Team Incentives and Leadership

Michalis Drouvelis[*]

Daniele Nosenzo[**]

Martin Sefton[**]

Abstract
We study, experimentally, how two alternative incentive mechanisms affect team performance and how a team chooses between alternative mechanisms. We study a group incentive mechanism (team output is shared equally among team members) and a hierarchical mechanism (team output is allocated by a team leader). We find that output is higher when a leader has the power to allocate output, but this mechanism also generates large differences between earnings of leaders and other team members. When team members can choose how much of team output is to be shared equally and how much is to be allocated by a leader they tend to restrict the leader's power to distributing less than half of the pie.

Keywords: Team Production; Leadership; Reward Power; Delegation; Experiment.

JEL codes: C92, L22, M5

Acknowledgements: We thank the Editor, Ananish Chaudhuri, and three anonymous referees for helpful comments. We also received useful suggestions from workshop participants at the ETH Social Norms and Institutions Conference and University of Amsterdam. We acknowledge support from the University of Birmingham and the Economic and Social Science Research Council (ES/K002201/1)

[*] Department of Economics, University of Birmingham.
[**] School of Economics, University of Nottingham.

1. Introduction

Many organizations are hierarchical in structure, where leaders have power over subordinates and can influence them, and hence organizational performance, in a variety of ways. Vast literatures in management and social psychology have studied various dimensions of leadership. Bass and Bass (2008) extensively survey the literatures on leadership and broadly classify theories of leadership as *informal* (dealing with the emergence of leaders who lack formal authority), *inspirational* (focusing on leaders' ideological or emotional appeals to followers' intrinsic motivation), or *instrumental* (focusing on outcome-directed leaders who have formal disciplining powers). Our paper contributes to the latter class of theories, in particular by examining the role of *contingent-positive reinforcement*, whereby a leader encourages compliance from subordinates by appealing to their self-interested response to material rewards. In fact, *reward power* is often recognized as a crucial dimension of leadership (see, e.g., French and Raven, 1959; Yukl, 1989) and is indeed a cardinal principle of the path-goal and operant conditioning approaches to leadership, that posit that whether a leader can successfully motivate followers depends on her ability to make rewards contingent on followers' performance (Jago, 1982). Relatedly, Hermalin (2013) surveys the leadership literature from an economic perspective and discusses various roles of leaders, one of which is to be responsible for monitoring and administering incentives within a group.

Although the use of rewards may successfully motivate subordinates, the availability of leadership reward power may introduce other sorts of incentive problems. Most obviously, opportunistic leaders may have an incentive to abuse their power and use group resources to advance their private interest. This threat of opportunism may have less force when the allocation of power is endogenous within the organization, e.g. when subordinates have a say on how much power is invested in the leader. In such cases, will leaders resist the temptation to abuse their power, and will subordinates be able to correctly anticipate the benefits, as well as potential perils, of leadership, and thus voluntarily grant power to the leader? This is an important question, especially because in some theories of leadership (e.g., Fiedler, 1967's contingency model of leadership) the effectiveness of a leader does not merely depend on her traits and behaviors, but also on the "favorableness of the leadership situation", including the availability of *position power*, i.e. the extent to which the leader is vested with authority to *"… direct, evaluate, reward and punish group members*" (Jago, 1982; p. 323).

In this paper we examine these issues using the methodology of experimental economics. Thus, we contribute to the existing management and social psychology literature, that have documented positive correlations between contingent-positive reinforcements and subordinates' performance (for reviews see, e.g., Podsakoff and Schriesheim, 1985; Podsakoff et al., 2006), by designing tightly-controlled laboratory experiments that allow for causal inferences and where subjects' decisions are elicited in an incentive-compatible way.[1]

Our experimental design begins with a 10-round repeated team production game where team members incur individual effort costs but share team output equally with all team members. Since the benefits of a team member's efforts are shared with the rest of the team this introduces an externality that will result in excessive shirking and welfare loss if decisions are guided by a comparison of private costs and benefits. Thus, our model of team production follows the tradition of using a Voluntary Contributions Mechanism (VCM) to capture the essence of the free-rider problem in teams (see Charness and Kuhn, 2011 for a review of this approach).[2] We observe substantial free-riding in this baseline treatment, in line with the large experimental literature on VCM games (see for example the recent review in Camerer and Weber, 2012).

Next, we ask whether installing a leader who allocates rewards to team members improves team performance. To do this we conduct treatments that complement and extend the recent experimental work on leadership with distributive power by Heijden et al. (2009) and Stoddard et al. (2014). In these treatments all team output accrues to a leader, who can decide how to distribute it after observing individual team members' efforts. Importantly, any output not allocated to other team members is retained by the leader. In this setting, a leader might induce efficient team production by compensating team members appropriately for the costs they incur from their productive efforts, and furthermore she has an incentive to do so as efficient team production will increase her residual claim. However, leaders also have an incentive to appropriate all the team output for themselves. In theory, assuming standard selfish preferences, a leader will keep all team output and, in anticipation of this, team members supply minimum effort. In contrast to this theoretical prediction, but in line with previous experimental work, in our experiment we find that installing a leader does indeed promote effort and increase efficiency. Leaders use simple strategies that reward workers

[1] A word of caution about external validity is necessary. In order to observe behavior in a controlled environment we use abstract settings that remove many of the complexities present in the field. Moreover, in natural organizations individuals self-select into leadership positions, a feature that is absent in our study. Thus, caution should be exercised in extrapolating insights from our lab setting to naturally-occurring environments.

[2] See also Guillen et al. (2014) for a discussion of the close relation between team production and VCM games. Bartling et al. (2010) provide a behavioral foundation for the use of the equal sharing rule in team production settings.

who supply high effort and withhold rewards from shirkers. This in turn encourages effort and results in substantial increases in team production and earnings.

Successful leadership may be more challenging when, as in many natural settings, workers vary in their productivity. Indeed, related experiments have shown that asymmetries between workers reduce the effectiveness of other forms of leader power (Levati et al. 2007). In our context, what constitutes "compensating team members appropriately" may be less straightforward if productivities vary among workers. Should compensation reflect the costs that a worker incurs from her efforts, or the output that she produces? If workers are concerned about equity and fairness, and if there are competing notions of fairness, it may be particularly difficult to provide the correct incentives. Thus, we also ran treatments with heterogeneous worker productivities. Again, we find low effort in the absence of a leader and substantially higher effort and efficiency with a leader. Thus, just as in the case of homogeneous teams, with heterogeneous teams we find that installing a leader with power to distribute the proceeds of team production is successful in promoting efficiency.

However, in both homogeneous and heterogeneous teams, we also find that the gains of leadership are distributed asymmetrically within a team: leaders reap most of the gains, but team members are no better off with than without a leader. This raises the question whether team members would actually *prefer* to install a leader if they could choose to do so.[3]

This question cannot be addressed in our initial treatments because a feature of these treatments is that the institutional setting – either a group incentive scheme or a leader reward scheme – is exogenously imposed on a team as part of our experimental design. Therefore, we designed further treatments to examine the endogenous emergence of the leadership institution. In these treatments we ask whether in our setting team members will voluntarily cede reward power to a leader, and whether this affects the leader's performance in terms of encouraging team production. To do this, we allowed team members to repeatedly decide what proportion of team output will be given to the leader to distribute, with the remainder shared equally among the team members.

Making the leader's power endogenous in this way could represent an obstacle to successful leadership. Even if a leader would be willing and able to compensate workers, team members may not support leadership because they fail to anticipate the leader's behavior, or because they prefer to retain some control over part of their earnings from

[3] This is a key question also in the social psychology and management literatures, which have studied extensively the conditions under which group members prefer to appoint a leader to solve the free-rider problem, see, e.g., Messick and Brewer (1983); Samuelson (1991); van Vugt and De Cremer (1999).

providing effort. This may reflect a non-pecuniary benefit from partially controlling incentives, similar to the desire to retain authority seen in other studies of principal-agent relationships (e.g., Bartling and Fischbacher, 2012; Fehr et al., 2013; Bartling et al., 2014). On the other hand, the endogenous emergence of leadership may facilitate its success. First, it gives team members an opportunity to punish leaders who abuse their power, and so it may be a useful mechanism for constraining opportunistic leaders.[4] Second, as seen in other studies of endogenous institutions (e.g., Dal Bó et al., 2010; Sutter et al., 2010) giving agents a voice in the institution may in and of itself foster a more cooperative environment.

We find that, as in our previous treatments, team production and earnings are higher when leaders are given more power. However, despite the success of the leadership institution when it emerges, we find that team members delegate too little power to leaders and so the potential gains from leadership are not realized. The main reason for this appears to be the way the benefits of leadership are shared. Although the leaders' rewarding strategies are well-calibrated to make it pay to work rather than shirk, the rents from work go mainly to the leader. That is, the leader rewards enough to compensate team members for their effort costs, but takes the lion's share of any remaining output. Given that team members do slightly better than predicted under group incentives, the leaders' rewarding strategies do not make leadership an attractive proposition.

Overall, our study shows that concentrating reward power in the hands of a leader can have beneficial effects on team production and efficiency. However, we also find that opportunism and abuse of power are tangible impediments to the success of leadership. This issue may be particularly serious in settings where leaders are endogenously appointed by team members, as the fear of exploitation may induce subordinates to resist delegation of agency to central authorities.

The remainder of the paper is organized as follows. In Section 2 we present our initial 2x2 design where we vary across treatments (i) the presence or absence of a leader, and (ii) whether team members are homogeneous or heterogeneous in their productivities. In Section 3 we present two additional treatments (with either homogeneous or heterogeneous productivities) where the leader's reward power is delegated by team members. In Section 4 we discuss how our results compare with other experimental findings, and conclude.

[4] Also see Bartling et al. (2013) on the use and abuse of power in employment contracts, and the means for agents to resist power exploitation by the principal.

2. Study 1: The effectiveness of exogenously imposed leadership

2.1 Experiment design

In our initial study we examine the effectiveness of exogenously imposed leadership using a ten-round, five-person team production game. Teams are randomly formed in the first round and remain fixed across rounds. Within each team, four subjects are randomly assigned the role of "worker" and one subject the role of "leader", and these roles are kept fixed across rounds.[5] Subjects earned points in each round and at the end of each round were informed about all decisions and earnings for all team members for the round, as well as accumulated point earnings from the current and previous rounds. At the end of the game subjects were paid based on their accumulated point earnings from all rounds.

Each round consists of two stages. In stage one each worker is endowed with 10 tokens and chooses how many to contribute to team production. Worker choices are made simultaneously. Leaders are also endowed with 10 tokens but they cannot contribute these to team production. Each token kept by a team member yields 30 points to that member, whereas each token contributed to team production by worker i generates θ_i points of "team output". We will refer to contributions of tokens in terms of the supply of "effort", though we recognize that this stylized setting where efforts correspond to a chosen number differs from natural settings in many ways. Importantly, however, our chosen effort design means that supplying "effort" in our experiment is individually costly in a tangible and transparent way and allows us to control the effort cost function. We adopted constant marginal costs and constant marginal productivity of effort for simplicity (see Falk and Fehr, 2003 for a discussion of the relative merits of chosen effort and real effort designs).

In stage two of each round the team output is redistributed among team members according to the following rules: a share γ of the team output ($0 \leq \gamma \leq 1$) is transferred to the team leader who decides how to redistribute the output among all team members (including the leader), and the remaining share $1-\gamma$ of team output is equally redistributed among the four workers. The leader is informed of the individual efforts in stage one before making her redistribution decisions by assigning reward points to workers. The point earnings of the team members are as follows:

[5] We did not use the labels "worker" and "leader" in the instructions. Instead, subjects were identified through letters A to E and group members A to D acted as workers and group member E as leader in the experiment. The instructions are reproduced in Appendix A.

$$\pi_{Leader} = 300 + \gamma \cdot \sum_{i=1}^{4} \theta_i e_i - \sum_{i=1}^{4} r_i$$

$$\pi_{Worker\ i} = 300 - 30e_i + \frac{1-\gamma}{4} \cdot \sum_{i=1}^{4} \theta_i e_i + r_i$$

where $e_i \in \{0, 1, \dots, 10\}$ denotes worker i's effort in stage one and r_i denotes the reward points assigned by the leader to worker i in stage two. The leader cannot assign negative rewards, $r_i \geq 0$, and total rewards cannot exceed the share of team output controlled by the leader, $\sum_{i=1}^{4} r_i \leq \gamma \cdot \sum_{i=1}^{4} \theta_i e_i$. At the end of each round all team members are informed of all stage one and stage two choices, and the resulting payoffs.

This initial study has four treatments in a 2x2 between-subject design. In our **NoLeader** treatments leaders have no reward power: no share of team output is transferred to the leader (i.e., $\gamma = 0$), the leader cannot assign any rewards (i.e., $r_i = 0$) and all team output is equally shared by the four workers. In our **Leader** treatments leaders have instead full reward power as they receive the whole team output to redistribute (i.e., $\gamma = 1$).[6] In our **Homogeneous** treatments productivity is homogeneous within a team, i.e. for all workers each unit of effort generates $\theta_i = 60$ points of team output. In our **Heterogeneous** treatments workers are heterogeneous in their productivity: two workers have high productivity ($\theta_i = 80$), while the other two have low productivity ($\theta_i = 40$). Table 1 summarizes the design of the experiment.

Table 1 – Study 1 experimental design

Treatment	Share of team output redistributed by leader (γ)	Workers' productivity (θ_i)	Number of subjects (teams)
NoLeader_Homogeneous	$\gamma = 0$	$\theta_i = 60$ for $i = \{1, 2, 3, 4\}$	60 (12)
Leader_Homogeneous	$\gamma = 1$	$\theta_i = 60$ for $i = \{1, 2, 3, 4\}$	60 (12)
NoLeader_Heterogeneous	$\gamma = 0$	$\theta_i = 80$ for $i = \{1,2\}$ and $\theta_i = 40$ for $i = \{3,4\}$	55 (11)
Leader_Heterogeneous	$\gamma = 1$	$\theta_i = 80$ for $i = \{1,2\}$ and $\theta_i = 40$ for $i = \{3,4\}$	55 (11)

In all treatments team output is maximized when workers supply maximum effort. However, assuming that it is common knowledge that players maximize own earnings, workers have no incentive to supply effort in any of the treatments. Consider a one-round

[6] In the NoLeader treatments leaders have no decision-making role and simply observe the decisions of the other group members. This "dummy" player is included to enhance comparability with the Leader treatment.

version of our team production game. In the NoLeader treatment, the game is a standard VCM game where a self-interested worker has a dominant strategy to provide zero effort. In our Leader treatments a self-interested leader will keep any team output produced by the workers and, anticipating this, workers do not supply any effort in stage one. These predictions for the single round game also carry over to the finitely repeated game: in the unique subgame perfect equilibrium workers provide zero effort in every round.

