard errors clustered at the group level.

Table 5 reports the regression and Figure 4 displays the predicted punishment costs across treat-

[7] Table A.9 in the Appendix reports the probit estimates for the winning proportional benchmarks. The below-benchmark indicator is positive in every treatment and target-endowment type, indicating that falling below one-half of one's endowment systematically increases the likelihood of punishment.

ments.[8] Predicted average punishment costs imposed are 1.21 EDs in *Equal*, compared with 1.79 EDs in *Low Inequality* and 1.81 EDs in *High Inequality*. Relative to *Equal*, predicted punishment expenditures are 0.584 EDs higher in *Low Inequality* ($p = 0.133$) and 0.597 EDs higher in *High Inequality* ($p = 0.062$). Although these estimated differences are not jointly statistically significant ($\chi^2(2) = 3.82$, $p = 0.148$), punishment expenditures are moderately higher in the unequal treatments. Thus, although subjects in unequal groups appear to devote more resources to punishment, these additional expenditures are not accompanied by comparable improvements in contributions or earnings. The following analyses examine whether this reflects differences in how punishment is directed and how subjects respond to sanctions.

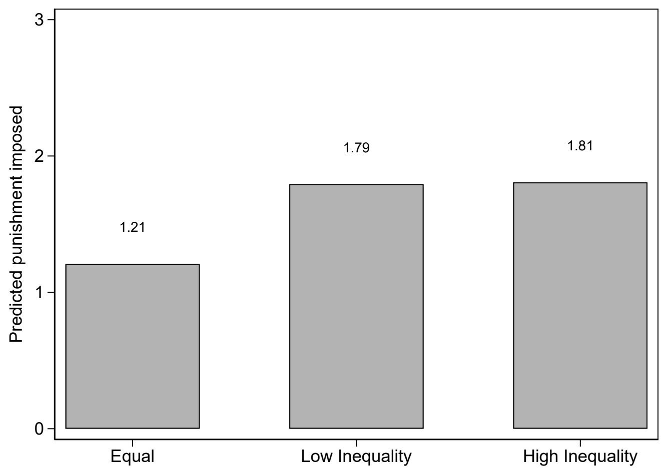


Figure 4: Predicted punishment cost imposed by treatment.

4.3.3 Enforcement of Norms

We next ask whether deviations from the proportional benchmark are sanctioned with the same strength across treatments. Unlike the benchmark-selection exercise, which uses punisher–target observations to identify the best-fitting contribution rule, the analysis here is conducted at the subject-period level and measures punishment received by each potential target. For each subject-period, we define negative deviation from the proportional norm as the amount by which the subject's proportional contribution falls below 0.5. This measure equals zero for subjects who contribute at or above one-half of their endowment and increases with the extent of undercontribution relative to the proportional benchmark. To facilitate interpretation, we rescale the measure so that a one-unit increase corresponds to a 10 percentage-point larger shortfall from the benchmark. Thus, a subject contributing 40% of their endowment has a rescaled negative deviation of 1, while a subject contributing nothing has a value of 5.

We estimate two models of punishment received. First, we estimate a random-effects probit

[8] Post-estimation comparisons are reported in Table A.10. Table A.11 provides nonparametric analysis for the amount of punishment sent and received.

Table 6: Punishment of Negative Deviations from the Proportional Norm

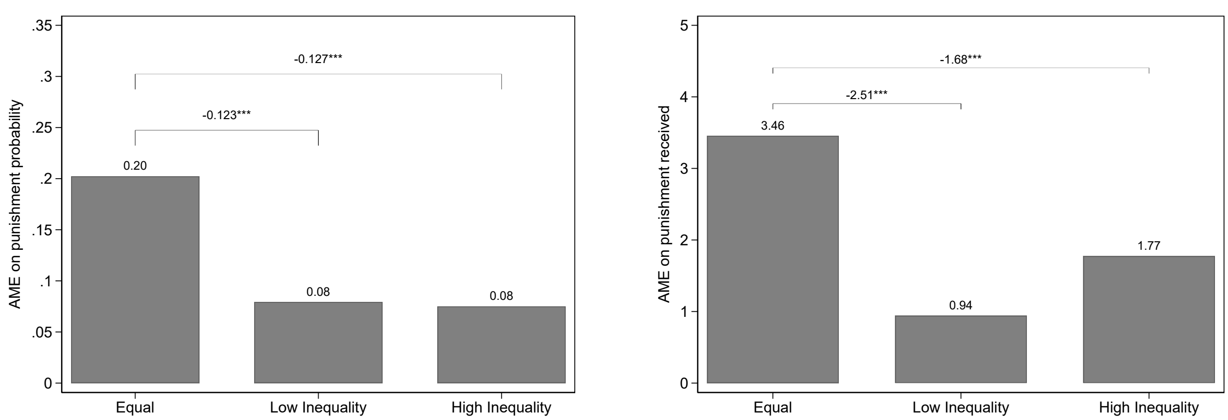
	Any punishment	Punishment cost
Low Inequality	0.584	3.191*
	(0.369)	(1.645)
High Inequality	0.789**	0.400
	(0.358)	(1.417)
Negative deviation from prop. norm	1.160***	3.456***
	(0.147)	(0.471)
Low Inequality × Negative deviation	-0.863***	-2.513***
	(0.161)	(0.575)
High Inequality × Negative deviation	-0.886***	-1.681***
	(0.160)	(0.592)
Punishment period trend	-0.079***	0.001
	(0.021)	(0.153)
Total group contribution	-0.002	0.167**
	(0.012)	(0.069)
Constant	-0.594	0.022
	(0.610)	(3.218)
Observations	2,250	958
Subjects	225	195
Wald χ^2	106.21***	101.94***

Column (1) reports random-effects probit estimates for whether a subject receives any punishment in a period. Column (2) reports random-effects GLS estimates for punishment cost received, conditional on positive punishment. Standard errors are clustered at the group level and reported in parentheses. *Equal* is the omitted treatment. Negative deviation from the proportional norm is measured as $\max(0, 0.5 - x_i/e_i)$ and rescaled so that one unit equals a 10 percentage-point shortfall from the proportional benchmark. Because coefficients in Column (1) are probit-index coefficients, treatment-specific marginal effects are reported in Figure 5 and Table A.12. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

model for whether a subject receives any punishment, capturing the extensive margin of enforcement. Second, restricting the sample to observations with positive punishment received, we estimate a random-effects linear model of punishment cost received, capturing the intensive margin. Both models interact negative deviation with treatment indicators and include a linear period trend within the punishment phase and total group contributions, with standard errors clustered at the group level.

Table 6 reports both regressions. In both specifications, below-norm deviations are positively associated with punishment in *Equal*. The interaction terms are negative and statistically significant in both inequality treatments, indicating that punishment is less responsive to proportional shortfalls when endowments are unequal.

Figure 5 displays the treatment-specific marginal effects.[9] Panel A shows that a 10 percentage-point larger shortfall from the proportional benchmark increases the probability of receiving punishment in all treatments. The marginal effect is 0.202 in *Equal*, compared with 0.079 in *Low Inequality* and 0.075 in *High Inequality* (all $p < 0.01$). Both inequality-treatment effects are significantly smaller than the effect in *Equal*: by 0.123 in *Low Inequality* ($p < 0.01$) and by 0.127 in *High Inequality* ($p < 0.01$). Panel B shows the same pattern for punishment intensity conditional on receiving punishment. A 10 percentage-point larger shortfall increases punishment cost received by 3.46 EDs in *Equal*, compared with 0.94 EDs in *Low Inequality* and 1.77 EDs in *High Inequality* (all $p < 0.01$). Again, both inequality-treatment effects are significantly smaller than the effect in *Equal*: by 2.51 EDs in *Low Inequality* ($p < 0.01$) and by 1.68 EDs in *High Inequality* ($p < 0.01$).



