universität innsbruck

Research Platform
Empirical and Experimental Economics

Competitive versus cooperative incentives in team production with heterogeneous agents

Glenn E. Dutcher, Regine Oexl, Dmitry Ryvkin, Tim Salmon

Working Papers in Economics and Statistics

2021-26

University of Innsbruck
http://uibk.ac.at/eeecon/

University of Innsbruck
Working Papers in Economics and Statistics

The series is jointly edited and published by

- Department of Banking and Finance
- Department of Economics
- Department of Public Finance
- Department of Statistics

Contact address of the editor:
research platform “Empirical and Experimental Economics”
University of Innsbruck
Universitaetsstrasse 15
A-6020 Innsbruck
Austria
Tel: + 43 512 507 71022
Fax: + 43 512 507 2970
E-mail: eeecon@uibk.ac.at

The most recent version of all working papers can be downloaded at
https://www.uibk.ac.at/eeecon/wopec/

For a list of recent papers see the backpages of this paper.

Competitive versus cooperative incentives in team production with heterogeneous agents.*

E. Glenn Dutcher[†], Regine Oexl[‡], Dmitry Ryvkin[§], Timothy C.Salmon[¶]

This version: September 11, 2021

Abstract

A debate among practicing managers is whether to use cooperative or competitive incentives for team production. While competitive incentives may drive individual effort higher, they may also lead to less help and more sabotage; an issue exacerbated when team members' abilities are varied. Using a lab experiment, we examine how increasing competitive incentives affects performance as team composition changes. We find that competitive incentives generally underperform noncompetitive incentives and a larger bonus does not generate enough effort to compensate for a loss in help. Our results help understand better how to balance out individual versus team rewards and how firms could structure teams when employees have heterogeneous abilities.

Keywords: contest, help and sabotage, team composition, incentive structure
JEL classification codes: C92, D01

*Financial support from the Austrian Science Fund (FWF, P27912-G27 and SFB F63) is gratefully acknowledged. We are grateful for valuable comments from the conference participants at the 2018 ESA World Conference in Berlin, Germany; 2018 GSDS Conference on Decision Sciences, Konstanz, Germany; 2018 THEEM in Kreuzlingen, Switzerland; 2018 ASFEE in Nice, France; 2018 LAWEBESS in Santiago, Chile; 2019 Coller Conference on Behavioral Economics in Tel Aviv, Israel; Texas Experimental association conference at Baylor University, Waco, USA, as well as seminar participants at the University of Alaska, USA, and Ohio University, USA.

[†]Ohio University; corresponding author, dutcherg@ohio.edu
[‡]University of Innsbruck
[§]Florida State University
[¶]Southern Methodist University

1 Introduction

Many tasks are best handled by a team.[1] The success of a team often relies on cooperation among team members such as when one member of the team helps another by sharing knowledge or takes on part of a teammate's task (Lazear and Shaw, 2007). Such cooperation is often fostered by team incentives, with some portion of pay based on collective output.[2] While team-based incentives may promote cooperation, they may also dampen individual effort due to free-riding concerns. The alternative approach to motivating employees, advocated by the likes of Jack Welch, argues that competition among employees and appropriate rank-based rewards are the best way to motivate workers and drive the organization toward constant improvement.[3] Competitive mechanisms, however, may diminish the willingness of individuals to help others, and may even incentivize intra-team sabotage[4], leading to potentially lower total output.[5] The question of whether rank-based reward mechanisms lead to increased or decreased production in a team setting have been

[1] The number of organizations utilizing teamwork has been growing since the 1980s (Lazear and Shaw, 2007). Between 1987 and 1999, the percentage of firms with at least 20% of employees working in teams increased from 37 to 61% (Lawler, Mohrman and Benson, 2001).