However, there is by now abundant evidence that not all individuals are exclusively motivated by their self-interest (see, e.g., Camerer, 2003; Fehr and Schmidt, 2006 for reviews). In a one-shot game without a leader, workers may supply effort to team production if they have other-regarding preferences (see, e.g., Fehr and Schmidt, 1999). Moreover, in a repeated version even selfishly motivated workers may supply effort in early rounds, if they believe that some workers have other–regarding preferences. Thus, the combination of repetition and incomplete information about preferences can sustain rational cooperation for several rounds (Kreps et al., 1982). Likewise, in the one-shot game with a leader, a leader with other-regarding preferences may be willing to reward workers for their effort in such a way that even a selfish worker finds it pays to supply effort (see Drouvelis et al., 2015). In the repeated game, a selfish leader may act in this way in early rounds for strategic reasons, waiting until towards the end of the game to exploit the workers (see Heijden et al., 2009 for a formal model).

2.2 Experiment procedures

The experiments were carried out at [OMITTED FOR PEER REVIEW] with 230 subjects recruited from a campus-wide distribution list.[7] Three sessions were conducted for each treatment, with either 20 or 15 subjects per session. No subject participated in more than one session. We had 60 subjects participate in each of our Homogeneous treatments, and 55 subjects participate in each of our Heterogeneous treatments.

At the beginning of a session, subjects were randomly allocated to visually-isolated computer terminals. They received written instructions that the experimenter read aloud. The instructions contained a set of control questions to test subjects' understanding of the experimental setting. Answers were checked in private by the experimenter. Once all subjects had answered all questions correctly, they were randomly allocated to teams and randomly assigned a role within the team. Subjects then played 10 rounds of the team production game described above.

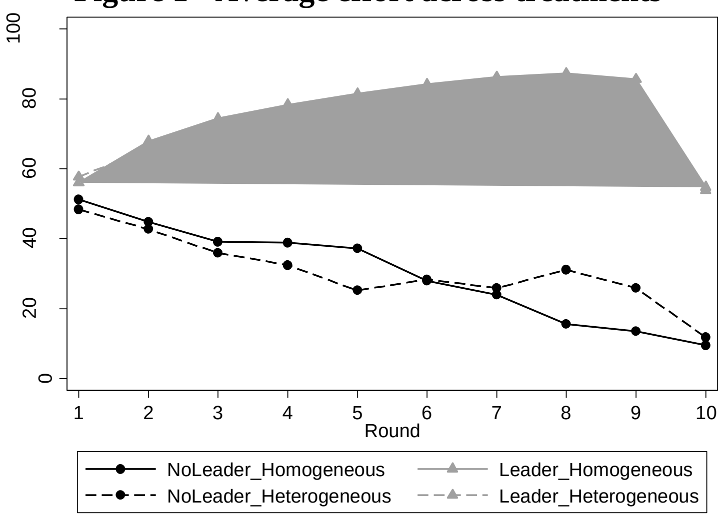
[7] Subjects were recruited through the online recruitment system ORSEE (Greiner, 2015). The experiment was computerized using the software z-Tree (Fischbacher, 2007).

At the end of round 10, subjects filled in a short questionnaire eliciting basic socio-demographic information. Subjects were then paid in cash and in private according to the sum of their earnings across the 10 rounds at a rate of £0.10 per 100 points, plus a £3.50 show-up fee. Earnings ranged between £6.18 and £14.47, averaging £7.76. On average sessions lasted about 75 minutes.

2.3 Results

We analyze our results using non-parametric tests and regressions. All reported p-values are two-sided and, for non-parametric tests, we use the team average as the independent unit of observation unless otherwise stated, so that we have 12 (Homogeneous treatments) and 11 (Heterogeneous treatments) observations per treatment. Figure 1 shows the average effort supplied by workers across our four treatments. In round 1 efforts range from 48% to 58% across treatments, but these differences are not statistically significant (Kruskal-Wallis test: p = 0.588).[8] Marked differences between treatments then emerge from round 2 onwards: efforts gradually decrease over rounds in our NoLeader treatments, whereas they increase in our Leader treatments. This pattern is observed both in the Homogeneous and Heterogeneous treatments. In round 10 of all treatments we observe an end-game drop in effort, which is particularly sharp in our Leader treatments.

Figure 1 - Average effort across treatments



Averaging across all rounds, subjects supply about 30% and 31% of maximum possible effort in NoLeader_Homogeneous and NoLeader_Heterogeneous, respectively. In contrast,

[8] This test is based on 48 individual observations in each of our NoLeader_Homogeneous and Leader_Homogeneous treatments, and 44 individual observations in each of our NoLeader_Heterogeneous and Leader_Heterogeneous treatments.

average effort is substantially higher in the treatments with leader: 76% of maximum possible effort in Leader_Homogeneous and 70% in Leader_Heterogeneous. These differences between treatments with and without leaders are statistically significant (Mann-Whitney tests: NoLeader_Homogeneous vs. Leader_Homogeneous p < 0.001; NoLeader_Heterogeneous vs. Leader_Heterogeneous p = 0.001). Thus, replacing the group incentive scheme of our No_Leader treatment with the discretionary incentives of our Leader treatment increases effort by 46 percentage points in homogeneous teams and by 39 percentage points in heterogeneous teams.

Note that the Leader and NoLeader treatments differ in two dimensions: 1) whether leaders receive a share of the total output; and 2) whether they have reward power.[9] To estimate the impact of each dimension on our treatment effect, we conducted additional control treatments identical to the NoLeader condition except that the leader also receives a share of team output (team output is shared equally among the five group members). We conducted two sessions with 30 subjects in Control_Homogeneous (6 independent observations) and two sessions with 35 subjects in Control_Heterogeneous (7 independent observations). We find that team output in these control treatments is significantly lower than in our NoLeader treatments (12% in both Homogeneous and Heterogeneous, averaging across all rounds; Mann-Whitney tests: p = 0.007 NoLeader_Homogeneous vs. Control_Homogeneous; p = 0.001 NoLeader_Heterogeneous vs. Control_Heterogeneous).[10] Thus, relative to our No_Leader treatment, giving the leader an equal share of output, but no reward power, reduces effort by 18 percentage points in Homogeneous and by 19 percentage points in Heterogeneous teams. Relative to the control treatments, where all team members including the leader get an equal share of output, we find that replacing group incentives with discretionary incentives increases effort by 64 percentage points in homogeneous teams and by 58 percentage points in heterogeneous teams.

It is interesting to observe that the effectiveness of leadership does not seem to be diminished in the treatments with heterogeneous worker productivities. In fact, we observe very similar efforts in homogeneous and heterogeneous teams (Mann-Whitney tests: NoLeader_Homogeneous vs. NoLeader_Heterogeneous p = 0.580; Leader_Homogeneous vs. Leader_Heterogeneous p = 0.853). Within the Heterogeneous treatments, we observe only small differences in effort between workers with high and low productivity. Averaging

[9] We thank an anonymous referee for pointing this out.

[10] These tests are based on comparisons between 12 and 6 observations in the Homogeneous treatments and 11 and 7 observations in the Heterogeneous treatments. See Appendix B for additional analysis of these treatments.

across rounds of our NoLeader treatment, high productivity workers supply 34% of maximum possible effort, while low productivity workers supply 28%. High productivity workers supply more effort in our Leader treatment as well, 73% compared with 66% by low productivity workers. However, none of these differences are statistically significant (Wilcoxon signed-rank tests: high vs. low productivity workers p = 0.477 in NoLeader_Heterogeneous and p = 0.286 in Leader_Heterogeneous).[11]

Overall, these findings suggest that in our Leader treatments leaders use their reward power effectively and adopt redistribution strategies that induce workers to contribute to team output. But how do leaders distribute rewards across team members? In Tables 2 and 3 we start addressing this question by plotting the average share of team output received by workers disaggregated by level of effort. Table 2 presents data from the Homogeneous treatment, while Table 3 presents data from the Heterogeneous treatment, distinguishing between high and low productivity workers. The tables distinguish between cases where a worker's own effort is above or below the average effort of the rest of their team in a given round of the experiment.[12]

The tables show that leaders reward relatively high effort with higher shares of team output. In Homogeneous workers who supply low effort (0 to 4 units) receive on average between 4% and 17% of team output. In contrast, workers who supply high effort (5 units or more) receive between 17% and 22% of team output. Similarly, in Heterogeneous high productivity workers who supply low effort receive between 5% and 15% of team output, whereas high productivity workers supplying high effort receive between 18% and 21% of team output. Similar patterns emerge for low productivity workers: low effort receives between 2% and 17% of team output, while high effort receives between 16% and 21%. In both treatments, the differences between shares of team output assigned to workers who supply low and high effort are highly significant (Wilcoxon signed-rank tests, p = 0.002 for Homogeneous; p = 0.004 for Heterogeneous high productivity workers; p = 0.007 for Heterogeneous low productivity workers). Notably, leaders assign similar rewards to low and high productivity workers. On average, low and high productivity workers receive 17% and

[11] These tests use average effort across the 10 rounds computed separately for high and low productivity workers within a team. This generates a set of 11 paired observations that we use as the independent unit of observation. The Wilcoxon signed-rank tests reported elsewhere in the paper are performed in a similar fashion.
[12] In the tables we focus on shares of team output instead of the number of reward points received by workers to account for the fact that the amount of reward points that the leader can redistribute to workers is endogenous and depends on the amount of team output produced.

18% of team output, respectively. The difference is not statistically significant (Wilcoxon signed-rank test, p = 0.424).

Table 2 – Share of team output received by effort level - Homogeneous

Worker's effort	At or above average team effort	Below average team effort	Overall
0	- *(0)*	0.04 *(39)*	0.04 *(39)*
1 or 2	0.28 *(3)*	0.11 *(18)*	0.13 *(21)*
3 or 4	0.25 *(10)*	0.14 *(24)*	0.17 *(34)*
5 or 6	0.25 *(17)*	0.13 *(28)*	0.17 *(45)*
7 or 8	0.28 *(35)*	0.17 *(26)*	0.23 *(61)*
9 or 10	0.22 *(264)*	0.17 *(16)*	0.22 *(280)*

Notes: The number of underlying observations is reported in parentheses.

Table 3 – Share of team output received by effort level - Heterogeneous

Worker's effort	At or above average team effort	Below average team effort	Overall
0	- / - *(2) / (2)*	0.05 / 0.02 *(28) / (14)*	0.05 / 0.02 *(30) / (16)*
1 or 2	0.05 / 0.16 *(2) / (3)*	0.10 / 0.12 *(11) / (13)*	0.09 / 0.13 *(13) / (16)*
3 or 4	0.14 / 0.20 *(5) / (7)*	0.15 / 0.14 *(25) / (10)*	0.15 / 0.17 *(30) / (17)*
5 or 6	0.19 / 0.17 *(12) / (9)*	0.17 / 0.16 *(9) / (21)*	0.18 / 0.16 *(21) / (30)*
7 or 8	0.21 / 0.24 *(7) / (14)*	0.12 / 0.16 *(5) / (5)*	0.17 / 0.22 *(12) / (19)*
9 or 10	0.21 / 0.22 *(111) / (120)*	0.15 / 0.03 *(3) / (2)*	0.21 / 0.21 *(114) / (122)*

Notes: The first numbers of each cell refer to high productivity workers and the second numbers refer to low productivity workers. The number of underlying observations is reported in parentheses. Note that we have two groups where all workers exerted effort equal to 0 (column 2) and so the leader had no team output to allocate.

Tables 2 and 3 also show that the share of team output assigned to a worker depends not just on that worker's effort, but also on how that compares to the effort of other team members. For instance, in Homogeneous, workers who exert low levels of effort receive on average between 4% and 14% of team output when their effort is below the average effort in their team. However, when the same levels of effort are at or above the team average, they are rewarded with 25% to 28% of team output. In fact, for any category of effort in Table 2, workers always receive a larger reward when their effort is at or above, rather than below, the team average. Analogous patterns emerge in Table 3 for the Heterogeneous treatment, for both low and high productivity workers.

The relation between a worker's effort and average effort in their team is illustrated in Figure 2, which shows the average share of team output received by workers in each treatment, disaggregated based on how their efforts relate to the average effort in their team. The figure also shows the average share of team output received by leaders in each treatment.

Figure 2 – Share of team output received

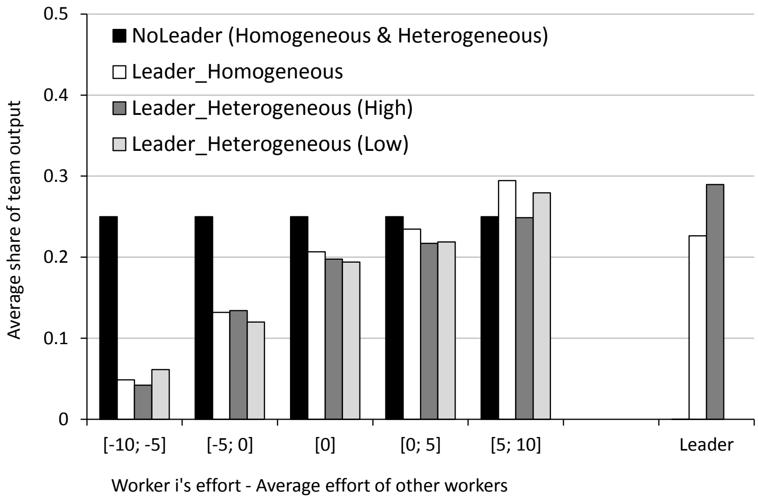


Figure 2 clearly shows that leaders follow a strategy that rewards contributions and punishes free-riding in both Leader treatments. In Homogeneous, workers who supply effort at or above the average of their team in a round receive on average 23% of team production, while those supplying effort below the average receive 12% of team production. These differences are highly significant (Wilcoxon signed-rank test, p = 0.002). The same pattern is observed in the Heterogeneous treatment, for both high and low productivity workers. Here, high productivity workers who supply effort at or above the team average receive on average 21% of team production, while those supplying effort below the team average receive 11% of team production (Wilcoxon signed-rank test, p = 0.004). Similarly, low productivity workers receive on average 20% of team production when their effort is at or above the team average, and 10% when it is below the average (Wilcoxon signed-rank test, p = 0.007).