(a) Marginal effect on punishment probability

(b) Marginal effect on punishment received

Figure 5: Punishment responsiveness to negative deviations from the proportional contribution norm. Panel A reports the average marginal effect on the probability of receiving punishment. Panel B reports the estimated effect on punishment cost received, conditional on receiving positive punishment. Effects are reported for a 10 percentage-point larger shortfall from the proportional contribution benchmark. Brackets in both panels indicate statistically significant differences in treatment effects relative to *Equal*.

These results show that the proportional benchmark organizes punishment behavior in all treat-

[9]Post-estimation comparisons are reported in Table A.12. Table A.13 in the Appendix reports a linear probability model as a complement to the probit specification.

ments, but that enforcement around this benchmark is weaker when endowments are unequal. Punishment is more responsive to the size of below-benchmark deviations in *Equal*: larger shortfalls are more likely to trigger punishment and, conditional on punishment, receive more severe sanctions. Taken together, the evidence indicates that inequality weakens both the likelihood and intensity of punishment directed toward below-benchmark contributions. The next section examines whether punished subjects also differ in how they adjust their subsequent contributions.

Table 7: Response to Punishment by Norm Deviation

	Below Proportional Norm	At or above Proportional Norm
Low Inequality	-0.015	-0.047**
	(0.035)	(0.020)
High Inequality	-0.028	-0.014
	(0.033)	(0.020)
Punishment received, $t-1$	0.003***	0.002
	(0.001)	(0.002)
Low Inequality $\times$ Pun. rec., $t-1$	-0.003	-0.001
	(0.002)	(0.002)
High Inequality $\times$ Pun. rec., $t-1$	-0.003*	-0.002
	(0.002)	(0.003)
Punishment period trend	-0.006*	0.002
	(0.003)	(0.002)
Lagged own prop. cont.	-0.875***	-1.116***
	(0.096)	(0.078)
Lagged mean others' prop. cont.	0.268**	0.111
	(0.115)	(0.073)
Constant	0.236***	0.456***
	(0.058)	(0.047)
Observations	554	1,471
Subjects	142	213
Wald χ^2	229.97***	285.14***

Random-effects GLS estimates with standard errors clustered at the group level in parentheses. The dependent variable is the change in the proportion of a subject's endowment contributed from period $t-1$ to period t. The sample is restricted to punishment-phase observations for periods 12–20. Column (1) includes observations in which the subject's proportional contribution in period $t-1$ was below 0.5. Column (2) includes observations in which the subject's proportional contribution in period $t-1$ was at or above 0.5. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

4.3.4 Response to Punishment

Finally, we examine whether received punishment produces corrective changes in subsequent proportional contributions. For each subject-period in the punishment phase, we define contribution change as the difference between the proportion of the subject's endowment contributed in period t and the corresponding proportion in period $t-1$. We regress this change on punishment received in period $t-1$, treatment indicators, and their interactions. Because punishment is expected to be most corrective when directed at subjects who were below the proportional benchmark, we estimate separate models for subjects whose proportional contribution in period $t-1$ was below 0.5 and for subjects whose proportional contribution was at or above 0.5. Both specifications include a linear period trend within the punishment phase, the subject's lagged proportional contribution, and the lagged mean proportional contribution of the other group members, with individual random effects and standard errors clustered at the group level.

Table 7 reports the response-to-punishment regressions separately by prior norm status. Figure 6 illustrates the response to punishment among subjects who were below the proportional benchmark in the previous period.[10] In *Equal*, punishment received is associated with a positive and statistically significant increase in the subsequent proportional contribution. The estimated marginal effect is 0.0035 ($p < 0.01$), indicating that each additional ED of punishment received is associated with a 0.35 percentage-point larger increase in the proportion of the subject's endowment contributed in the next period. The estimated response is 0.0029 smaller in *Low Inequality* ($p = 0.139$) and 0.0033 smaller in *High Inequality* ($p = 0.053$). Consequently, the treatment-specific responses under inequality are close to zero and statistically insignificant: 0.0005 in *Low Inequality* ($p = 0.698$) and 0.0002 in *High Inequality* ($p = 0.903$).

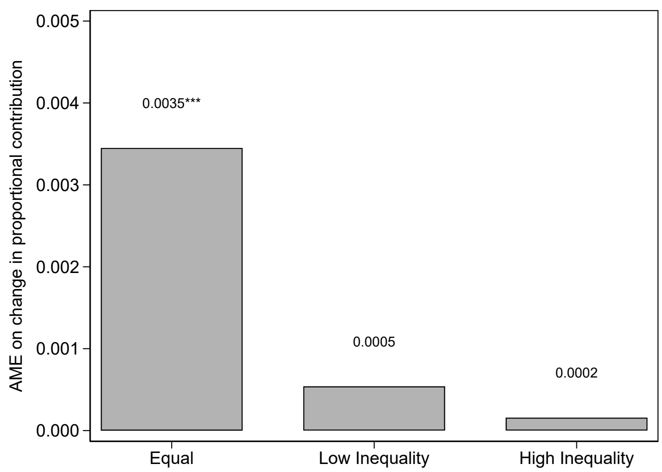


Figure 6: Average marginal effect of punishment on the subsequent change in proportional contribution for subjects below the proportional contribution norm. Bars show the treatment-specific marginal effects of punishment received in the previous period.

For subjects who were at or above the proportional benchmark in the previous period, punishment has no systematic effect on subsequent proportional contributions in any treatment. The estimated marginal effects of lagged punishment received are small and statistically insignificant: 0.0016 in *Equal* ($p = 0.411$), 0.0008 in *Low Inequality* ($p = 0.565$), and -0.0008 in *High Inequality* ($p = 0.657$). The treatment differences in these marginal effects are also statistically insignificant, both individually and jointly.

Taken together, the response results suggest that punishment is corrective primarily when below-

[10] Complete post-estimation comparisons are reported in Table A.14.

norm contributions occur in *Equal*. Under inequality, the estimated responses of below-norm subjects are substantially smaller, close to zero, and statistically insignificant. The cross-treatment evidence remains suggestive rather than conclusive, however: the difference between *High Inequality* and *Equal* is significant only at the 10 percent level, the difference between *Low Inequality* and *Equal* is insignificant, and the treatment interactions are not jointly significant. Combined with the punishment-targeting and intensity results, these estimates indicate that inequality weakens decentralized norm enforcement primarily by reducing the strength with which deviations from the proportional norm are sanctioned, with additional suggestive evidence that punishment is also less effective at inducing corrective responses from undercontributors.

5 Discussion

This paper examines how endowment inequality affects cooperation and peer enforcement in an interior-solution public good environment. The main result is that punishment increases contributions in all treatments, but is substantially less effective when endowments are unequal. Under equality, punishment produces the largest contribution increase and does not significantly change earnings. Under inequality, punishment still raises contributions, but the increase is smaller and earnings decline significantly. Consequently, during the punishment phase, both contributions and earnings are lower in the unequal treatments than in the equal treatment.