[2] Following the notion from Alchian and Demsetz (1972) that team production can be modeled as a public good problem, Dickinson and Isaac (1998) examine how absolute versus relative individual rewards affect contributions to a team when revenue sharing is present and the individuals are heterogeneous. Likewise, Lawler, Mohrman and Benson (2001) report some level of "gainsharing" in 53% of surveyed firms, indicating this is a common practice.

[3] "Jack Welch: 'Rank-and-Yank'? That's Not How It's Done, "The Wall Street Journal, November 14, 2013, https://www.wsj.com/articles/8216rankandyank8217-that8217s-not-how-it8217s-done-1384473281.

[4] Even though sabotage has been more broadly defined (Charness and Levine, 2004; Hollinger and Clark, 1983), in a contest setting, it is an act that reduces a rival's likelihood of winning and may entail actions such as, reducing the output of one's opponent, denying access to resources, withholding of information, mobbing, harassment or physical sabotage (Münster, 2007; Gürtler, Münster and Nieken, 2013). For summary articles, see Chowdhury and Gürtler (2015) and Dechenaux, Kovenock and Sheremeta (2015), chap. 6.1.

[5] Sabotage in contests was introduced theoretically by Shubik (1954), Dye (1984), Nalebuff and Stiglitz (1983), and Lazear (1989). Using a Lazear and Rosen (1981) rank-order tournament, Lazear (1989) highlights sabotage among fellow workers as a key concern for this incentive mechanism. He makes explicit the theoretical tradeoff between efficiency and sabotage, and models "negative sabotage", that is, helping behavior. Empirical research that tests theories on sabotage typically relies on data from field or laboratory experiments (Garicano and Palacios-Huerta, 2005; del Corral, Prieto-Rodriguez and Simmons, 2010; Carpenter, Matthews and Schirm, 2010; Deutscher et al., 2013; Balafoutas, Lindner and Sutter, 2012; Harbring and Irlenbusch, 2008, 2011); by and large, these empirical studies confirm the above theories.

much debated in many large organizations over the past few decades, and while the ranking mechanisms may be on the decline, which approach corporations should adopt has yet to be settled. For example, while Microsoft recently decided to scrap their rank-order system, Yahoo! announced they were implementing one.[6] Given the relevant implications of these compensation issues, a more complete understanding of how they actually affect employee behavior is warranted.

Evaluating these claims and determining the effectiveness of rank-based mechanisms is vital to unlocking team dynamics and ultimately in resolving questions regarding how employers might want to assemble their workers into teams. The dual incentive problem that firms face in trying to incentivize team production—incentivizing individual effort while also trying to encourage cooperation—becomes even more complicated once the realistic assumption of worker heterogeneity is considered. This is because workers of varying ability may respond to these rank-based mechanisms differently and their response may also differ based on the composition of their team. This makes determining the effectiveness of payment schemes more difficult and compounds the problem by introducing a new puzzle regarding the optimal way to compose teams. The issue is no longer one related only to incentives, but now involves a question of whether it is best to construct homogeneous or heterogeneous teams where the answer very likely depends on the nature of the task and the rank-based reward system.

Our goal in this paper is to assess the validity of the competing claims regarding the effect of competitive incentives on individual effort and helping behavior in a team production setting with heterogeneous agents, and to determine what those results suggest about the optimal structure of teams. Exploring these questions using field data would be

[6]"Microsoft Ditches the Stack Ranking System. Yahoo! Lays off 600 because of It," InfoQ, November 16, 2013, https://www.infoq.com/news/2013/11/stack-ranking-microsoft-yahoo; "'Because Marissa Said So' – Yahoo's Bristle at Mayer's QPR Ranking System and 'Silent Layoffs,"' All Things D, November 8, 2013, http://allthingsd.com/20131108/because-marissa-said-so-yahoos-bristle-at-mayers-new-qpr-ranking-system-and-silent-layoffs/; "Microsoft axes its controversial employee-ranking system," The Verge, November 12, 2013, https://www.theverge.com/2013/11/12/5094864/microsoft-kills-stack-ranking-internal-structure.

difficult because such data rarely contains information on effort and abilities, and is almost guaranteed to omit information on behaviors involving help and sabotage. Further, the endogeneity of team construction and design of the compensation mechanism would make it difficult to identify causal relationships. We will, therefore, investigate these questions through the use of controlled laboratory experiments guided by a theoretical model of the underlying incentives.