Overall, these results suggest that leaders provide strong incentives for effort provision, both in our Homogeneous and Heterogeneous treatments. We conclude our analysis by studying the impact of leadership on earnings and efficiency. For a measure of efficiency we use attained team earnings in excess of Nash equilibrium earnings as a percentage of maximum possible excess earnings. Table 4 reports individual earnings, combined earnings, and efficiencies per round across treatments.

Table 4 – Individual earnings and efficiency

	NoLeader Homogeneous	Leader Homogeneous	NoLeader Heterogeneous	Leader Heterogeneous
Leader's Earnings	300 *(0.0)*	686.8 *(203.4)*	300 *(0.0)*	746.0 *(190.8)*
Workers' Earnings	390.6 *(38.9)*	430.2 *(57.1)*	-	-
Low Productivity			407.1 *(62.8)*	406.4 *(68.4)*
High Productivity			389.5 *(38.6)*	401.2 *(60.2)*
Combined Earnings	1862.5 *(155.6)*	2407.5 *(193.7)*	1893 *(174.1)*	2361.3 *(220.5)*
Efficiency	30%	76%	33%	72%

Notes: "Combined Earnings" are the sum of leader's and workers' earnings. "Efficiency" is computed as (combined earnings – 1500) / (2700 – 1500), where 1500 are the theoretical earnings under zero contributions and 2700 are maximum possible combined earnings. Standard deviations based on team averages in parentheses.

The earnings analysis confirms the effectiveness of leadership in promoting this measure of efficiency in our team production setting. In the Homogeneous treatment combined earnings increase from 1862.5 when leaders have no reward power (an efficiency of 30%) to 2407.5 when leaders have reward power (an efficiency of 76%). This difference is statistically significant according to a Mann-Whitney test (NoLeader_Homogeneous vs. Leader_Homogeneous $p < 0.001$). Similarly, in the Heterogeneous treatment combined earnings increase from 1893 when leaders have no reward power (an efficiency of 33%) to 2361.3 when leaders have reward power (an efficiency of 72%). This difference is statistically significant (Mann-Whitney test: NoLeader_Heterogeneous vs. Leader_Heterogeneous $p = 0.001$).[13]

Table 4, however, also shows that the efficiency gains of leadership are redistributed very unequally between leaders and workers: most of the efficiency gains accrue to leaders, whereas workers' earnings are not very different between NoLeader and Leader treatments. Indeed, for leaders there is a statistically significant difference between earnings in NoLeader and Leader treatments (Mann-Whitney tests: NoLeader_Homogeneous vs.

[13] We do not detect statistically significant differences in efficiency between our Homogeneous and Heterogeneous treatments (Mann-Whitney tests: NoLeader_Homogeneous vs. NoLeader_Heterogeneous $p = 0.975$; Leader_Homogeneousvs. Leader_Heterogeneous $p = 0.460$).

Leader_Homogeneous $p < 0.001$; NoLeader_Heterogeneous vs. Leader_Heterogeneous $p < 0.001$), while for workers earnings in the Leader treatment are not significantly different from earnings in the NoLeader treatment (Mann-Whitney tests: NoLeader_Homogeneous vs. Leader_Homogeneous $p = 0.106$; NoLeader_Heterogeneous vs. Leader_Heterogeneous low productivity $p = 0.974$; high productivity $p = 0.577$).

3. Study 2: The emergence of endogenous leadership

Our initial study raises a natural question about whether leadership will emerge in environments where leaders are not exogenously imposed on a team, but are endogenously appointed by team members. In such environments, the success of leadership may be hindered: team members may refuse to support leadership if leaders behave too opportunistically and do not share enough of the proceeds of team production. On the other hand, the endogenous appointment of leaders may constrain their opportunism and thus facilitate the emergence of successful leadership. Our follow-up study addresses these questions by studying a setting where leadership is not imposed on teams, but instead may emerge endogenously through the support of team members.

3.1 Experiment design and procedures

The follow-up study is based on the same ten-round, five-person team production game that we used in our initial study and that we described above. However, differently from the initial study, at the beginning of the game the four workers simultaneously express a preference for the leader's reward power, i.e. the share γ of team output that will be redistributed by the leader. Workers can choose between one of six possible levels of γ: 0, 0.2, 0.4, 0.6, 0.8, or 1, one of which will be implemented across all ten rounds. For example, with $\gamma = 0.4$, in each round 40% of team output is transferred to the leader, whereas 60% of team output is evenly distributed among workers. Note that when leaders are granted no reward power ($\gamma = 0$) or full reward power ($\gamma = 1$) we have cases that correspond to the NoLeader and Leader treatments of the initial study. The level of γ implemented in a team is decided using a "random dictator" rule: after each worker has expressed his/her preference, one worker from each team is selected at random and his/her choice of γ is implemented.[14]

[14] We implemented a random dictator rule because we are interested in team members' preferences for leadership. The random dictator rule is the simplest mechanism that gives workers an incentive to reveal their most preferred level of reward power. While the use of alternative voting rules, such as a majority rule, may increase the legitimacy of the voting outcome, this would also introduce strategic considerations in voting decisions that would confound the identification of workers' preferences for leadership.

We study a repeated version of this game across three blocks of 10 rounds each. Subjects were informed of this structure at the beginning of the experiment. Group composition and roles were kept fixed across blocks and rounds. At the beginning of each block, all team members were notified of the randomly selected level of γ and knew that it would apply to the subsequent 10 rounds of the block. Note that this introduces an incentive for leaders not to abuse their power in the early blocks in order to increase the probability of being granted power in later bocks, and this may represent an effective mechanism for constraining the opportunism of leaders. As in the initial study, we conduct a **Homogeneous** treatment where all workers have the same productivity ($\theta_i = 60$) and a **Heterogeneous** treatment where two workers have high productivity ($\theta_i = 80$) and two have low productivity ($\theta_i = 40$).

The experiments were carried out at [OMITTED FOR PEER REVIEW] using the same procedures as in the initial study. We had 195 subjects in total, 100 in the Homogeneous treatment (20 teams) and 95 in the Heterogeneous treatment (19 teams). Subjects were paid according to their accumulated earnings across the 30 rounds of the experiment at a rate of £0.10 per 100 points. Earnings ranged between £12.23 and £24.89, averaging £16.04, including a £3.50 show-up free. On average session lasted about 150 minutes.

3.2 Results

Figure 3 shows the reward power that workers are willing to delegate to leaders in each of the three blocks of the experiment in the Homogeneous and Heterogeneous treatments. The figure is based on the levels of γ submitted by all workers at the beginning of each block, and not only on those that were actually implemented in the experiment.

Initially, delegation of reward power is similar across our two treatments: in block 1 of both treatments workers are willing to delegate about one-third of team output to leaders, 29% in Homogeneous and 33% in Heterogeneous. In the Homogeneous treatment the share of team output delegated to leaders increases to 42% in block 2 and 45% in block 3. In contrast, in Heterogeneous workers delegate less in blocks 2 and 3 than in block 1: 26% in block 2 and 29% in block 3. The difference between treatments in block 1 is not significant (Mann-Whitney test: p = 0.422), whereas the differences in block 2 and 3 are (Mann-Whitney tests: block 2 p = 0.034; block 3 p = 0.026).[15]

[15] The test for block 1 is based on 80 individual observations in Homogeneous and 76 individual observations in Heterogeneous. The other tests use team averages as the independent unit of observation and are therefore based on 20 (Homogeneous) and 19 (Heterogeneous) observations per treatment.

Figure 3 – Average reward power delegated by workers

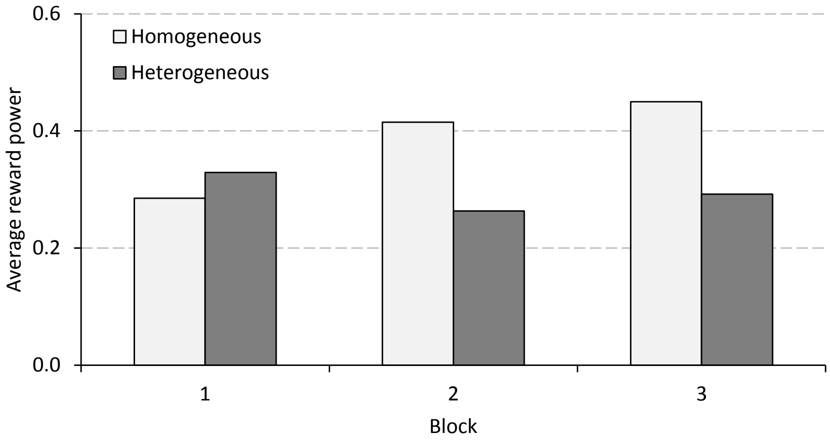
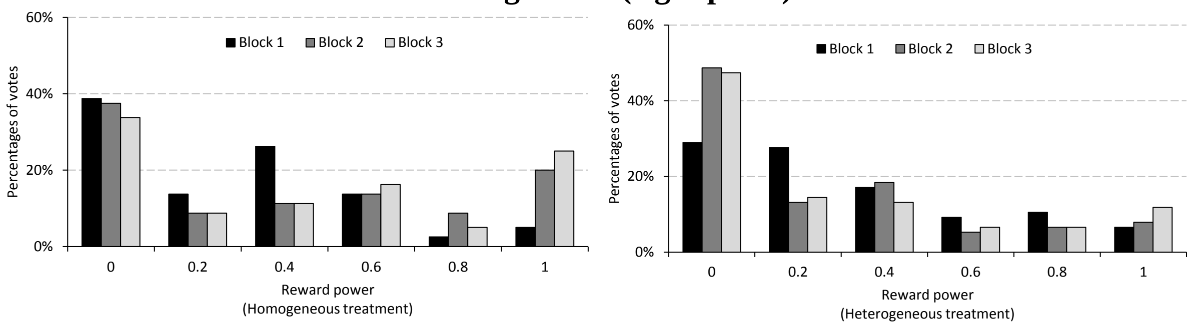


Figure 4 shows the distribution of delegation choices across the three blocks of the experiment in the Homogeneous (left panel) and Heterogeneous (right panel) treatments.

Figure 4 – Distribution of delegation choices in Homogeneous (left panel) and Heterogeneous (right panel)



In both treatments, and in all blocks, the modal worker delegates no power to the leader. In the Homogeneous treatment the proportion of workers who are opposed to delegation is stable across blocks and varies between 34% and 39%. In Heterogeneous this proportion is more volatile across blocks: in block 1 29% of workers are opposed to delegation, and this proportion increases to 49% in block 2 and 47% in block 3. The fraction of workers in favor of delegation also varies across blocks. In both treatments, workers are initially cautious in delegating reward power: the most popular choices in favor of delegation in block 1 are for a reward power of 0.2 or 0.4. In later blocks workers tend to delegate more power to the leader, especially in the Homogeneous treatment where the fraction of workers in favor of full delegation ($\gamma = 1$) increases from 5% in block 1 to 25% in block 3.

Overall, these results show that workers delegate some power to leaders, but they do not fully exploit the potential for delegation. In both treatments, and in all blocks, workers transfer less than half of team output to the leader. Moreover, a substantial fraction of team members (between one-third and one-half) are opposed to any transfer of power to the leader.

This reluctance to delegate power to the leader is not due to the fact that endogenous leaders are unable to provide appropriate incentives for team members to supply effort. In fact, in Appendix C we show that effort and efficiency are considerably higher in teams with than without leaders. Moreover, as in our initial study, the positive effect of leadership on effort is driven by the redistribution strategies adopted by leaders, which reward workers who supply effort and withhold rewards from those who shirk (see Appendix C). However, also as in our initial study, leaders tend to appropriate large fractions of the proceeds from team production (about 20% of the team output in Homogeneous, and about 30% in Heterogeneous, see Appendix C, Figure C.3). As in Study 1, this leaves workers no better off with than without a leader and generates strong asymmetries between leaders' and workers' earnings (see Appendix C, Tables C.2 and C.3).

4. Discussion and conclusions

Our studies investigate a team production game that exhibits the fundamental incentive problems facing teams when agents are purely self-interested. When the benefits of team production are distributed using group incentive schemes, such as the simple revenue-sharing scheme we employ, there are well-known free-riding incentives. We observe substantial levels of shirking under simple revenue-sharing. When a leader can allocate the benefits, the leader can incentivize workers by allocating benefits to those who contribute the most and withholding them from those who free-ride. However, in theory this hierarchical solution to the free-rider problem introduces further incentive problems. In particular, a self-interested leader has an incentive to keep all the benefits for herself. Indeed, this is identified by Miller (1992, p. 154) as the central dilemma in a hierarchy: "*how to constrain the self-interest of those with a stake in the inevitable residual generated by an efficient incentive system*".

Our research suggests that self-restraint by leaders can resolve this dilemma. We find that leaders do reward those who contribute to team output, and do so in a way that incentivizes efficient effort provision. To this extent, our research suggests that hierarchical institutional structures can resolve free-rider problems in teams. However, our research also identifies another problem. Leaders who calibrate rewards so as to maximize their residual claim while giving away just enough to incentivize work will tend to distribute the benefits of leadership unevenly. Workers are (just) compensated for their efforts, and the majority of the rents accrue to the leader. Given this, workers are barely better off in a well-functioning, hard-working team than in a dysfunctional team of shirkers. This asymmetric distribution of the leadership rents limits the benefits that workers get from a hierarchical structure. Thus we find that when the leader's reward power requires the acquiescence of workers, workers are less willing to empower the leader. Thus, endogenously arriving at an efficient incentive scheme is a more difficult challenge.