The evidence suggests that this difference is not driven primarily by disagreement over which contribution norm to enforce. The benchmark-identification exercise shows that punishment is organized around a proportional contribution norm in all treatments: subjects are more likely to be punished when they contribute less than one-half of their endowment. Negative deviations from this benchmark increase both the likelihood and intensity of punishment in every treatment, but these relationships are considerably stronger under equality. When endowments are unequal, punishment is less responsive to whether a subject falls below the benchmark and, conditional on punishment occurring, less responsive to the size of the deviation. Thus, inequality does not appear to change the content of the norm around which punishment is organized. Instead, it weakens the enforcement of that norm. There is also suggestive evidence that punished undercontributors adjust their subsequent contributions less under inequality.

These findings contribute to the literature on punishment in heterogeneous public good environments. Previous work in linear public good games shows that peer punishment can sustain cooperation when contribution norms are clear and behavior can be evaluated relative to members' capacities. The present results qualify this conclusion in an interior-solution environment. Even when punishment behavior is organized around a common proportional benchmark, inequality weakens the ability of decentralized sanctions to sustain that benchmark. The challenge created by inequality is therefore not limited to selecting an appropriate burden-sharing rule. Inequality may also make an otherwise recognizable rule more difficult to enforce effectively.

The interior structure is important because it separates group efficiency from a unique individual contribution profile. In the equal-endowment treatment, equal contributions, equal proportional contributions, and equal earnings all prescribe the same allocation. Under inequality, these benchmarks diverge even though each can support efficient provision. This feature allows the experiment to distinguish between two potential effects of inequality: difficulty coordinating on a contribution norm and difficulty enforcing that norm once it has been identified. The results point primarily to the latter mechanism.

The mechanism identified here may extend beyond public good provision to social dilemmas more generally. Social dilemmas arise when individually costly cooperation produces collective benefits, and their resolution often depends on informal norms and decentralized responses to noncooperation. Peer punishment is one such response: individuals incur costs to sanction norm violations in ways that can sustain cooperation. The present results suggest that inequality may weaken this form of informal governance even when group members appear to share an understanding of the cooperative obligation. In that sense, inequality need not undermine cooperation by creating wholly different views of appropriate behavior. It may instead reduce group members' willingness or ability to enforce a commonly recognized rule, or reduce the effectiveness of their enforcement efforts.

This interpretation provides a cautious connection to the literature on social capital. Social capital is commonly understood to encompass the norms, trust, and networks that facilitate collective action (Coleman, 1990; Putnam et al., 1993). Within this broad conception, the willingness to support cooperative norms and sanction violations can be viewed as part of a group's informal capacity for self-governance (Ostrom, 2000; Fehr and Gächter, 2002). The experiment does not measure social capital or economic growth directly, but it identifies a behavioral mechanism through which inequality may affect both. If unequal groups enforce commonly understood cooperative norms less effectively, inequality may weaken one component of the informal institutional capacity on which sustained cooperation depends. This interpretation is consistent with evidence linking inequality to lower trust and weaker social capital, as well as with research associating inequality with slower or less durable economic growth (Aghion et al., 1999; Knack and Keefer, 1997; Zak and Knack, 2001; Ostry et al., 2014; Fehr et al., 2020). The results therefore provide micro-level evidence for a channel through which inequality may weaken the informal institutions that support cooperation and economic performance.

Several limitations point to useful directions for future research. The experiment studies fixed groups with observable, exogenously assigned endowments and a particular punishment technology, so the results may differ when groups can communicate or choose their enforcement institutions. In addition, because subjects do not earn their endowments, the experiment does not examine how perceived property rights or the legitimacy of inequality affect norm selection and enforcement. When endowment differences reflect prior effort or performance, high-endowment members may feel more entitled to retain their resources, while other group members may view unequal contributions as more justified. These considerations could alter both the contribution benchmark that groups enforce and the willingness of subjects to sanction deviations from it. Future work could therefore examine whether communication helps unequal groups coordinate enforcement, whether centralized institutions outperform peer punishment in this setting, and whether earned inequality changes the norms around which enforcement is organized. The central conclusion is that inequality does not eliminate cooperation or prevent groups from identifying a proportional contribution norm. Rather, it weakens the decentralized enforcement of that norm. As a result, punishment raises contributions less and fails to offset the costs of sanctions when endowments are unequal.

References

Aghion, P., E. Caroli, and C. García-Peñalosa (1999). Inequality and economic growth: The perspective of the new growth theories. *Journal of Economic Literature 37*(4), 1615–1660.

Almås, I., A. W. Cappelen, and B. Tungodden (2020). Cutthroat capitalism versus cuddly socialism: Are Americans more meritocratic and efficiency-seeking than Scandanavians. *Journal of Political Economy 128*(5), 1753–1787.

Ambrus, A. and B. Greiner (2012). Imperfect public monitoring with costly punishment: An experimental study. *The American Economic Review 102*(7), 3317–3332.

Bicchieri, C. (2006). *The grammar of society.* Cambridge University Press.

Bochet, O., T. Page, and L. Putterman (2006). Communication and punishment in voluntary contribution experiments. *Journal of Economic Behavior & Organization 60*(1), 11–26.

Buckley, E. and R. Croson (2006). Income and wealth heterogeneity in the voluntary provision of linear public goods. *Journal of Public Economics 90*(4-5), 935–955.

Cappelen, A. W., A. D. Hole, E. Ø. Sørensen, and B. Tungodden (2007). The pluralism of fairness ideals: An experimental approach. *American Economic Review 97*(3), 818–827.

Carpenter, J. and P. H. Matthews (2009). What norms trigger punishment? *Experimental Economics 12*(3), 272–288.

Cason, T. N. and L. Gangadharan (2015). Promoting cooperation in nonlinear social dilemmas through peer punishment. *Experimental Economics 18*(1), 66–88.

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

Cinyabugama, M., T. Page, and L. Putterman (2006). Can second-order punishment deter perverse punishment? *Experimental Economics 9*, 265–279.

Coleman, J. S. (1990). *Foundations of social theory.* Harvard University Press.

De Geest, L. R. and D. C. Kingsley (2021). Norm enforcement with incomplete information. *Journal of Economic Behavior & Organization 189*(September), 403–430.

Egas, M. and A. Riedl (2008). The economics of altruistic punishment and the maintenance of cooperation. *Proceedings of the Royal Society of London B: Biological Sciences 275*(1637), 871–878.

Elster, J. (1989). Social norms and economic theory. *Journal of Economic Perspectives 3*(4), 99–117.

Fehr, D., H. Rau, S. T. Trautmann, and Y. Xu (2020). Inequality, fairness and social capital. *European Economic Review 129*, 980–994.

Fehr, E. and S. Gächter (2000). Cooperation and punishment in public goods experiments. *American Economic Review 90*(4), 980–994.

Fehr, E. and S. Gächter (2002). Altruistic punishment in humans. *Nature 415*(6868), 137–140.

Fehr, E. and I. Schurtenberger (2018). Normative foundations of human cooperation. *Nature Human Behaviour 2*(7), 458–468.

Fischbacher, U. (2007). z-tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics 10*(2), 171–178.

Frohlich, N., J. Oppenheimer, and A. Kurki (2004). Modeling other-regarding preferences and an experimental test. *Public Choice 119*(1-2), 91–117.