To highlight the basic questions, it is worth discussing the competing schools of thought in more detail to understand the essential elements of their claims and what evidence might exist for or against these claims. Fundamentally, the two sides of this debate make conflicting claims regarding how workers might react to different incentive schemes in regard to both their individual effort and their tendencies to help fellow co-workers. Those who believe strongly in the importance of competitive incentives are implicitly claiming that such incentives lead to large increases in individual effort, but do not significantly reduce helping behaviors or at least do not lead to such a large reduction that overall productivity is harmed. The group who believes that competitive incentives damage teamwork and impair overall productivity are essentially claiming the opposite which is that while the competitive incentives might or might not increase individual effort, those incentives will ruin the willingness of co-workers to help each other and could even lead to acts of sabotage.[7] This viewpoint will usually come with an optimistic view of the ability of people to cooperate with each other as well.

One of the reasons that neither viewpoint has achieved dominance is that there is substantial prior evidence favoring the core arguments from both. For example, there is a long literature examining behavior in tournaments and contests, and the most common result found in that literature is that competitive incentives drive individuals to exert substantially more effort than predicted in a standard model (see, e.g., a review by

[7] See, e.g., "Companies Revisit 'Rank And Yank' of 1980s," NPR, December 2, 2013, https://www.npr.org/2013/12/02/248151316/companies-revisit-1980s-rank-and-yank, and "Why Stack Ranking Is a Terrible Way To Motivate Employees," Business Insider, November 15, 2013, https://www.businessinsider.com/stack-ranking-employees-is-a-bad-idea-2013-11.

Dechenaux, Kovenock and Sheremeta, 2015). These results lend credence to the use of competitive incentives in the field. On the other hand, there is also a very long literature showing that individuals are much better at cooperating than one would expect given the predictions from a standard model (see, e.g., Ledyard, 1995). This literature showing that individuals will often contribute much more than expected in a public goods setting suggests that in corporate team production settings, teammates may also be able to solve the cooperation problems and, therefore, generate high output without the need for the competitive incentives. These base studies on behavior in contests and public goods environments do not, however, provide clear answers to the question of how the competitive incentives affect cooperative behavior because in most of these prior studies the two issues are examined separately. What remains largely unexplored is if these behavioral deviations will exist when both incentives mechanisms are present and agents vary in their ability.

There are a few prior studies which examine more directly the effect of competition on cooperation. Buser and Dreber (2016) examine the issue in a setting where participants either compete for a prize or engage in piece-rate work prior to playing a public goods game. That study finds that people do tend to cooperate less after they have competed with each other. These findings do not seem conclusive, though, due to the fact that the study finds similar results when the prize is allocated purely randomly rather than through a competition. This suggests the lack of cooperation in the public goods game may simply be due to unequal endowments rather than the experience of competition, as has been found in other studies (Heap, Ramalingam and Stoddard, 2016). Drago and Garvey (1998) suggest that helping effort is reduced when incentives in promotion tournaments are strong. A similar finding is shown by Brown and Heywood (2009), using a survey of finance industry employees. Hamilton, Nickerson and Owan (2003) provide indirect evidence that high-ability subjects were willing to help low ability subjects at a garment factory when the incentives are switched from individual- to team-based. In a somewhat similar study, Johnson and Salmon (2016) examine heterogeneity and sabotage, but sabotage occurs via a post-tournament choice, not simultaneously to taking the effort decision.