It is interesting to relate our findings to the existing experimental literatures on the effectiveness of leadership in solving the free-rider problem. Our results on the effectiveness of leadership are in line with the findings of the few experimental studies in economics that have focused on central leaders who have the power to allocate group resources. Heijden et al. (2009) and Stoddard et al. (2014), for example, compare settings where group output is shared equally among group members to settings where group output is distributed by a group leader who has the power to monitor other group members' efforts. Both studies differ in numerous ways from ours: for example, in Heijden et al. (2009) leaders have a productive as well as allocative role, effort decisions are binary, and there are complementarities in team production; in Stoddard et al. (2014) subjects play a common-property resource game, leaders can only choose between a limited and exogenously determined set of allocation rules, and they cannot misappropriate team output. Moreover, neither of these studies considers asymmetries in productivity or endogenous incentive mechanisms. Despite the several differences between designs, these studies, like us, find that the introduction of a leader strongly and significantly increases cooperation within the group and improves efficiency.[16] Also related is the paper by Abeler et al. (2010), who study a principal-agent setting where the principal can distribute wages to the agents after having observed their efforts. They find that allowing principals to pay different wages to the agents has a beneficial effect on effort relative to a setting where principals are forced to pay equal wages, suggesting that the power to discretionally distribute resources is an important feature of successful leadership.[17]

We show that these positive effects of leadership on cooperation are present in settings where leaders manage heterogeneous as well as homogeneous teams. In fact, we do not observe any significant differences in team production between homogeneous and

[16] Also related is Grosse et al. (2011), who study the role of central agents in investing in the monitoring of workers' effort in a team production experiment. They compare teams with central monitors with teams with peer monitoring as well as teams without monitoring possibilities. They then let workers choose their preferred monitoring system (if any). They find that teams are very successful at self-monitoring and thus often opt for peer monitoring systems. However, when peer monitoring is comparatively more costly than central monitoring, peer monitoring become ineffective, and workers tend to choose central monitoring.

[17] See also Bartling and von Siemens (2011), who find that wage inequality has no detrimental effect on effort provision in a team production setting.

heterogeneous teams in our experiment, nor do we observe differences in the efforts of high and low productivity workers in heterogeneous teams.

The fact that we do not find differences between homogeneous and heterogeneous teams when there is no leader is not necessarily surprising given that there is mixed evidence about the impact of heterogeneity in VCM settings (see the summary of the literature in Reuben and Riedl, 2013). The studies most related to ours are Tan (2008) and Kölle (2015), who also study heterogeneity in productivities. Tan (2008) finds homogeneous groups contribute more than heterogeneous groups, whereas Kölle (2015) finds the opposite.

Previous experiments have not studied the effect of heterogeneity on team performance when leaders have reward power. However, heterogeneity has been found to undermine the effectiveness of "leading-by-example" in public good games.[18] Levati et al. (2007) show that when group members have different endowments and are aware of the distribution of endowments, leadership is still beneficial relative to a setting without a leader, but not as effective as in the case of equal endowments. When group members are not aware of the distribution of endowments in the group, this small positive effect vanishes and contributions are not different than in the absence of a leader. Levati et al. (2007) argue that the ineffectiveness of leadership in heterogeneous groups stems from the inability of leaders to send a signal about the appropriate contribution levels when there is incomplete information about the distribution of endowments. This could explain the differences between their findings and ours: in our setting, the nature and extent of heterogeneity is common knowledge among team members and thus the leaders can use rewarding strategies to send clear messages about appropriate contributions.

While our results suggest that central leaders can successfully promote pro-social behavior and cooperation, Nosenzo and Sefton (2014) show that centralized leadership may not always be successful. They study voluntary contribution games where leaders can assign rewards or punishments to the other team members after having observed their contributions. They find that reward/punishment power does not increase contributions relative to a setting where no rewards or punishments are possible.[19] Nosenzo and Sefton note that one of the

[18] In these games contributions to a public good are made sequentially so that the contributions made by leaders may, in principle, influence those made by followers (see, e.g., Gächter and Renner, 2003; Güth et al., 2007; Potters et al., 2007; Drouvelis and Nosenzo, 2013; Pogrebna et al., 2011; Haigner and Wakolbinger, 2010; Rivas and Sutter, 2011). See also van Vugt and De Cremer (1999), van Vugt and De Cremer (2003), and van Vugt et al. (2004) for a social psychology perspective on leadership and leading-by-example in social dilemma games.
[19] Gürerk et al. (2015) report similar findings in a 20-round 4-person public goods game. They find that contributions are somewhat higher in treatments with leaders with reward/punishment power than in a treatment without reward/punishment possibilities. However, they also show that this effect is driven by the

reasons that could explain the failure of leadership in their study is the existence of heterogeneity in the quality of leaders in their experiment. While some leaders used effective reward or punishment strategies to encourage cooperation, other leaders used punishments or rewards ineffectively, or did not try to use them at all and kept for themselves the resources that were earmarked for punishing or rewarding team members.[20] Thus, opportunism and, more generally, differences in the quality of leaders can hinder the potential for leadership to overcome the free-rider problem.

These issues are also present in our study, as leaders in our experiments also face opportunistic incentives to appropriate all the team output for themselves rather than to redistribute it and encourage team production. While our leaders seem to be able to eschew the temptation to seize all team resources for themselves and dispense enough rewards to repay workers for their efforts, they still take the lion's share of team output. As a consequence, and despite the large differences in team production, workers are not better off with than without a leader. This constitutes a significant impediment to the success of leadership in our setting since, as our second study shows, workers are reluctant to delegate power to the leader and thus the benefits of leadership are only partly realized.

These findings are in line with previous studies of leadership and delegation of agency. For example, Hamman et al. (2011) study a linear public goods production setting where contributions are either selected individually by group members, or by a designed "allocator" who is granted complete authority over the choice of contributions of all group members. They find that delegating authority to a central allocator has a strong, positive effect on public good provision. However, in a second study Hamman et al. (2011) allow each group member to choose whether or not to cede agency to the central allocator. Similar to our study, they find that groups fail to seize the benefits of leadership as not enough group members are willing to transfer their agency to the central allocator.

Overall, these findings point to a potential problem with hierarchical solutions to the free-rider problem. Leaders face strong incentives to behave opportunistically and abuse their power. Even when leaders manage to eschew the temptation of opportunism, the fear of exploitation may induce followers to resist delegation of agency to central authorities.

---

reward/punishment incentives rather than by leadership per se: contributions in the leader treatments are in fact not different from those in two additional treatments with peer-to-peer rewards and punishment.

[20] This finding is also emphasized by Heijden et al. (2009), who also observe a mixture of "good" and "bad" leaders in their setting. Similarly, Brandts and Cooper (2007) examine the role of leadership in a turnaround game and show that while leaders are on average effective in increasing coordination in groups, not all leaders are equally effective and a third of groups experience complete coordination failure in the final rounds of the experiment.

References

Abeler, J., S. Altmann, S. Kube, and M. Wibral. 2010. Gift Exchange and Workers' Fairness Concerns - When Equality is Unfair. *Journal of the European Economic Association* 8(5), 1299–1324.

Bartling, B., E. Fehr, and H. Herz. 2014. The Intrinsic Value of Decision Rights. *Econometrica* 82(6), 2005–2039.

Bartling, B., E. Fehr, and K.M. Schmidt. 2013. Use and Abuse of Authority: A Behavioral Foundation of the Employment Relation. *Journal of the European Economic Association* 11(4), 711–742.

Bartling, B., and U. Fischbacher. 2012. Shifting the Blame: On Delegation and Responsibility. *The Review of Economic Studies* 79(1), 67–87.

Bartling, B., and F.A. von Siemens. 2011. Wage Inequality and Team Production: An Experimental Analysis. *Journal of Economic Psychology* 32, 1–16.

Bartling, B., F.A. von Siemens, and F.A. von Siemens. 2010. Equal Sharing Rules in Partnerships. *Journal of Institutional and Theoretical Economics (JITE) / Zeitschrift für die gesamte Staatswissenschaft* 166(2), 299–320.

Bass, B., and R. Bass. 2008. *The Bass Handbook of Leadership*. Fourth Edition. New York: Free Press.

Brandts, J., and D.J. Cooper. 2007. It's What You Say, Not What You Pay: An Experimental Study of Manager Employee Relationships in Overcoming Coordination Failure. *Journal of the European Economic Association* 5(6), 1223–1268.

Camerer, C., and R. Weber. 2012. Experimental Organizational Economics. In *The Handbook of Organizational Economics*, 213–262. Princeton: Princeton University Press.

Camerer, C.F. 2003. *Behavioral game theory*. Princeton: Princeton University Press.

Charness, G., and P. Kuhn. 2011. Lab Labor: What Can Labor Economists Learn from the Lab? In *Handbook of Labor Economics*, ed by. O. Ashenfelter and D. Card, 4A:229–330. Elsevier B.V.

Dal Bó, P., A. Foster, and L. Putterman. 2010. Institutions and Behavior: Experimental Evidence on the Effects of Democracy. *American Economic Review* 100(5), 2205–2229.

Drouvelis, M., and D. Nosenzo. 2013. Group identity and leading-by-example. *Journal of Economic Psychology* 39, 414–425.

Drouvelis, M., D. Nosenzo, and M. Sefton. 2015. Team Incentives and Leadership. CeDEx Discussion Paper 2015-05.

Falk, A., and E. Fehr. 2003. Why labour market experiments? *Labour Economics* 10(4), 399–406.

Fehr, E., H. Herz, and T. Wilkening. 2013. The Lure of Authority: Motivation and Incentive Effects of Power. *American Economic Review* 103(4), 1325–59.

Fehr, E., and K.M. Schmidt. 1999. A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics* 114(3), 817–68.

Fehr, E., and K.M. Schmidt. 2006. The Economics of Fairness, Reciprocity and Altruism - Experimental Evidence and New Theories. In *Handbook of the Economics of Giving, Reciprocity and Altruism*, ed by. Serge Kolm and Jean Mercier Ythier, Volume 1:615–691. Elsevier.

Fiedler, F.E. 1967. *A theory of leadership effectiveness*. New York: McGraw Hill.

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

French, J.R.P., and B. Raven. 1959. The bases of social power. In *Group dynamics*, ed by. D. Cartwright and A. Zander. New York: Harper & Row.

Gächter, S., and E. Renner. 2003. Leading by example in the presence of free rider incentives. Paper presented at the Workshop on 'Leadership and Social Interactions', March 2003, Lyon.

Greiner, B. 2015. Subject pool recruitment procedures: organizing experiments with ORSEE. *Journal of the Economic Science Association* 1(1), 114–125.

Grosse, S., L. Putterman, and B. Rockenbach. 2011. Monitoring in Teams: Using Laboratory Experiments to Study a Theory of the Firm. *Journal of the European Economic Association* 9(4), 785–816.

Guillen, P., D. Merrett, and R. Slonim. 2014. A New Solution for the Moral Hazard Problem in Team Production. *Management Science* 61(7), 1514–1530.

Gürerk, Ö., T. Lauer, and M. Scheuermann. 2015. Leadership with Individual Rewards and Punishments. MPRA Paper No. 65691.

Güth, W., M.V. Levati, M. Sutter, and E. van der Heijden. 2007. Leading by example with and without exclusion power in voluntary contribution experiments. *Journal of Public Economics* 91(5–6), 1023–1042.

Haigner, S.D., and F. Wakolbinger. 2010. To lead or not to lead: Endogenous sequencing in public goods games. *Economics Letters* 108(1), 93–95.

Hamman, J.R., R.A. Weber, and J. Woon. 2011. An Experimental Investigation of Electoral Delegation and the Provision of Public Goods. *American Journal of Political Science* 55(4), 738–752.

Heijden, E. van der, J. Potters, and M. Sefton. 2009. Hierarchy and opportunism in teams. *Journal of Economic Behavior & Organization* 69(1), 39–50.

Hermalin, B.E. 2013. Leadership and Corporate Culture. In *The Handbook of Organizational Economics*, ed by. Robert Gibbons and John Toberts, 432–478. Princeton: Princeton University Press.

Jago, A.G. 1982. Leadership: Perspectives in Theory and Research. *Management Science* 28(3), 315–336.

Kölle, F. 2015. Heterogeneity and cooperation: The role of capability and valuation on public goods provision. *Journal of Economic Behavior & Organization* 109, 120–134.

Kreps, D.M., P. Milgrom, J. Roberts, and R. Wilson. 1982. Rational Cooperation in the Finitely Repeated Prisoners-Dilemma. *Journal of Economic Theory* 27(2), 245–252.

Levati, M.V., M. Sutter, and E. van der Heijden. 2007. Leading by example in a public goods experiment with heterogeneity and incomplete information. *Journal of Conflict Resolution* 51(5), 793–818.

Messick, D.M., and M.B. Brewer. 1983. Solving Social Dilemmas; A Review. In *Review of Personality and Social Psychology*, ed by. L. Wheeler and P. Shaver. Beverly Hills, CA: Sage.

Miller, G. 1992. *Managerial Dilemmas: The Political Economy of Hierarchy*. New York: Cambridge University Press.

Nosenzo, D., and M. Sefton. 2014. Promoting Cooperation: The distribution of reward and punishment power. In *Reward and Punishment in Social Dilemmas*, ed by. Paul A. M. Van Lange, Bettina Rockenbach, and Toshio Yamagishi. Oxford University Press.

Podsakoff, P.M., W.H. Bommer, N.P. Podsakoff, and S.B. MacKenzie. 2006. Relationships between leader reward and punishment behavior and subordinate attitudes, perceptions, and behaviors: A meta-analytic review of existing and new research. *Organizational Behavior and Human Decision Processes* 99(2), 113–142.

Podsakoff, P.M., and C.A. Schriesheim. 1985. Leader reward and punishment behavior: A methodological and substantive review. In *Research in organizational behavior*, ed by. B. Staw and L.L. Cummings. San Francisco: Jossey-Bass.