Gächter, S., E. Renner, and M. Sefton (2008). The long-run benefits of punishment. *Science 322*(5907), 1510–1510.

Grechenig, K., A. Nicklisch, and C. Thöni (2010). Punishment despite reasonable doubt? A public goods experiment with sanctions under uncertainty. *Journal of Empirical Legal Studies 7*(4), 847–867.

Hargreaves Heap, S. P., A. Ramalingam, and B. V. Stoddard (2016). Endowment inequality in public goods games: A re-examination. *Economics Letters 146*, 4–7.

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

Kingsley, D. C. and L. R. De Geest (2025). Property rights and public goods. *Journal of the Economic Science Association 11*, 233–243.

Knack, S. and P. Keefer (1997). Does social capital have an economic payoff? A cross-country investigation. *The Quarterly Journal of Economics 112*(4), 1251–1288.

Konow, J. (2000). Fair shares: Accountability and cognitive dissonance in allocation decisions. *American Economic Review 90*(4), 1072–1091.

Krupka, E. L. and R. A. Weber (2013). Identifying social norms using coordination games: Why does dictator game sharing vary? *Journal of the European Economic Association 11*(3), 495–524.

Ledyard, J. O. (1995). Public goods: A survey of experimental research. In J. H. Kagel and A. E. Roth (Eds.), *Handbook of Experimental Economics*. Princeton: Princeton University Press.

Nikiforakis, N. and H.-T. Normann (2008). A comparative statics analysis of punishment in public-good experiments. *Experimental Economics 11*(4), 358–369.

Nikiforakis, N., C. N. Noussair, and T. Wilkening (2012). Normative conflict and feuds: The limits of self-enforcement. *Journal of Public Economics 96*(9-10), 797–807.

Ostrom, E. (2000). Collective action and the evolution of social norms. *Journal of Economic Perspectives 14*(3), 137–158.

Ostrom, E., J. Walker, and R. Gardner (1992). Covenants with and without a sword: Self-governance is possible. *American Political Science Review 86*(02), 404–417.

Ostry, J. D., A. Berg, and C. G. Tsangarides (2014). Redistribution, inequality, and growth. *IMF Staff Discussion Note SDN/14/02*.

Putnam, R. D., R. Leonardi, and R. Y. Nanetti (1993). *Making Democracy Work: Civic Traditions in Modern Italy*. Princeton University Press.

Reuben, E. and A. Riedl (2013). Enforcement of contribution norms in public good games with heterogeneous populations. *Games and Economic Behavior 77*(1), 122–137.

Schelling, T. C. (1960). *The strategy of conflict*. Harvard University Press.

Suchon, R. and V. Théroude (2022). Inequality and cooperation: Meta-analytical evidence from public good experiments. *BETA Working Paper No. 2022-29*.

Visser, M. and J. Burns (2015). Inequality, social sanctions and cooperation within South African fishing communities. *Journal of Economic Behavior & Organization 118*, 95–109.

Vostroknutov, A. (2020). Social norms in experimental economics: Towards a unified theory of normative decision making. *Analyse & Kritik 42*(1), 3–40.

Weng, Q. and F. Carlsson (2015). Cooperation in teams: The role of identity, punishment, and endowment distribution. *Journal of Public Economics 126*, 25–38.

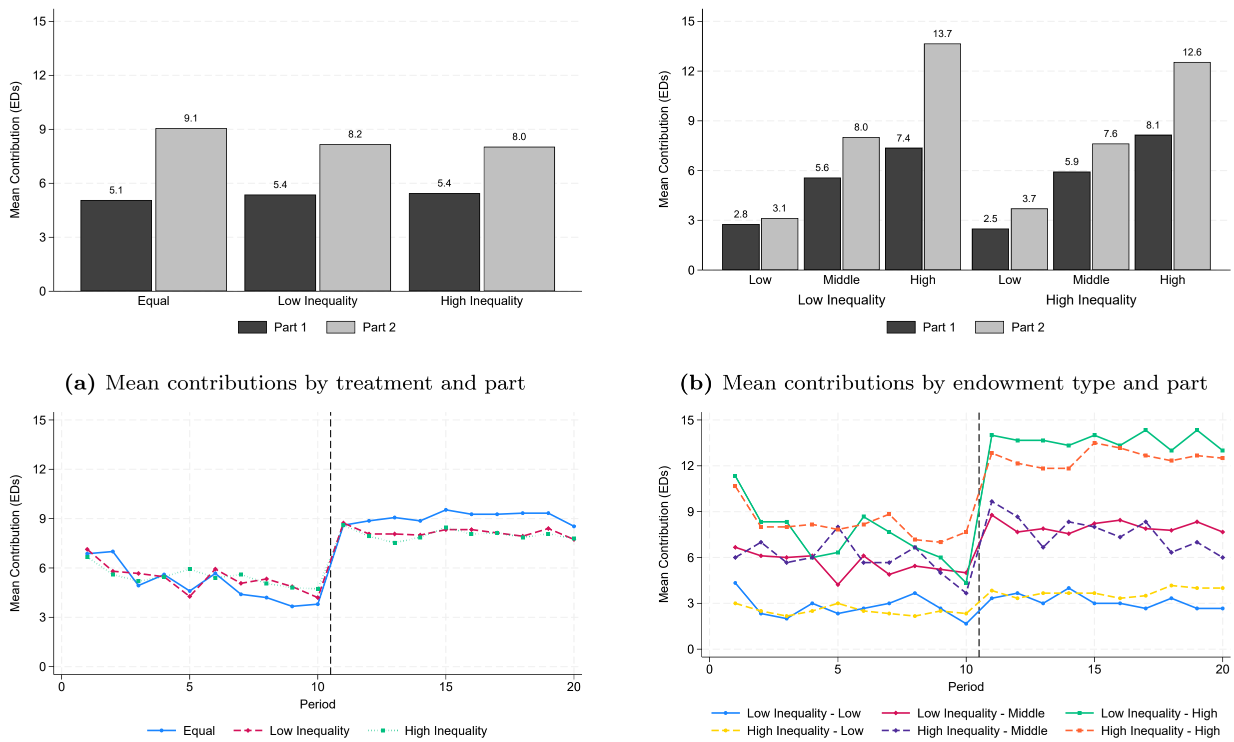
Young, H. P. (2008). Social norms. In *The New Palgrave Dictionary of Economics*, Volume 7, pp. 647–651. Palgrave Macmillan.

Zak, P. J. and S. Knack (2001). Trust and growth. *The Economic Journal 111*(470), 295–321.

Zelmer, J. (2003). Linear public good experiments: A meta-analysis. *Experimental Economics 6*(3), 299–310.

A Appendix

A.1 Contributions



(a) Mean contributions by treatment and part

(b) Mean contributions by endowment type and part

(c) Contributions over time by treatment

(d) Contributions over time by endowment type

Figure A.1: Average contributions in Experimental Dollars (EDs). Panels A and C summarize contributions by treatment; Panels B and D summarize contributions by endowment type in the unequal treatments. Vertical dashed lines mark the transition from Part 1 (periods 1–10) to Part 2 (periods 11–20).

Figure A.1 summarizes contribution patterns by treatment, endowment type, and period. Contributions increase after punishment is introduced in Part 2 and then remain relatively stable, suggesting that punishment primarily produces a level shift. In the unequal treatments, each endowment type increases contributions in Part 2, with the largest increases among High types.

Table A.1: Group-Level and By-Type Contributions

Panel A. Group-level contributions							
	Mean Contribution			Across	Pairwise Across		
Phase	Equal	Low Ineq.	High Ineq.	KW	Eq.–Low	Eq.–High	Low–High
Part 1	5.07	5.37	5.45	0.114	0.085	0.062	0.983
	(1.05)	(1.03)	(1.01)				
Part 2	9.07	8.17	8.03	0.011	0.011	0.009	0.803
	(0.88)	(1.30)	(1.39)				
Within	0.001	0.001	0.002				
Obs.	15	15	15				
Panel B. Contributions by endowment type in unequal treatments							
	Low Ineq.			High Ineq.			Across
Type	Part 1	Part 2	Within	Part 1	Part 2	Within	Part 1 / Part 2
Low	2.77	3.13	0.391	2.50	3.72	0.001	0.559 / 0.205
	(1.70)	(1.33)		(1.02)	(1.40)		
Middle	5.58	8.02	0.001	5.93	7.63	0.020	0.803 / 0.632
	(1.44)	(1.42)		(2.45)	(2.05)		
High	7.37	13.67	0.001	8.15	12.55	0.002	0.575 / 0.307
	(3.59)	(4.35)		(2.86)	(3.36)		
Obs.	15	15	15	15	15	15	

The unit of observation in Panel A is the group. The unit of observation in Panel B is the group-by-endowment-type cell. Entries report mean contributions, with standard deviations in parentheses. Columns labeled "Within" report p-values from Wilcoxon signed-rank tests comparing Part 1 and Part 2 within the relevant treatment or treatment-type cell. The column labeled "Across" in Panel A reports the Kruskal–Wallis p-value for equality across all three treatments within phase. Pairwise across-treatment p-values in Panel A are from Wilcoxon rank-sum tests. In Panel B, the "Across" column reports Wilcoxon rank-sum p-values comparing Low Inequality and High Inequality within endowment type; the first value is for Part 1 and the second is for Part 2.

Table A.1 shows that the introduction of punishment increases contributions in all treatments. However, the across-treatment comparisons show that contributions are significantly lower in both inequality treatments than in the equal treatment during the punishment phase, while the two inequality treatments do not differ significantly from each other. The by-type results show that this pattern is not driven by a significant difference within any endowment type. Contributions generally increase within endowment type after punishment is introduced, although the increase for low-endowment members in *Low Inequality* is not statistically significant.

Table A.2: Robustness Checks for Proportional Contributions

	(1)	(2)
Low Inequality	–	0.009
		(0.015)
High Inequality	–	0.012
		(0.012)
Punishment phase (Part 2)	0.152***	0.141***
	(0.018)	(0.015)
Low Inequality × Part 2	-0.062***	-0.051***
	(0.022)	(0.019)
High Inequality × Part 2	-0.065***	-0.051***
	(0.020)	(0.017)
VCM period trend	-0.006***	-0.005***
	(0.002)	(0.002)
Punishment period trend	-0.003*	-0.005***
	(0.002)	(0.002)
Lagged own prop. cont.	0.057***	0.267***
	(0.021)	(0.031)
Lagged mean others' prop. cont.	0.163***	0.102**
	(0.043)	(0.046)
Age		-0.001
		(0.001)
Gender		0.000
		(0.008)
Semesters		-0.001
		(0.003)
Economics		-0.002
		(0.002)
Constant	0.235***	0.203***
	(0.021)	(0.040)
Observations	4,275	4,142
Subjects	225	218
F/Wald statistic	123.20***	1815.45***

Column (1) reports subject fixed-effects estimates. Column (2) reports random-effects estimates with demographic controls. The dependent variable is individual contribution measured as a proportion of endowment. Standard errors clustered at the group level are reported in parentheses. *Equal* is the omitted treatment in Column (2). Treatment main effects are absorbed by subject fixed effects in Column (1). Demographic controls include age, gender, semesters completed, and economics courses. The final row reports the model F statistic in Column (1) and Wald statistic in Column (2). *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table A.2 reports robustness checks for the contribution regression presented in Table 3. The treatment-by-punishment-phase interactions remain negative and statistically significant with subject fixed effects and after adding demographic controls. Thus, the finding that peer punishment increases proportional contributions less under inequality is robust to absorbing subject-level heterogeneity and controlling for observed demographics.

Table A.3: Proportional Contributions by Treatment

	Treatment Mean			Rank-Sum p-value		
	Equal	Low Ineq.	High Ineq.	Eq.–Low	Eq.–High	Low–High
No punishment	0.254	0.272	0.268	0.071	0.146	0.431
Punishment	0.453	0.394	0.392	0.004	0.010	0.836

The unit of observation is the group. Proportional contribution is contribution divided by endowment. Pairwise p-values are from Wilcoxon rank-sum tests using group-level averages.

Table A.4: Proportional Contributions by Endowment Type in Unequal Treatments

		Treatment Mean		Rank-Sum p-value
Phase	Endowment Type	Low Ineq.	High Ineq.	Low–High
No punishment	Low	0.277	0.250	0.572
	Middle	0.279	0.297	0.814
	High	0.246	0.272	0.616
Punishment	Low	0.313	0.372	0.212
	Middle	0.401	0.382	0.644
	High	0.456	0.418	0.317

The unit of observation is the group-by-endowment-type average. Proportional contribution is contribution divided by endowment. Pairwise p-values are exact Wilcoxon rank-sum tests comparing Low and High Inequality.

As a nonparametric complement to the regression analysis presented in Table 3, Table A.3 reports group-level average proportional contributions by treatment and phase. The table shows little evidence of treatment differences before punishment is introduced. In the punishment phase, proportional contributions are significantly higher in *Equal* than in either unequal treatment, while the two unequal treatments are nearly identical. Table A.4 reports the same proportional contribution measure separately by endowment type within the unequal treatments. The type-level comparisons show no statistically significant differences across the unequal treatments for Low, Middle, or High types in either phase.

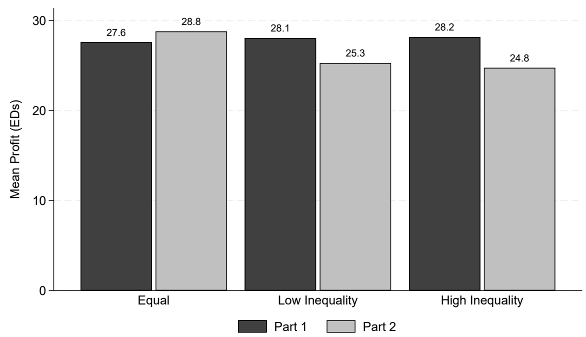
Table A.5: Post-Estimation Comparisons for Proportional Contributions