Pogrebna, G., D.H. Krantz, C. Schade, and C. Keser. 2011. Words versus actions as a means to influence cooperation in social dilemma situations. *Theory and Decision* 71(4), 473–502.

Potters, J., M. Sefton, and L. Vesterlund. 2007. Leading-by-example and signaling in voluntary contribution games: an experimental study. *Economic Theory* 33, 169–182.

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.

Rivas, M.F., and M. Sutter. 2011. The benefits of voluntary leadership in experimental public goods games. *Economics Letters* 112(2), 176–178.

Samuelson, C.D. 1991. Perceived Task Difficulty, Causal Attributions, and Preferences for Structural Change in Resource Dilemmas. *Personality and Social Psychology Bulletin* 17(2), 181–187.

Stoddard, B., J.M. Walker, and A. Williams. 2014. Allocating a voluntarily provided common-property resource: An experimental examination. *Journal of Economic Behavior & Organization* 101, 141–155.

Sutter, M., S. Haigner, and M.G. Kocher. 2010. Choosing the Carrot or the Stick? Endogenous Institutional Choice in Social Dilemma Situations. *The Review of Economic Studies* 77(4), 1540–1566.

Tan, F. 2008. Punishment in a Linear Public Good Game with Productivity Heterogeneity. *De Economist* 156(3), 269–293.

van Vugt, M., and D. De Cremer. 1999. Leadership in social dilemmas: The effects of group identification on collective actions to provide public goods. *Journal of Personality and Social Psychology* 76(4), 587–599.

van Vugt, M., and D. De Cremer. 2003. Leader endorsement in social dilemmas: Comparing the instrumental and relational perspectives. *European Review of Social Psychology* 13(1), 155–184.

van Vugt, M., S.F. Jepson, C.M. Hart, and D. De Cremer. 2004. Autocratic leadership in social dilemmas: A threat to group stability. *Journal of Experimental Social Psychology* 40(1), 1–13.

Yukl, G.A. 1989. *Leadership in Organizations*. 4th ed. Prentice Hall.

Online Supplementary Materials

Online Appendix A: Instructions

[All treatments: Preliminary Instructions

Welcome! You are about to take part in a decision-making experiment. This experiment is run by the [OMITTED FOR PEER REVIEW] and has been financed by various research foundations. Just for showing up you have already earned £3.50. You can earn additional money depending on the decisions made by you and other participants. It is therefore very important that you read these instructions with care.

It is important that you remain silent and do not look at other people's work. If you have any questions, or need assistance of any kind, please raise your hand and an experimenter will come to you. If you talk, laugh, exclaim out loud, etc., you will be asked to leave and you will not be paid. We expect and appreciate your following of these rules.

We will first jointly go over the instructions. After we have read the instructions, you will have time to ask clarifying questions. We would like to stress that any choices you make in this experiment are entirely anonymous. Please do not touch the computer or its mouse until you are instructed to do so. Thank you.

]

[Exogenous treatments: Instructions

At the beginning of the experiment you will be matched with four other people, randomly selected from the participants in this room, to form a group of five. **The composition of your group will stay the same throughout the experiment**, i.e. you will form a group with the same four other participants during the whole experiment.

Each person in the group will be randomly assigned a role, either 'group member A', 'group member B', 'group member C', 'group member D', or 'group member E'. **Your role will stay the same throughout the experiment**. You will not learn the identity of the other participants in your group. Participants will be identified simply as 'group member A', 'group member B' etc..

When we have finished reading the instructions you will be informed of your role. **The experiment will then consist of 10 periods**. In each period you can earn points. Your point earnings will depend on the decisions made within your group, and will not be affected by decisions made in other groups.

At the end of the experiment your accumulated point earnings from all periods will be converted into cash at the exchange rate of **10 pence per 100 points**. You will be paid, in cash and in private, this amount in addition to the show-up fee of £3.50.

Decisions and Earnings

The experiment consists of ten periods in which you can earn points. Every period has the same structure and has two stages.

Stage One

At the beginning of Stage One each group member will be given 10 tokens.

Group members A, B, C and D must choose how many of these tokens to invest in a group project and how many to keep in their private accounts. Group member E has no choice to make: the computer will place all ten of his or her tokens in his or her private account.

Each token a group member keeps in his or her private account yields a return of thirty points to that group member.

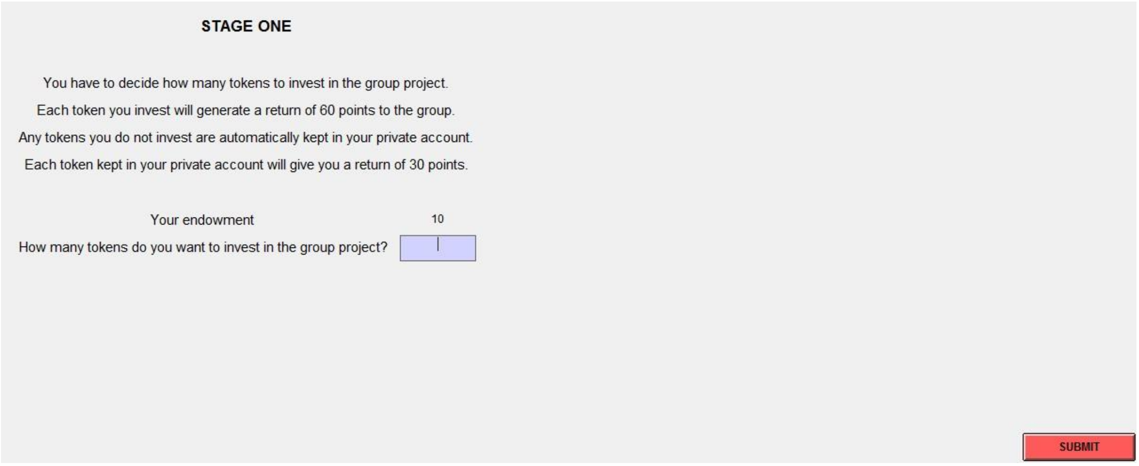
[Homogeneous: Each token a group member invests in the group project yields a return of sixty points to the group. How these points are allocated among group members will be determined in Stage Two.**]**

[Heterogeneous: Each token group member A or B invests in the group project yields a return of forty points to the group. Each token group member C or D invests in the group project yields a return of eighty points to the group. How these points are allocated among group members will be determined in Stage Two.**]**

Your point earnings for the period will be your point earnings from your private account plus your point earnings from the group project.

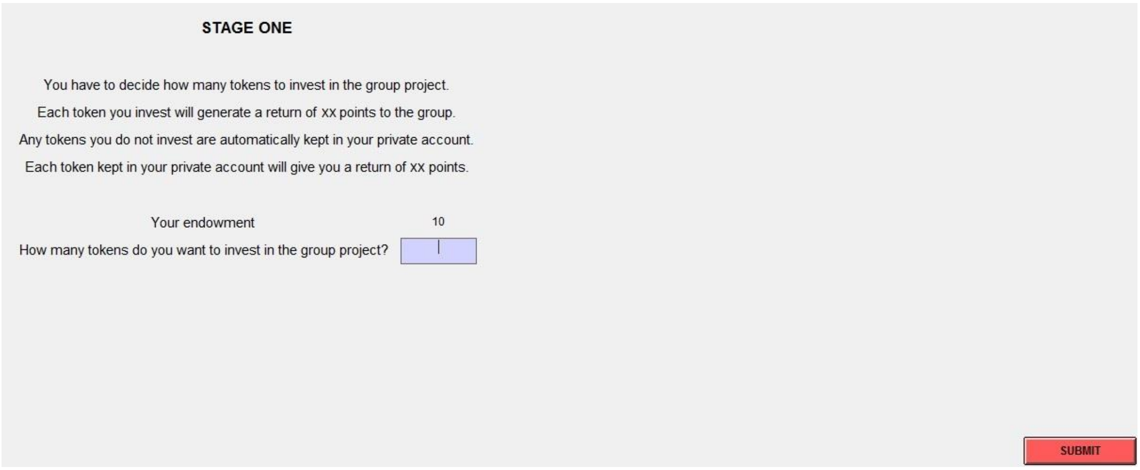
If you are one of group members A, B, C or D, you will make your decision by entering the number of tokens you invest in the group project. Any tokens you do not invest will automatically be kept in your private account. You will enter your decisions on a screen like the one shown below.

[Homogeneous:



]

[Heterogeneous:



]

When you have made your decision you must click on the *SUBMIT* button. Once group members A, B, C and D have submitted their decisions Stage Two will begin.

Stage Two

[Homogeneous: Remember, each token invested in the group project generates a return of sixty points to the group. Thus, the total return from the group project will be 60 times the total number of tokens invested in the group project. How this return is distributed among group members is determined as follows.**]**

[Heterogeneous: Remember, each token invested in the group project by group member A or B yields a return of forty points to the group. Each token invested in the group project by group member C or D yields a return of eighty points to the group. How this return is distributed among group members is determined as follows.**]**

[NoLeader_Homogeneous: Each of group members A, B, C, and D will receive an equal share of the total return from the group project. This means that each group member receives 60 / 4 = 15 points per token for each token invested in the group project.

Thus, **if you are group member A, B, C or D**, your earnings for the period will be:

> **Your point earnings = 30 x (number of tokens kept in your private account) + 15 x (total number of tokens invested in the group project by your group).]**

[**Control_Homogeneous:** All group members (A, B, C, D and E) will receive an equal share of the total return from the group project. This means that each group member receives 60 / 5 = 12 points per token for each token invested in the group project.

Thus, your earnings for the period will be:

> **Your point earnings = 30 x (number of tokens kept in your private account) + 12 x (total number of tokens invested in the group project by your group).]**

[Control_Heterogeneous: Each of group members A, B, C, and D will receive an equal share of the total return from the group project. This means that for each token invested in the group project by A or B each group member receives 40 / 4 = 10 points per token, and for each token invested in the group project by C or D each group member receives 80 / 4 = 20 points per token.

Thus, **if you are group member A, B, C or D**, your earnings for the period will be:

Your point earnings = 30 x (number of tokens kept in your private account) +

10 x (total number of tokens invested in the group project by A and B) +

20 x (total number of tokens invested in the group project by C and D).]

[NoLeader Heterogeneous Additional: All group members (A, B, C, D and E) will receive an equal share of the total return from the group project. This means that for each token invested in the group project by A or B each group member receives 40 / 5 = 8 points per token, and for each token invested in the group project by C or D each group member receives 80 / 5 = 16 points per token.

Thus, your earnings for the period will be:

Your point earnings = 30 x (number of tokens kept in your private account) +
8 x (total number of tokens invested in the group project by A and B) +
16 x (total number of tokens invested in the group project by C and D).]

[NoLeader: If you are group member E you do not get any of the return from the group project. Thus, **if you are group member E** your earnings for the period will be

Your point earnings = 30 x 10 = 300.]

[Leader: Group member E will be informed of the decisions of the other group members and must decide how to allocate the total return from the group account among all group members, including himself or herself. Group member E is free to choose any allocation he or she wants, as long as each group member receives at least zero points from the group project and the total received by all group members is equal to the total return from the group project.

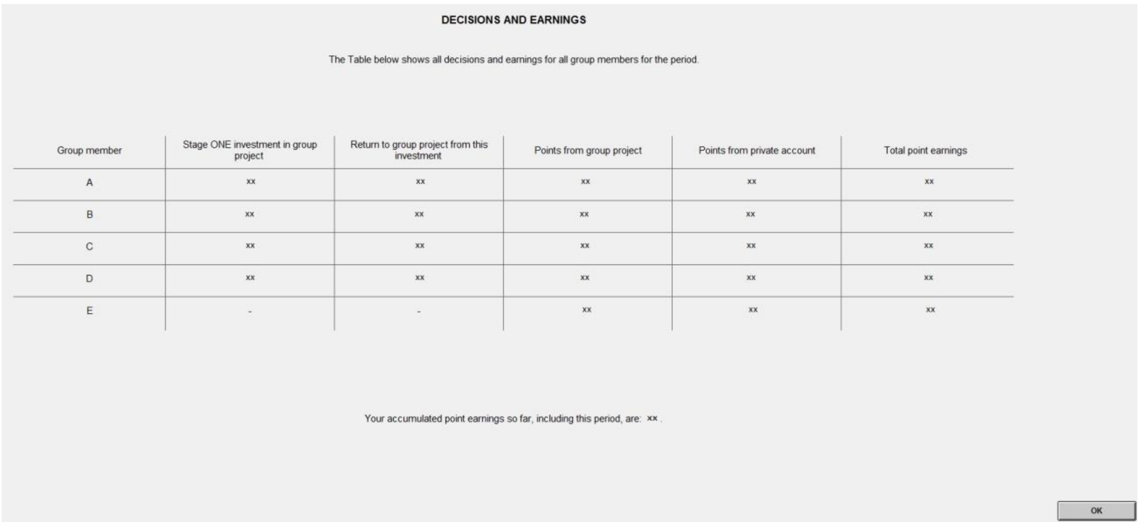
To do this Group Member E will complete a screen like the one shown below.

The first column shows the group member ID and the second column shows how many tokens this group member invested in the group project in Stage One. The third column shows the resulting return to the group from this investment decision. We will explain the fourth column in a moment. The fifth column is to be completed by group member E and shows how many points from the group project that this group member receives. The sixth column shows this group member's point earnings from his or her private account and the last column shows this group member's point earnings for the period.

Group member E completes the fifth column by allocating the total return from the group project. The total return is shown above the table in bold, and the amount that still has to be allocated is shown at the top of the fourth column. In each row of the fourth column there are add (+) / subtract (-) buttons that group member E can use to allocate the total return. For example, if group member E clicks on the first add button 100 points will be added to group member A's point earnings from the group project. At the same time the entry in the final column for group member A will increase by 100 points and the amount still to be allocated will be reduced by 100 points. Thus, Group member E can easily see how total earnings and the amount left to be allocated change as he or she allocates the return. Group member E must allocate the total return from the group project among the five group members. Once group member E has allocated the total return he or she can either amend her decisions using the add/subtract buttons, or submit the decision by clicking the *SUBMIT* button.

]

At the end of the period all group members will be shown a *Decision and Earnings* screen like the one shown below.


DECISIONS AND EARNINGS

The Table below shows all decisions and earnings for all group members for the period.

Group member	Stage ONE investment in group project	Return to group project from this investment	Points from group project	Points from private account	Total point earnings
A	xx	xx	xx	xx	xx
B	xx	xx	xx	xx	xx
C	xx	xx	xx	xx	xx
D	xx	xx	xx	xx	xx
E	-	-	xx	xx	xx

Your accumulated point earnings so far, including this period, are: xx .

OK

The screen shows all decisions and earnings for all group members for the period. At the bottom of the screen you will also see your total point earnings that you have accumulated from this and previous periods.

After you have read the information on the Decisions and Earnings screen you can click the *OK* button to continue. Once all group members have done this, the next period will begin.

At the end of period ten you will see your total point earnings from all periods and you will be paid 10p for every 100 points, in addition to your £3.50 show-up fee. You will be paid in private and in cash.

Questions

Please answer the questions below. The example in the questions is purely hypothetical. In the actual experiment the investments in the group project **[Leader:** and the allocation of the total return**]** will be determined by the decisions made in your group. In a couple of minutes someone will come to your desk to check your answers. When each participant has answered all questions correctly we will continue with the experiment.

1. How many periods will there be? __________

2. Will the people in your group be the same for period to period or change from period to period? Same / change

Suppose that in a period A invests 0, B invests 4 tokens, C invests 6 and D invests 10 tokens in the group project.

3. What will be the total return from the group project in that period? __________

[NoLeader:

4. How many points will each group member earn from the group project in that period?

A: __________

B: __________

C: __________

D: __________

E: __________

5. How many points will each group member earn from his or her private account in that period?

A: __________

B: __________

C: __________

D: __________

E: __________

6. What will be each group member's point earnings for the period?

A: __________

B: __________

C: __________

D: __________

E: __________

]

[Leader:

4. Suppose E distributes the total return from the group project as shown in the Table below. Complete the blanks in the last two columns.

Group member	Stage ONE investment in group project	Return to group project	Points from group project	Points from private account	Total point earnings
A	4	240 [160]	300		
B	6	360 [240]	100		
C	10	600 [800]	240		
D	0	0	360		
E	-	-	200		

]

[Endogenous: Instructions

At the beginning of the experiment you will be matched with four other people, randomly selected from the participants in this room, to form a group of five. **The composition of your group will stay the same throughout the experiment**, i.e. you will form a group with the same four other participants during the whole experiment.

Each person in the group will be randomly assigned a role, either 'group member A', 'group member B', 'group member C', 'group member D', or 'group member E'. **Your role will stay the same throughout the experiment**. You will not learn the identity of the other participants in your group. Participants will be identified simply as 'group member A', 'group member B' etc..

When we have finished reading the instructions you will be informed of your role. **The experiment will then consist of 3 blocks of 10 periods each**. In each period you can earn points. Your point earnings will depend on the decisions made within your group, and will not be affected by decisions made in other groups.

At the end of the experiment your accumulated point earnings from all periods will be converted into cash at the exchange rate of **10 pence per 100 points**. You will be paid, in cash and in private, this amount in addition to the show-up fee of £3.50.

Decisions and Earnings

The experiment consists of three blocks of ten periods in which you can earn points. Every period has the same structure and has two stages.

Stage One

At the beginning of Stage One each group member will be given 10 tokens.

Group members A, B, C and D must choose how many of these tokens to invest in a group project and how many to keep in their private accounts. Group member E has no choice to make: the computer will place all ten of his or her tokens in his or her private account.

Each token a group member keeps in his or her private account yields a return of thirty points to that group member.

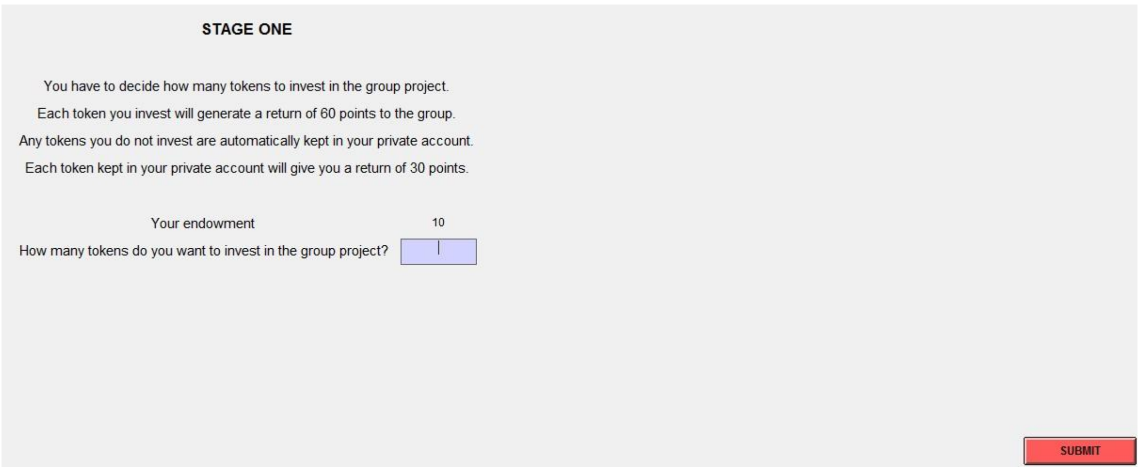
[Homogeneous: Each token a group member invests in the group project yields a return of sixty points to the group. How these points are allocated among group members will be determined in Stage Two.**]**

[Heterogeneous: Each token group member A or B invests in the group project yields a return of forty points to the group. Each token group member C or D invests in the group project yields a return of eighty points to the group. How these points are allocated among group members will be determined in Stage Two.**]**

Your point earnings for the period will be your point earnings from your private account plus your point earnings from the group project.

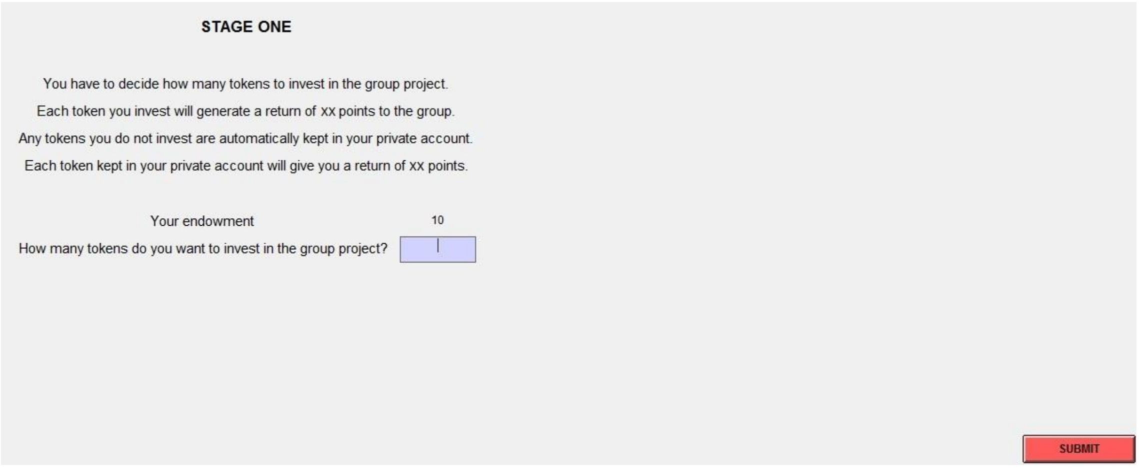
If you are one of group members A, B, C or D, you will make your decision by entering the number of tokens you invest in the group project. Any tokens you do not invest will automatically be kept in your private account. You will enter your decisions on a screen like the one shown below.

[Homogeneous:



]

[Heterogeneous:



]

When you have made your decision you must click on the *SUBMIT* button. Once group members A, B, C and D have submitted their decisions Stage Two will begin.

Stage Two

[Homogeneous: Remember, each token invested in the group project generates a return of sixty points to the group. How this return is distributed among group members is determined as follows.**]**

[Heterogeneous: Remember, each token group member A or B invests in the group project yields a return of forty points to the group. Each token group member C or D invests in the group project yields a return of eighty points to the group. How this return is distributed among group members is determined as follows.**]**

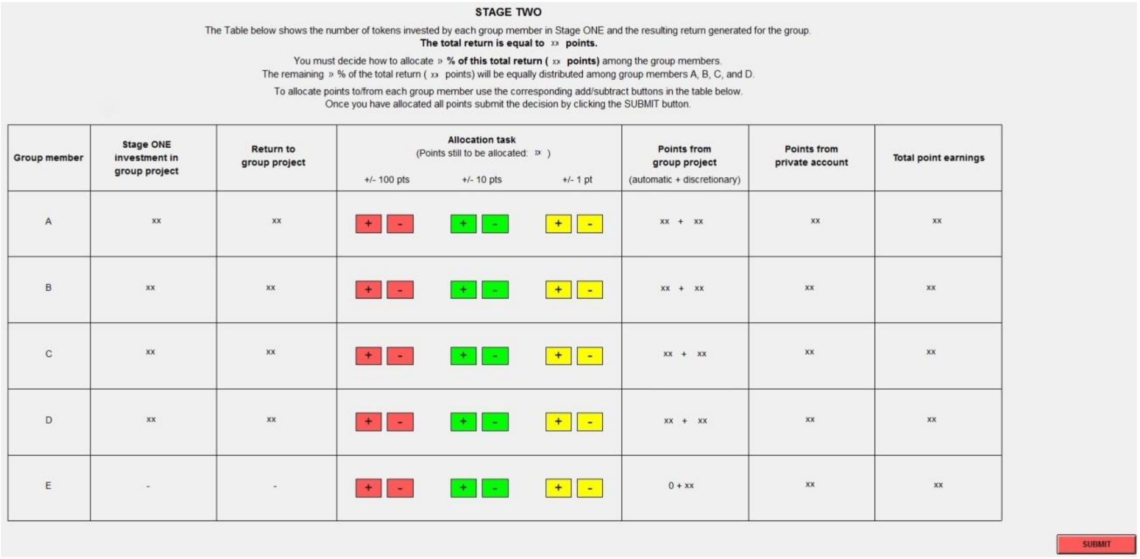
The total return will be divided into two parts:

total return = automatic return + discretionary return.

The **automatic return** will be equally distributed among group members A, B, C and D. The **discretionary return** will be allocated by group member E. We will explain what part of the total return will be automatic and what part will be discretionary later.

For the discretionary part, group member E will be informed of the decisions of the other group members and must decide how to allocate the discretionary return among all group members, including himself or herself. Group member E is free to choose any allocation he or she wants, as long as each group member receives at least zero discretionary points from the group project and the total discretionary points received by all group members equals the discretionary return.

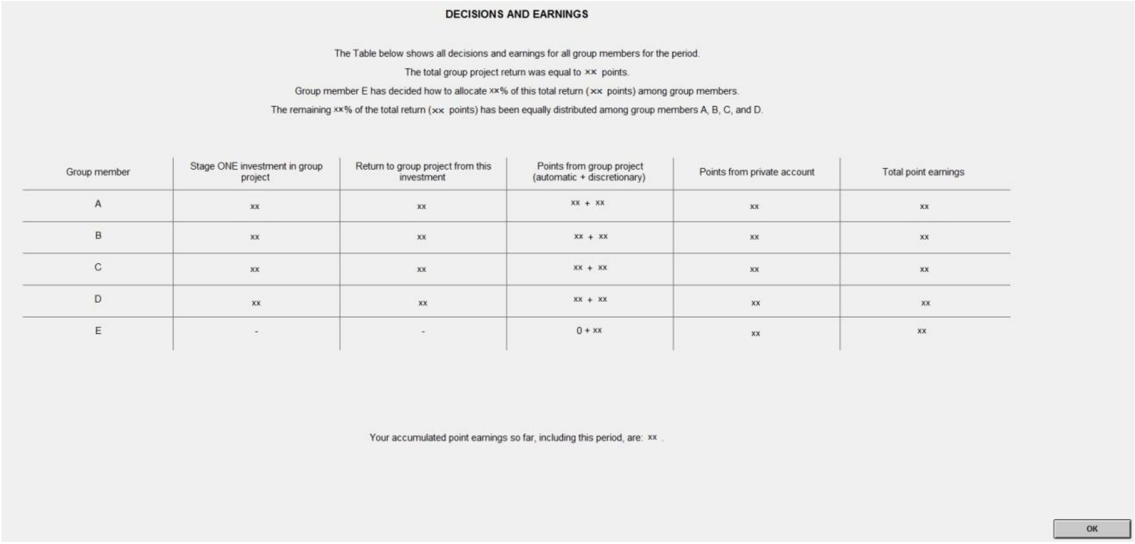
To do this group member E will complete a screen like the one shown below.



The first column shows the group member ID and the second column shows how many tokens this group member invested in the group project in Stage One. The third column shows the resulting return to the group from this investment decision. We will explain the fourth column in a moment. The fifth column is to be completed by group member E and shows how many points from the group project that this group member receives. Note that in each row of this column there are two numbers, the first is the earnings from the group project this group member automatically receives, and the second is the discretionary earnings from the group project that this group member is allocated by group member E. The sixth column shows this group member's point earnings from his or her private account and the last column shows this group member's point earnings for the period.

Group member E completes the fifth column by allocating the discretionary return from the group project. The total return from the group project is shown above the table in bold, and the discretionary return that group member E must allocate is shown on the next line. The amount that still has to be allocated by group member E is shown at the top of the fourth column. In each row of the fourth column there are add (+) / subtract (-) buttons that group member E can use to allocate the return. For example, if group member E clicks on the first add button 100 points will be added to group member A's point earnings from the group project. At the same time the entry in the final column for group member A will increase by 100 points and the amount still to be allocated will be reduced by 100 points. Thus, group member E can easily see how total earnings and the amount left to be allocated change as he or she allocates the return. Group member E must allocate all of the discretionary return among the five group members. Once group member E has allocated all of the discretionary return he or she can either amend his or her decisions using the add/subtract buttons, or submit the decision by clicking the *SUBMIT* button.

At the end of the period all group members will be shown a *Decision and Earnings* screen like the one shown below.