	Estimate	p-value
Panel A. Predicted proportional contributions		
Equal, Part 1	0.2723	< 0.001
Equal, Part 2	0.4181	< 0.001
Low Inequality, Part 1	0.2838	< 0.001
Low Inequality, Part 2	0.3779	< 0.001
High Inequality, Part 1	0.2864	< 0.001
High Inequality, Part 2	0.3765	< 0.001
Panel B. Pairwise comparisons within Part 1		
Low Inequality – *Equal*	0.0115	0.417
High Inequality – *Equal*	0.0141	0.244
High Inequality – *Low Inequality*	0.0026	0.840
Panel C. Pairwise comparisons within Part 2		
Low Inequality – *Equal*	-0.0401	0.002
High Inequality – *Equal*	-0.0416	0.004
High Inequality – *Low Inequality*	-0.0015	0.933
Panel D. Change from Part 1 to Part 2		
Equal	0.1458	< 0.001
Low Inequality	0.0941	< 0.001
High Inequality	0.0900	< 0.001
Panel E. Difference-in-differences		
Low Inequality – *Equal*	-0.0517	0.005
High Inequality – *Equal*	-0.0557	0.001
High Inequality – *Low Inequality*	-0.0041	0.851
Joint test of inequality treatment differences	$\chi^2(2) = 14.67$	0.001

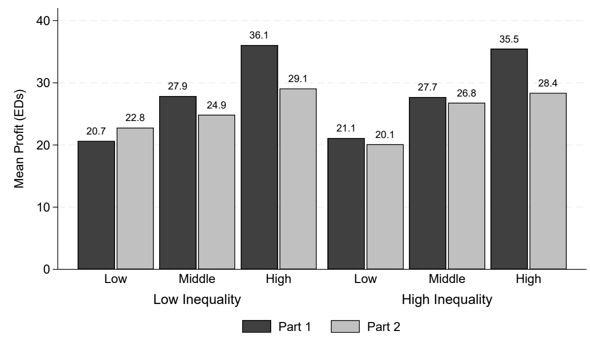
This table reports post-estimation results from the random-effects GLS model in Table 3. The dependent variable is individual contribution measured as a proportion of endowment. Part 1 is the VCM phase and Part 2 is the punishment phase. Panel A reports regression-adjusted predictive margins by treatment and part. Panels B and C report pairwise comparisons within each part. Panel D reports the estimated change from Part 1 to Part 2 within each treatment. Panel E reports difference-in-differences estimates comparing this change across treatments. Standard errors are clustered at the group level in the underlying model.

Table A.5 reports the post-estimation comparisons underlying the contribution results presented in Figure 1. The estimated increase from Part 1 to Part 2 is positive in all treatments, but significantly smaller in both inequality treatments than in *Equal*.

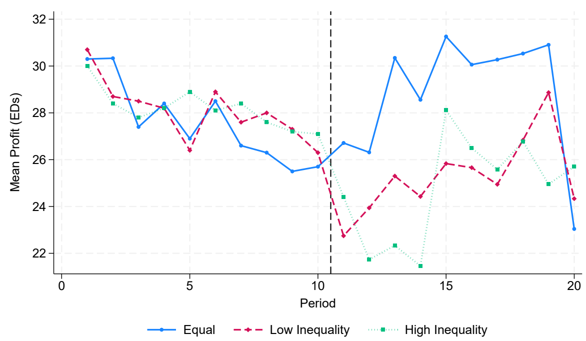
A.2 Earnings



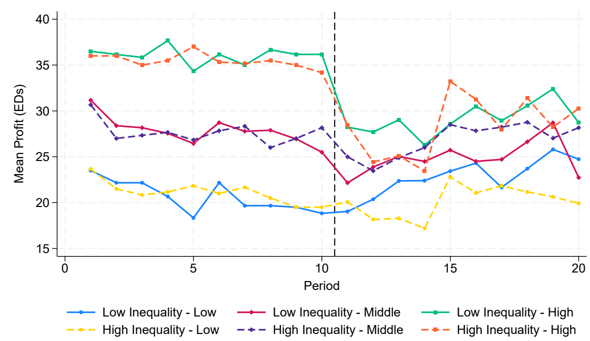
(a) Mean earnings by treatment and part



(b) Mean earnings by endowment type and part



(c) Earnings over time by treatment



(d) Earnings over time by endowment type

Figure A.2: Average earnings in Experimental Dollars (EDs). Panels A and C summarize earnings by treatment; Panels B and D summarize earnings by endowment type in the unequal treatments. Vertical dashed lines mark the transition from Part 1 (periods 1–10) to Part 2 (periods 11–20).

Figure A.2 summarizes earnings by treatment, endowment type, and period. Average earnings are similar across treatments in Part 1, consistent with the similarity in contribution behavior before punishment is introduced. In Part 2, earnings are lower in the unequal treatments than in *Equal*, reflecting both weaker contribution gains and the costs of punishment.

Table A.6: Group-Level and By-Type Earnings

Panel A. Group-level earnings

	Mean Earnings			Across	Pairwise Across		
Phase	Equal	Low Ineq.	High Ineq.	KW	Eq.–Low	Eq.–High	Low–High
Part 1	27.59	28.06	28.17	0.105	0.074	0.062	0.983
	(1.56)	(1.54)	(1.51)				
Part 2	28.80	25.29	24.76	0.028	0.040	0.015	0.443
	(5.19)	(5.01)	(3.39)				
Within	0.256	0.100	0.009				
Obs.	15	15	15				

Panel B. Earnings by endowment type in unequal treatments

	Low Ineq.			High Ineq.			Across
Type	Part 1	Part 2	Within	Part 1	Part 2	Within	Part 1 / Part 2
Low	20.67	22.78	0.281	21.12	20.11	0.394	0.493 / 0.171
	(2.80)	(6.16)		(2.84)	(4.06)		
Middle	27.86	24.86	0.100	27.68	26.79	0.281	0.361 / 0.330
	(1.69)	(5.20)		(3.14)	(4.46)		
High	36.07	29.09	0.001	35.47	28.39	0.001	0.740 / 0.619
	(3.48)	(5.45)		(1.55)	(4.92)		
Obs.	15	15	15	15	15	15	

The unit of observation in Panel A is the group. The unit of observation in Panel B is the group-by-endowment-type cell. Entries report mean earnings, with standard deviations in parentheses. Columns labeled "Within" report p-values from Wilcoxon signed-rank tests comparing Part 1 and Part 2 within the relevant treatment or treatment-type cell. The column labeled "Across" in Panel A reports the Kruskal–Wallis p-value for equality across all three treatments within phase. Pairwise across-treatment p-values in Panel A are from Wilcoxon rank-sum tests. In Panel B, the "Across" column reports Wilcoxon rank-sum p-values comparing Low Inequality and High Inequality within endowment type; the first value is for Part 1 and the second is for Part 2.

Table A.6 shows that the earnings consequences of punishment differ across treatments. At the group level, earnings do not significantly change from Part 1 to Part 2 in *Equal*, decline marginally in *Low Inequality*, and decline significantly in *High Inequality*. Across treatments, earnings are significantly lower in both unequal treatments than in *Equal* during the punishment phase, while the two inequality treatments do not differ significantly from each other. The by-type results indicate that the earnings decline is concentrated among high-endowment members, whose earnings fall significantly in both inequality treatments after punishment is introduced.