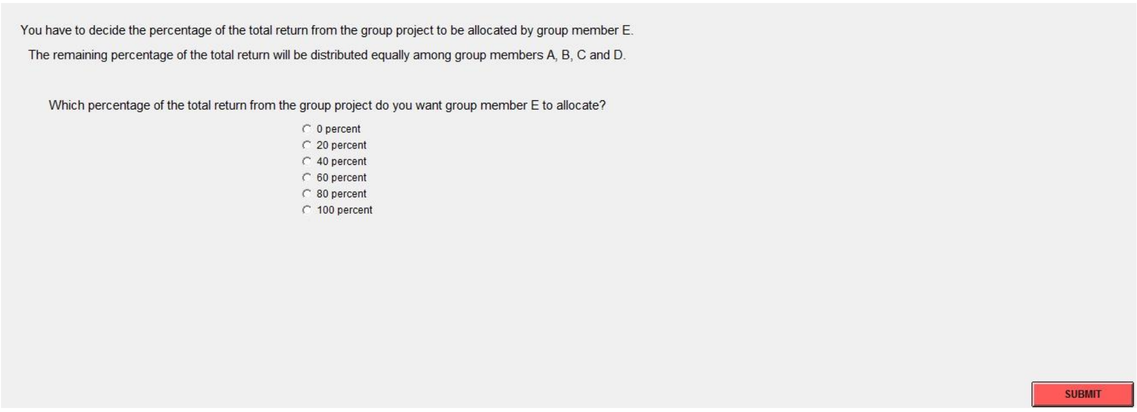


The screen shows all decisions and earnings for all group members for the period. At the bottom of the screen you will also see your total point earnings that you have accumulated from this and previous periods.

After you have read the information on the Decisions and Earnings screen you can click the *OK* button to continue. Once all group members have done this, the next period will begin.

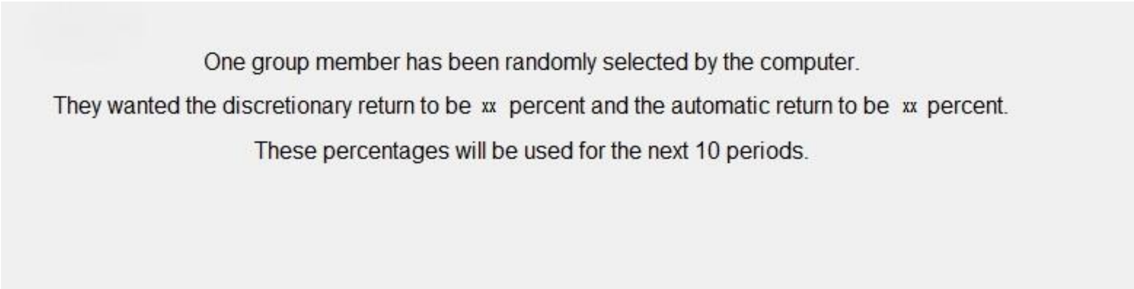
Beginning a Block

At the beginning of a block (that is, just before periods 1, 11 and 21), each of group members A, B, C and D will see the following screen.



Each group member has to indicate how he or she wants the total return from the group project to be divided between the discretionary return and automatic return. Each group member has to choose a percentage (out of 0%, 20%, 40%, 60%, 80% and 100%). This is the percentage of the total return from the group project that he or she wants to be the discretionary return, that is, the percentage to be allocated by group member E. The remaining percentage is the percentage of the total return from the group project that this group member wants to be distributed automatically.

After group members A, B, C and D have submitted their decisions, one of the four decisions will be selected at random by the computer and will be used to determine how the total return is distributed during the block of ten periods. The randomly selected group member's decision will be used for all ten periods in the block. All group members will be notified of the randomly selected decision on a screen like the one shown below.



Notice that if the discretionary return is 0% of the total return from the group project then group member E will have no decision to make. In this case group member E will still see the stage two decision screen, but he or she cannot change the entries in the table and he or she should just click on the *SUBMIT* button.

Ending the Experiment

At the end of period thirty you will see your total point earnings from all periods and you will be paid 10p for every 100 points, in addition to your £3.50 show-up fee. You will be paid in private and in cash.

Questions

Please answer the questions below. The example in the questions is purely hypothetical. In the actual experiment the investments in the group project and the allocation of the total return will be determined by the decisions made in your group. In a couple of minutes someone will come to your desk to check your answers. When each participant has answered all questions correctly we will continue with the experiment.

1. How many blocks will there be there be in this experiment? __________

2. How many periods will there be there be in each block? __________

3. Will the people in your group be the same from block to block or change from block to block? Same / change

4. Will the people in your group be the same from period to period or change from period to period? Same / change

5. Suppose that just before period 21 group member A chose the discretionary return 0%, B chose 20%, C chose 80%, and D chose 100%. Group member C's choice was randomly selected by the computer. What percentage of total return from the group project will be allocated by group member E in periods 21 to 30? __________

Suppose that for this block of periods the discretionary return is 0% and the automatic return is 100% of the total return from the group project. Suppose also that in a period A invests 4, B invests 6 tokens, C invests 10 and D invests 0 tokens in the group project.

6. What will be the return to the group project from each group member's investment in that period?

A: __________

B: __________

C: __________

D: __________

E: __________

7. What will be the total return from the group project in that period? __________

8. How many points will each group member earn from the group project in that period?

A: __________

B: __________

C: __________

D: __________

E: __________

9. How many points will each group member earn from his or her private account in that period?

A: __________

B: __________

C: __________

D: __________

E: __________

10. What will be each group member's point earnings for the period?

A: __________

B: __________

C: __________

D: __________

E: __________

Suppose that for this block of periods the discretionary return is 100% and the automatic return is 0% of the total return from the group project. Suppose also that in a period A invests 4, B invests 6 tokens, C invests 10 and D invests 0 tokens in the group project.

11. Suppose E allocates the total return from the group project as shown in the Table below. Complete the blanks in the last two columns.

Group member	Stage ONE investment in group project	Return to group project	Points from group project	Points from private account	Total point earnings
A	4	240 [160]	300		
B	6	360 [240]	100		
C	10	600 [800]	240		
D	0	0	360		
E	-	-	200		

]

Online Appendix B: Additional treatments

Figure B.1 Effort choice in control and NoLeader treatments across rounds

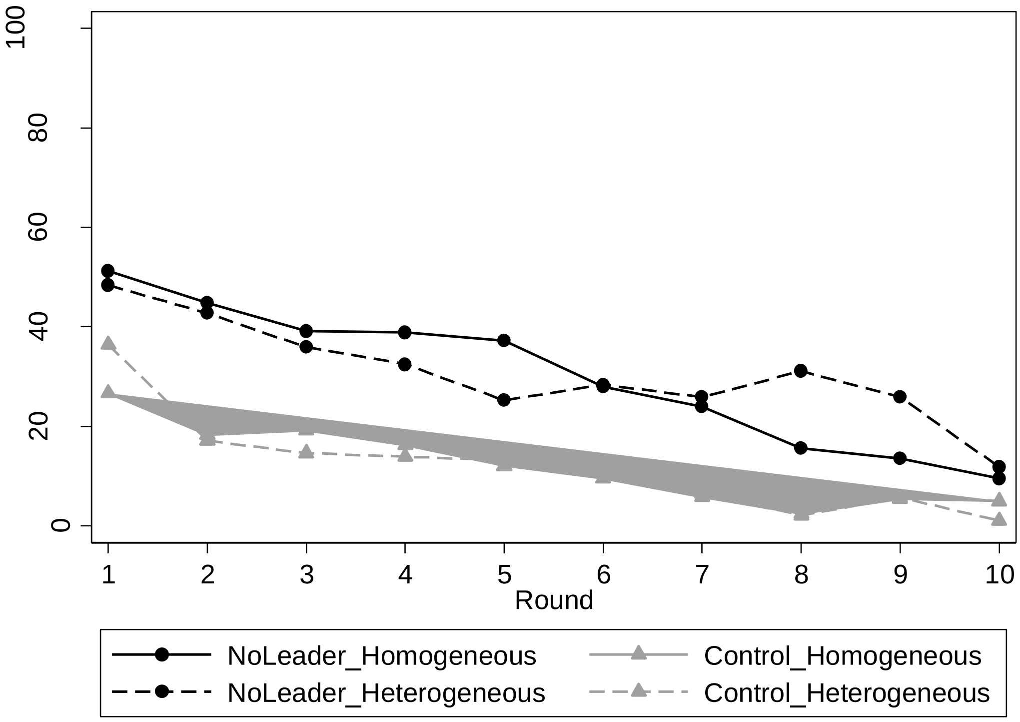


Table B.1 Summary statistics of efforts in the Control treatments and non-parametric comparisons with NoLeader treatments.

	Homogeneous			Heterogeneous		
	NoLeader	Control	p-value	NoLeader	Control	p-value
Effort (round 1)	51.25	26.67	0.002			
Low Productivity				48.12	41.43	0.623
High Productivity				48.64	31.43	0.102
Effort (all rounds)	30.21	12.08	0.007			
Low Productivity				27.86	11.57	0.021
High Productivity				33.73	12.93	0.007

Effort measured as a percentage of maximum. P-values based on Mann-Whitney two-sided tests.

Online Appendix C. Additional Analysis for Study 2

In this Appendix we report additional analysis of effort and redistribution behavior in Study 2. Starting with effort choices, Figures C.1 and C.2 report the average effort supplied in each of the three blocks of the Homogeneous (Figure C.1) and Heterogeneous (Figure C.2) treatments, disaggregated by the realized level of the leader's reward power.

Teams with powerful leaders supply considerably more effort than teams with powerless leaders. The separation between teams with powerful and powerless leaders is clearest in the Homogeneous treatment (Figure C.1), where the effort trajectories are generally higher the higher the share of team output allocated by the leader. There is a similar pattern in the Heterogeneous treatment (Figure C.2), although there are some anomalies (e.g. the low effort in teams with $\gamma = 1$ in Block 2); these anomalies may well reflect the low numbers of observations in some cases (the number of teams that have delegated a level of reward power is shown in the legend of each panel).

Figure C.1 – Average effort in Homogeneous

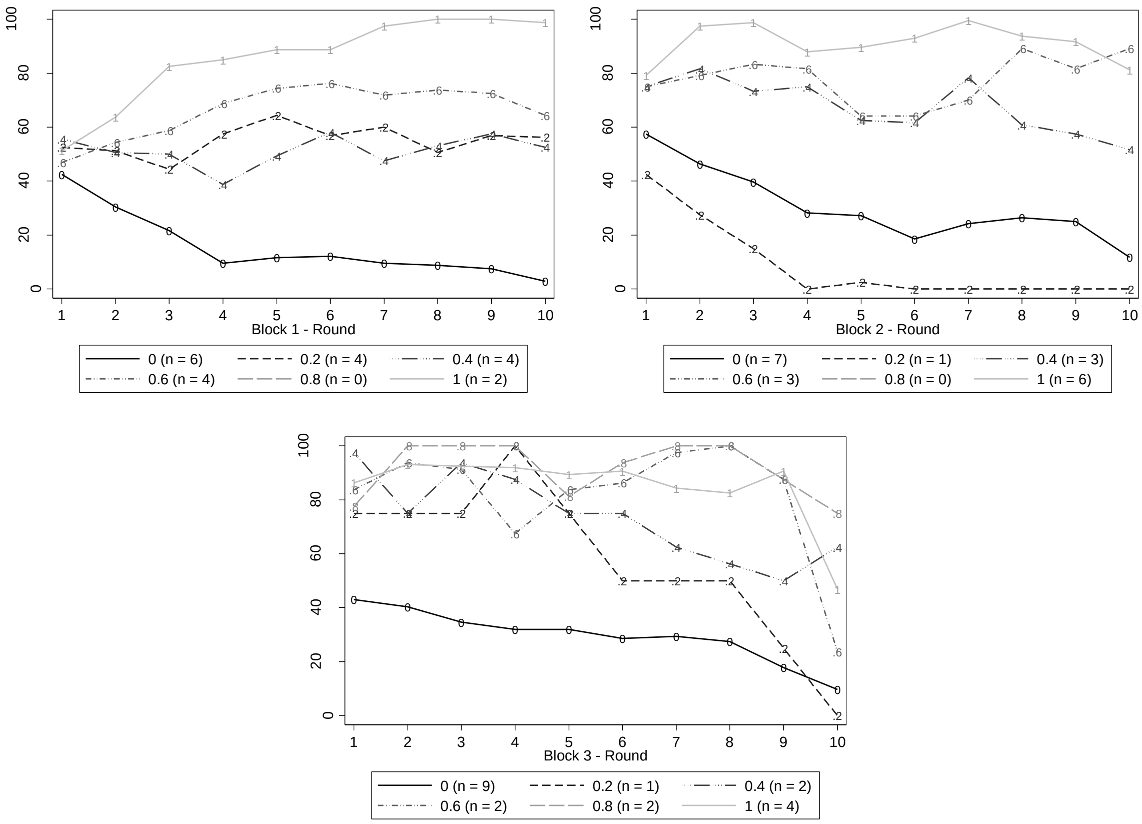
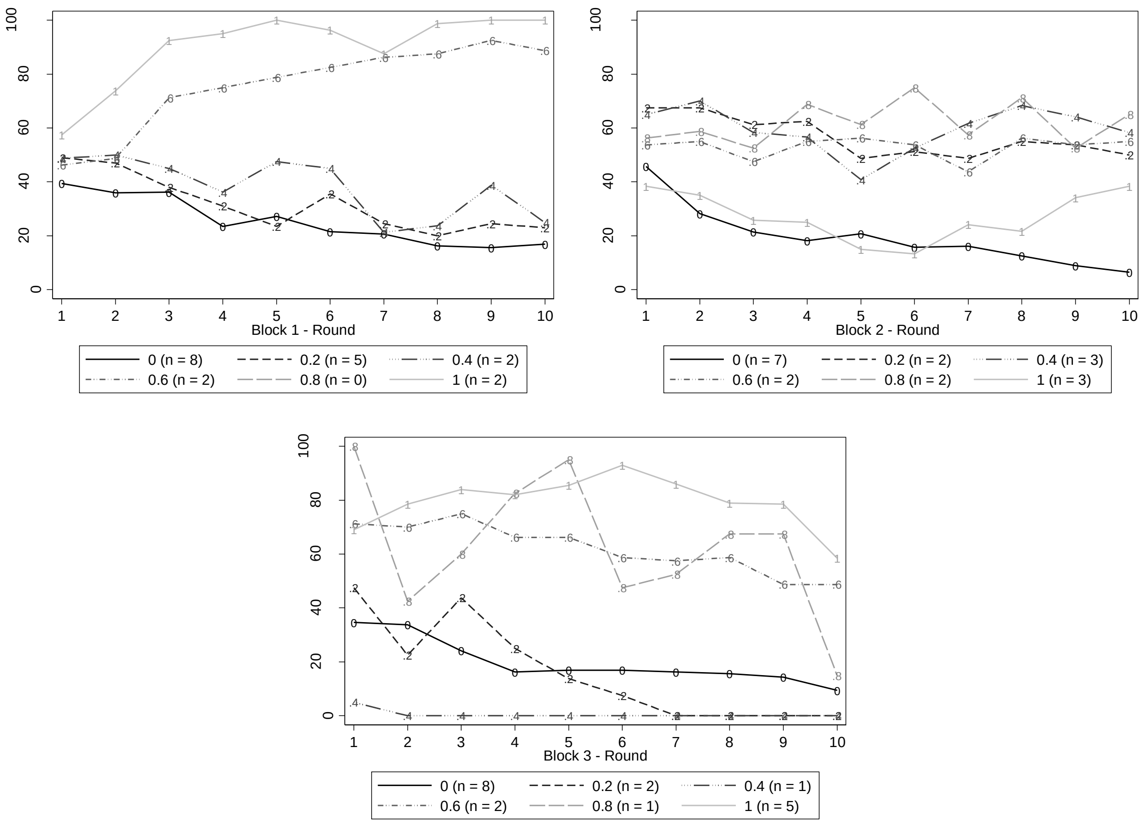


Figure C.2 – Average effort in Heterogeneous



We explore the effect of leader power on effort in Table C.1, where we report fixed-effects regressions of effort on dummy variables for the level of reward power delegated to the leader (note that the benchmark category is the case where $\gamma = 0$, i.e. the team has no leader). We report separate regressions for the Homogeneous and Heterogeneous treatments. For the Heterogeneous treatment, we report separate regressions for low and high productivity workers. All regressions also include round and block variables as well as a dummy variable assuming value 1 for observations in the last round of each block. [1]

Table C.1 – Determinants of workers' effort

	Homogeneous treatment	Heterogeneous treatment	
		Low Productivity Workers	High Productivity Workers
0.2	6.37	0.77	3.63
	(6.89)	*(6.96)*	*(7.35)*
0.4	-13.52	14.16	12.71
	(10.70)	*(13.23)*	*(9.90)*
0.6	31.60***	12.71	17.29**
	(5.54)	*(13.67)*	*(7.88)*
0.8	52.97***	12.89	21.17**
	(7.18)	*(11.75)*	*(9.51)*
1	55.50***	38.53***	32.73***
	(9.05)	*(11.58)*	*(9.26)*
Round	-1.09**	-1.57***	-1.74***
	(0.51)	*(0.54)*	*(0.57)*
Last Round Dummy	-9.72***	-0.99	-0.00
	(2.56)	*(2.45)*	*(2.91)*
Block	0.95	-2.42	-2.69
	(3.31)	*(2.53)*	*(2.58)*
Constant	44.13***	39.35***	48.29***
	(10.08)	*(8.06)*	*(7.97)*
N. of observations	2400	1140	1140
R^2	.318	.190	.136

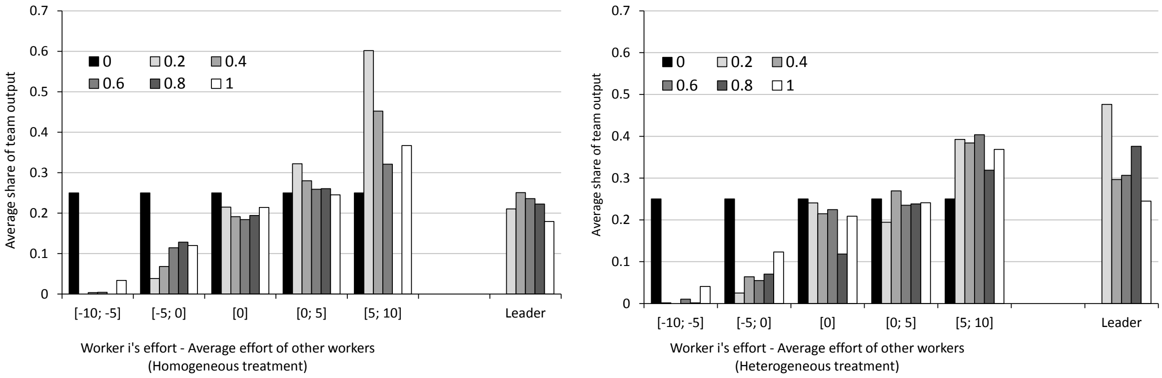
Notes: Fixed-effects regressions with robust standard errors clustered at the team level reported in parentheses (20 clusters in Homogeneous; 19 clusters in Heterogeneous). Dependent variable is workers' effort expressed as percentage of maximum possible effort. Significance levels: *** = 1%; ** = 5%; * = 10%.

The regressions confirm that delegation of reward power increases effort provision. The effect is particularly strong in Homogeneous where delegating full reward power to leaders ($\gamma = 1$) increases productivity by about 55 percentage points relative to teams without a leader. In Heterogeneous the effect of delegation is somewhat weaker: for low productivity workers the effect is only statistically significant when the team delegates full power to leaders; for high productivity workers the effect is statistically significant when 60% or more of the team output is delegated to the leader.

[1] Using random effects regressions with controls for individual characteristics (gender, risk attitude, field of study) produces very similar results.

The positive effect of leadership on effort is driven by the redistribution strategies adopted by leaders, which reward workers who supply effort and withhold rewards from those who shirk. Figure C.3 shows the average share of team output received by leaders and workers disaggregated by level of reward power of the leader. The left panel contains data from the Homogeneous treatment whereas the right panel contains data from the Heterogeneous treatment.[2] Workers' output shares are disaggregated based on how their efforts relate to the average effort of the rest of their team.

Figure C.3 – Share of team output received



Notes: When $\gamma = 0$ the leader is assigned no share of the team output and all output is redistributed in equal proportions among workers. When $\gamma > 0$ the figure shows how the share of team output assigned to the leader is redistributed among workers and leaders.

In both treatments, when the leader has no reward power ($\gamma = 0$) each worker automatically receives a quarter of the team output and the leader receives nothing. When leaders are granted some reward power ($\gamma > 0$), leaders redistribute output by giving higher rewards to those who supply more effort. In the Homogeneous treatment workers who supply effort at or above the average effort in their team receive on average 25% of the team output assigned to the leader. In contrast, those who provide less effort than the team average receive only 7% of team output. The difference is highly significant (Wilcoxon signed-rank test, $p < 0.001$). [3]Similar patterns emerge in the Heterogeneous treatment: workers supplying effort at or above the team average receive 23% of team output, whereas workers supplying effort below the team average receive 5% of team output (Wilcoxon signed-rank test, $p = 0.001$). This is true both in the case of leaders who have been allocated full reward power, i.e.

[2] For simplicity, in Figure C.3 we do not distinguish between high and low productivity workers. We do not observe substantial differences in rewards received by high and low productivity workers, although workers who contribute 5 or more units of effort in excess of team average tend to receive higher rewards if they are high rather than low productivity workers.

[3] This and subsequent non-parametric tests use team averages as the independent unit of observation.

$\gamma = 1$ (p = 0.012 for Homogeneous and p = 0.008 for Heterogeneous) and partial reward power, i.e. $1 > \gamma > 0$ (p = 0.001 for both Homogeneous and Heterogeneous).

Figure C.3 also shows that leaders tend to appropriate a large fraction of team output, about 20% in Homogeneous and about 30% in Heterogeneous. Thus, delegation of reward power is beneficial to the leaders rather than to workers, as shown in Table C.2. In both treatments, combined earnings and efficiency are positively related with leaders' power. However, in both treatments the leaders reap most of these efficiency gains. Leaders' earnings increase by about 119% from 300 points in the case where they are powerless ($\gamma = 0$) to 658 points in the case where they have full reward power ($\gamma = 1$). Workers' earnings also tend to be higher when leaders have more reward power, but the impact of leadership on workers' earnings is not nearly as strong as in the leaders' case. In Homogeneous workers' earnings increase on average by 26% between the $\gamma = 0$ and $\gamma = 1$ cases. In Heterogeneous, high productivity workers' earnings increase by 16% while low productivity workers' earnings increase by just 2%.

Table C.3 reports fixed-effects regressions of subjects' earnings on a set of dummy variables for the level of reward power delegated to the leader as well as controls for number of blocks and rounds. We report separate regressions for the Homogeneous and Heterogeneous treatments, and for leaders and workers.[4] The regressions confirm that leader power has a strong, positive effect on leaders' earnings. In both treatments, leaders earn significantly more when they are granted reward power, for nearly all levels of power. In the Homogeneous treatment workers' earnings are also positively associated with leader power, although the effect is small compared with the leaders' case. Moreover, workers' earnings when leaders have power greater than 0.2 are not significantly different from earnings when leaders have a power of 0.2 (F-tests: p = 0.202 for $\gamma = 0.4$; p = 0.239 for $\gamma = 0.6$; p =0.254 for $\gamma = 0.8$; p = 0.132 for $\gamma = 1$). Thus while delegating some power to the leader may be profitable, it does not seem to pay to delegate a lot of power. Leadership is not beneficial for low productivity workers in heterogeneous teams: earnings are not significantly different between the $\gamma = 0$ case and most of the cases where leaders have reward power. In one case ($\gamma = 0.8$), workers' earnings are lower than when $\gamma = 0$ and the effect is significant at the 10% level. For high productivity workers, leadership has no impact on workers' earnings.

[4] Again, we also conducted random effects regressions with controls for individual characteristics and obtained very similar results.

Table C.2 – Individual earnings and efficiency

	Homogeneous treatment γ					
	0	0.2	0.4	0.6	0.8	1
Leader's Earnings	300 *(0.0)*	350 *(31.6)*	439.8 *(89.4)*	533 *(112.6)*	691.9 *(18.5)*	658.7 *(153.9)*
Workers' Earnings	378.1 *(65.1)*	430.8 *(72.7)*	450.2 *(66.8)*	462.1 *(50.6)*	476.5 *(26.9)*	474.6 *(64.8)*
Combined Earnings	1812.6 *(260.6)*	2073 *(316.9)*	2240.7 *(340.7)*	2381.3 *(217.2)*	2598 *(89.1)*	2557.8 *(158.2)*
Efficiency	26%	48%	62%	73%	91%	88%

	Heterogeneous treatment γ					
	0	0.2	0.4	0.6	0.8	1
Leader's Earnings	300 *(0.0)*	365.4 *(30.5)*	437.9 *(115.5)*	464 *(100.9)*	772.5 *(193.2)*	658.4 *(192.3)*
Low Productivity Worker	383.3 *(97.3)*	399.6 *(94.0)*	413.8 *(86.2)*	439.3 *(106.0)*	353.9 *(44.3)*	390.1 *(59.9)*
High Productivity Worker	361.1 *(33.4)*	389.5 *(77.4)*	406.1 *(80.4)*	482.9 *(123.2)*	411.7 *(63.8)*	420 *(88.6)*
Combined Earnings	1788.8 *(238.8)*	1943.7 *(341.1)*	2077.7 *(401.4)*	2308.5 *(413.3)*	2303.7 *(198.3)*	2278.6 *(361.9)*
Efficiency	24%	37%	48%	67%	67%	65%

Notes: "Combined Earnings" are the sum of leader's and workers' earnings. "Efficiency" is computed as (combined earnings – 1500) / (2700 – 1500), where 1500 are the theoretical earnings under zero contributions and 2700 are maximum possible combined earnings. Standard deviations based on team averages per block in parentheses.

Table C.3 – Regression of earnings on leaders' power

	Homogeneous		Heterogeneous		
	Leaders	Workers	Leaders	Low Productivity Workers	High Productivity Workers
0.2	-1.40 *(30.73)*	19.47 *(23.17)*	38.63 *(30.73)*	-3.24 *(25.3.)*	2.86 *(16.03)*
0.4	60.20 *(62.81)*	-55.59 *(41.61)*	142.06** *(54.05)*	-16.35 *(39.81)*	23.02 *(28.56)*
0.6	200.15*** *(41.70)*	44.77** *(20.15)*	99.37 *(68.23)*	-4.31 *(32.70)*	53.79 *(43.98)*
0.8	366.74*** *(42.34)*	67.21** *(26.71)*	364.49*** *(83.98)*	-71.30* *(35.31)*	7.79 *(50.83)*
1	390.23*** *(55.49)*	68.93* *(34.09)*	358.36*** *(59.73)*	-11.67 *(30.34)*	34.65 *(25.77)*
Round	6.51* *(3.53)*	-4.90*** *(1.48)*	-0.05 *(2.53)*	-3.93* *(2.17)*	-6.33*** *(1.65)*
Last Round Dummy	-29.40 *(34.08)*	-21.81** *(9.01)*	39.02 *(49.88)*	-19.48 *(17.37)*	-1.03 *(12.50)*
Block	11.47 *(9.97)*	-0.01 *(9.76)*	22.81 *(15.09)*	-6.92 *(10.07)*	-20.35*** *(6.36)*
Constant	260.74*** *(35.90)*	442.20*** *(31.39)*	267.04*** *(42.55)*	440.53*** *(26.51)*	456.76*** *(15.33)*
N. of observations	600	2400	570	1140	1140
R^2	.399	.146	.281	.054	.103

Notes: Fixed-effects regressions with robust standard errors clustered at the team level reported in parentheses (20 clusters in Homogeneous; 19 clusters in Heterogeneous). Dependent variable is earnings. Significance levels: *** = 1%; ** = 5%; * = 10%.