Table A.7: Robustness Checks for Earnings

	(1)	(2)
Low Inequality	–	0.365
		(0.583)
High Inequality	–	0.284
		(0.578)
Punishment phase (Part 2)	-2.130	-2.280
	(1.556)	(1.567)
Low Inequality × Part 2	-3.979**	-3.910**
	(1.832)	(1.841)
High Inequality × Part 2	-4.620***	-4.591***
	(1.550)	(1.574)
VCM period trend	-0.346***	-0.352***
	(0.054)	(0.055)
Punishment period trend	0.260**	0.271**
	(0.123)	(0.124)
Age		-0.224**
		(0.109)
Gender		0.206
		(0.576)
Semesters		0.061
		(0.160)
Economics		0.123
		(0.124)
Constant	29.846***	33.987***
	(0.440)	(2.351)
Observations	4,500	4,360
Subjects	225	218
F/Wald statistic	13.89***	114.08***

Column (1) reports subject fixed-effects estimates. Column (2) reports random-effects estimates with demographic controls. The dependent variable is individual period earnings, net of punishment costs paid and received. Standard errors clustered at the group level are reported in parentheses. *Equal* is the omitted treatment in Column (2). Treatment main effects are absorbed by subject fixed effects in Column (1). Part 2 indicates the punishment phase. Demographic controls include age, gender, semesters completed, and economics courses. The final row reports the model F statistic in Column (1) and Wald statistic in Column (2). *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table A.7 reports robustness checks for the earnings regression presented in Table 4. The treatment-by-punishment-phase interactions remain negative and statistically significant with subject fixed effects and after adding demographic controls. Thus, the finding that earnings fall more sharply under inequality when punishment is introduced is robust to absorbing time-invariant subject characteristics and controlling for observed demographics.

Table A.8: Post-Estimation Comparisons for Earnings

	Estimate	p-value
Panel A. Predicted earnings		
Equal, Part 1	29.2621	< 0.001
Equal, Part 2	27.1319	< 0.001
Low Inequality, Part 1	29.7288	< 0.001
Low Inequality, Part 2	23.6199	< 0.001
High Inequality, Part 1	29.8388	< 0.001
High Inequality, Part 2	23.0886	< 0.001
Panel B. Pairwise comparisons within Part 1		
Low Inequality – Equal	0.4667	0.399
High Inequality – Equal	0.5767	0.293
High Inequality – Low Inequality	0.1100	0.840
Panel C. Pairwise comparisons within Part 2		
Low Inequality – Equal	-3.5120	0.054
High Inequality – Equal	-4.0433	0.010
High Inequality – Low Inequality	-0.5313	0.728
Panel D. Change from Part 1 to Part 2		
Equal	-2.1302	0.171
Low Inequality	-6.1089	< 0.001
High Inequality	-6.7502	< 0.001
Panel E. Difference-in-differences		
Low Inequality – Equal	-3.9787	0.030
High Inequality – Equal	-4.6200	0.003
High Inequality – Low Inequality	-0.6413	0.701
Joint test of inequality treatment differences	$\chi^2(2) = 9.40$	0.009

This table reports post-estimation results from the random-effects GLS model in Table 4. The dependent variable is individual period earnings, net of punishment costs paid and received. Part 1 is the VCM phase and Part 2 is the punishment phase. Panel A reports regression-adjusted predictive margins by treatment and part. Panels B and C report pairwise comparisons within each part. Panel D reports the estimated change from Part 1 to Part 2 within each treatment. Panel E reports difference-in-differences estimates comparing this change across treatments. Standard errors are clustered at the group level in the underlying model.

Table A.8 reports the post-estimation comparisons underlying the earnings results in Figure 2. The change from Part 1 to Part 2 is not statistically significant in *Equal*, but earnings fall significantly in both unequal treatments.

A.3 Punishment, Norms, and Enforcement

A.3.1 Punishment Benchmark

Table A.9: Winning-Benchmark Probit Regressions

	Equal		*Low Inequality*			*High Inequality*	
Target endowment	20	10	20	30	10	20	30
Winning benchmark	10	5	10	15	5	10	15
Below benchmark	2.02***	0.53*	1.63***	1.41***	1.44***	1.71***	2.69***
	(0.19)	(0.32)	(0.24)	(0.31)	(0.33)	(0.33)	(0.37)
Punisher prop. cont.	−0.25	−0.43	−1.24***	−1.09	−0.86**	−0.58	−1.62**
	(0.40)	(1.01)	(0.46)	(0.82)	(0.43)	(0.46)	(0.74)
Total group cont.	−0.01	−0.00	0.02*	0.01	0.02	0.03***	0.07***
	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)
Lagged pun. rec.	0.03***	0.02	0.01***	0.01	0.01	−0.01	0.02**
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)
Pun. period trend	0.01	−0.05	−0.02	−0.04	−0.06	−0.10***	−0.06*
	(0.02)	(0.06)	(0.02)	(0.04)	(0.04)	(0.03)	(0.04)
Constant	−1.86***	−1.22	−2.10***	−1.21	−1.55***	−1.33*	−3.64***
	(0.56)	(1.04)	(0.68)	(0.97)	(0.58)	(0.78)	(0.99)
Ave. M.E.	0.23***	0.06*	0.29***	0.25***	0.22***	0.28***	0.44***
Below benchmark	(0.02)	(0.03)	(0.04)	(0.06)	(0.05)	(0.05)	(0.06)
Observations	2700	540	1620	540	1080	540	1080
Punishers	75	60	75	60	75	60	75
Groups	15	15	15	15	15	15	15
Log likelihood	−608.48	−128.53	−570.85	−213.04	−367.35	−187.69	−385.51

The unit of observation is a punisher–target–period punishment opportunity. Entries report random-effects probit estimates with punisher-level random effects and robust standard errors clustered at the group level. The dependent variable equals one if the punisher assigns any punishment to the target. The below-benchmark indicator equals one when the target's contribution falls below the winning benchmark identified in Figure 3. Average marginal effects for the below-benchmark indicator are reported near the bottom of the table. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table A.9 reports the probit regressions for the winning proportional benchmarks identified in Figure 3. The below-benchmark indicator is positive in every treatment and target-endowment cell, indicating that targets are more likely to be punished when their contribution falls below one-half of their endowment.

Table A.10: Post-Estimation Comparisons for Punishment Imposed

	Estimate	p-value
Panel A. Predicted punishment cost imposed		
Equal	1.209	< 0.001
Low Inequality	1.793	< 0.001
High Inequality	1.807	< 0.001
Panel B. Treatment differences		
Low Inequality $-$ *Equal*	0.584	0.133
High Inequality $-$ *Equal*	0.597	0.062
Joint test of treatment differences from *Equal*	$\chi^2(2) = 3.82$	0.148

This table reports predicted margins and treatment contrasts from the random-effects GLS model in Table 5. The dependent variable is the cost paid by a subject to impose punishment on other group members in a period. The sample is restricted to the punishment phase, periods 11–20. Standard errors are clustered at the group level in the underlying model.

Table A.10 reports the post-estimation comparisons underlying Figure 4. Predicted punishment expenditures are moderately higher in both unequal treatments than in *Equal*, although the treatment differences are not jointly statistically significant. These results provide no evidence that the smaller contribution gains under inequality are driven by less punishment use. Instead, subjects in unequal groups appear to devote at least as many resources to punishment without achieving comparable improvements in contributions or earnings.

Table A.11: Group-Level and By-Type Punishment

Panel A. Group-level punishment sent

	Punishment Sent			Across	Pairwise Across		
Phase	Equal	Low Ineq.	High Ineq.	KW	Eq.–Low	Eq.–High	Low–High
Part 2	1.21	1.79	1.81	0.116	0.115	0.059	0.534
	(1.03)	(1.14)	(0.73)				
Obs.	15	15	15				

Panel B. Punishment by endowment type in unequal treatments

	Punishment Sent			Punishment Received		
Type	Low Ineq.	High Ineq.	Across	Low Ineq.	High Ineq.	Across
Low	1.91	1.77	0.507	2.64	4.39	0.018
	(2.16)	(1.06)		(3.08)	(2.45)	
Middle	1.88	1.25	0.146	5.88	4.42	0.110
	(1.37)	(0.90)		(4.08)	(4.75)	
High	1.40	2.12	0.383	6.62	6.95	0.678
	(0.76)	(1.53)		(5.18)	(4.48)	
Obs.	15	15		15	15	

The unit of observation in Panel A is the group. The unit of observation in Panel B is the group-by-endowment-type cell. Entries report mean punishment, with standard deviations in parentheses. Punishment is observed only in Part 2. The column labeled "Across" in Panel A reports the Kruskal–Wallis p-value for equality across all three treatments. Pairwise across-treatment p-values in Panel A are from Wilcoxon rank-sum tests. In Panel B, the columns labeled "Across" report Wilcoxon rank-sum p-values comparing Low Inequality and High Inequality within endowment type.

Table A.11 summarizes punishment sent and received at the group and endowment-type levels. At the group level, punishment sent is higher in both unequal treatments than in *Equal*, although the across-treatment difference is not statistically significant overall. The pairwise comparison between *Equal* and *High Inequality* is marginally significant. By endowment type, there are no significant differences between the unequal treatments in punishment sent. Punishment received is also generally similar across the two unequal treatments, with one exception: low-endowment members receive significantly more punishment in *High Inequality* than in *Low Inequality*.

A.3.2 Targeting and Intensity of Punishment

Table A.12: Post-Estimation Results: Punishment of Negative Deviations

	Estimate	Std. Error	p-value
Panel A. Effect of negative deviation on probability of receiving punishment			
Equal	0.202	0.024	< 0.001
Low Inequality	0.079	0.015	< 0.001
High Inequality	0.075	0.016	< 0.001
Panel B. Cross-treatment differences in probability effects			
Low Inequality − *Equal*	-0.123	0.029	< 0.001
High Inequality − *Equal*	-0.127	0.030	< 0.001
Joint test of treatment differences		$\chi^2(2) = 21.63$, $p < 0.001$	
Panel C. Effect of negative deviation on punishment cost received			
Equal	3.456	0.471	< 0.001
Low Inequality	0.942	0.309	0.002
High Inequality	1.775	0.368	< 0.001
Panel D. Cross-treatment differences in punishment-cost effects			
Low Inequality − *Equal*	-2.513	0.575	< 0.001
High Inequality − *Equal*	-1.681	0.592	0.005
Joint test of treatment differences		$\chi^2(2) = 19.15$, $p < 0.001$	

This table reports post-estimation estimates from the specifications in Table 6. Panels A and B report average marginal effects from the random-effects probit model for whether the subject receives any punishment. Panels C and D report marginal effects from the random-effects GLS model for punishment cost received, conditional on receiving positive punishment. Negative deviation is measured as the proportional shortfall below one-half of the subject's endowment and rescaled so that one unit equals a 10 percentage-point shortfall from the proportional benchmark. The sample is restricted to the punishment phase, periods 11–20.

Table A.12 reports the treatment-specific effects of negative deviations from the proportional benchmark. A 10 percentage-point larger shortfall from the proportional benchmark significantly increases both the probability of receiving punishment and the amount of punishment received in every treatment. However, both effects are significantly smaller in the inequality treatments than in *Equal*, confirming that deviations from the proportional norm are sanctioned less strongly when endowments are unequal.

Table A.13: Linear Probability Model: Punishment of Negative Deviations from the Proportional Norm

	Any punishment
Low Inequality	0.123
	(0.092)
High Inequality	0.184**
	(0.092)
Negative deviation from proportional norm	0.189***
	(0.029)
Low Inequality × Negative deviation	-0.102***
	(0.033)
High Inequality × Negative deviation	-0.112***
	(0.033)
Punishment period trend	-0.021***
	(0.005)
Total group contribution	-0.000
	(0.003)
Constant	0.353**
	(0.169)
Observations	2,250
Subjects	225
Wald χ^2	124.55***

The table reports random-effects linear probability estimates for whether a subject receives any punishment in a period. Standard errors are clustered at the group level and reported in parentheses. *Equal* is the omitted treatment. Negative deviation from the proportional norm is measured as $\max(0, 0.5 - x_i/e_i)$ and rescaled so that one unit equals a 10 percentage-point shortfall from the proportional benchmark. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table A.13 reports a linear probability model corresponding to the random-effects probit specification in Table 6. The results are very similar to the average marginal effects from the probit model. In the *Equal* treatment, a 10 percentage-point larger shortfall from the proportional benchmark increases the probability of punishment by 18.9 percentage points. The corresponding effects are 8.6 percentage points in *Low Inequality* and 7.7 percentage points in *High Inequality*. Both inequality-treatment effects are significantly smaller than the effect in *Equal*, while the two unequal treatments are similar. Thus, the conclusion that inequality weakens the enforcement of the proportional benchmark is robust to estimating the extensive-margin punishment model as a linear probability model.

A.3.3 Response to Punishment

Table A.14: Post-Estimation Results: Response to Punishment

	Estimate	Std. Error	p-value
Panel A. Effect of lagged punishment received: below proportional norm			
Equal	0.00345	0.00130	0.008
Low Inequality	0.00054	0.00139	0.698
High Inequality	0.00016	0.00129	0.903
Panel B. Cross-treatment differences: below proportional norm			
Low Inequality – Equal	-0.00291	0.00197	0.139
High Inequality – Equal	-0.00329	0.00170	0.053
Joint test of treatment differences	$\chi^2(2) = 4.17$, $p = 0.124$		
Panel C. Effect of lagged punishment received: at or above proportional norm			
Equal	0.00163	0.00199	0.411
Low Inequality	0.00080	0.00139	0.565
High Inequality	-0.00076	0.00171	0.657
Panel D. Cross-treatment differences: at or above proportional norm			
Low Inequality – Equal	-0.00083	0.00234	0.721
High Inequality – Equal	-0.00239	0.00264	0.365
Joint test of treatment differences	$\chi^2(2) = 0.91$, $p = 0.634$		

This table reports post-estimation estimates from the specifications in Table 7. Panels A and B report treatment-specific effects and cross-treatment contrasts for subjects whose proportional contribution in period $t-1$ was below 0.5. Panels C and D report the corresponding estimates for subjects whose proportional contribution in period $t-1$ was at or above 0.5. The dependent variable is the change in the proportion of a subject's endowment contributed from period $t-1$ to period t. The punishment variable is punishment received in period $t-1$. The sample is restricted to punishment-phase observations for periods 12–20.

Table A.14 reports the post-estimation comparisons for the response-to-punishment regressions presented in Table 7. Among subjects who were below the proportional norm, punishment received has a positive and statistically significant effect on subsequent changes in proportional contributions in *Equal*. The corresponding effects in both inequality treatments are close to zero and statistically insignificant. The estimated response is significantly smaller in *High Inequality* relative to *Equal* at the 10 percent level, while the difference between *Low Inequality* and *Equal* is not statistically significant. Among subjects who were at or above the proportional norm, punishment received has no statistically significant effect on subsequent proportional contributions in any treatment, and there is no evidence of cross-treatment differences.