To the best of our knowledge, Danilov, Harbring and Irlenbusch (2019) is the only other paper analyzing help and sabotage in a team production setting involving a combination of cooperative and competitive incentives. While the overall theme of their study is similar to ours, the two studies address different sets of research questions that call for significant differences in modeling and experimental design. Danilov, Harbring and Irlenbusch (2019) constructed a model of team production using a variation of a Lazear-Rosen tournament with symmetric agents. Their main interest was to test the model predictions in an experiment, which they carefully implemented with a one-shot incentivized decision preceded by an extensive training phase to ensure subjects understand the environment well. Indeed, Danilov, Harbring and Irlenbusch (2019) presents an experiment well designed to address their questions and find a good agreement between theory and observed behavior. In contrast, our focus is on understanding the behavior of heterogeneous agents and group composition effects. We are also interested in behavioral deviations from standard theory—mainly excessive competitiveness and/or cooperativeness—and in how those deviations develop over time in an organization-like setting with repeated interactions and feedback. To this end, we constructed a model based on a Tullock contest with asymmetric productivities. In the experiment, we use multiple decision rounds with partner matching and allow for help and sabotage to be type-dependent and allocated separately towards individual team members. Thus, our interest – and the contribution of this paper – is in taking the examination of these issues further by investigating teams with various compositions of *heterogeneous* workers and trying to identify systematic deviations from the standard theoretical predictions that will provide a deeper understanding of this behavior.[8]

[8]Literature that theoretically investigates heterogeneous agents in tournaments has found diverging results on the effect of heterogeneity on behavior. Chen (2003) and Münster (2007) investigate heteroge-

To begin to understand how heterogeneity complicates these issues, consider a heterogeneous team composed of one member whose ability is far superior to that of her teammates. Relative to a purely team-based incentive scheme, the introduction of a competitive bonus may not actually induce the better player to exert much higher effort if she expects to easily win the competition anyway. Even though the incentive may not induce higher effort, because she knows she faces very little competition for this prize, she may still help her less productive co-workers to keep team production high and secure higher team-level payments. Additionally, if her less productive teammates expect little chance of winning the bonus, it might also be the case that they too will not increase their effort, but this also implies that their willingness to help others is not diminished. If a manager wishes to increase the competitiveness of the environment, she may reassign workers so that the team includes multiple strong members. With this new configuration, introducing tournament incentives could yield a very different impact. The effort of the high-ability team members could increase substantially as each strives to win the prize, but they may no longer be willing to spend the time to help others as it could improve the competitiveness of others which indirectly decreases their own chance of winning. At the extreme, when all members of a team are of similar abilities, the addition of a competitive prize may actually lead to team members sabotaging each other. Of course the firm does not want to encourage such behavior, but more important to the firm is the total effect. Reductions in help resulting from more competitive settings may be optimal if they are more than offset by increases in effort.

What these few examples make clear is that when teams are comprised of workers of heterogeneous ability, their potential responses to the introduction of a competitive element are quite complicated and are potentially driven by a range of conflicting mo-

neous agents, who compete against each other for an outside reward. They find that high ability workers are more likely to attract sabotage. By contrast, Kräkel (2005*a*) finds that an underdog is more likely to choose help than a favorite, whereas a favorite is more likely to choose sabotage than an underdog. These basic predictions were supported in subsequent experiments (Harbring et al., 2007; Vandegrift and Yavas, 2010; Charness, Masclet and Villeval, 2014); yet, in all these papers, helping behavior is usually ignored or treated as "negative sabotage", which disregards behavioral aspects.

tives. Depending on how different motivations balance out, it may lead to firm managers preferring to try to form relatively homogenous teams, as this may maximize the effect of bonuses on individual effort. The firm may also wish to form heterogeneous teams, as this may better preserve the willingness of teammates to cooperate with each other. It is also possible that the optimal team configuration could depend on the level of competitive incentives.

To investigate these issues, we start by presenting a theoretical model of decision making for workers in a team where output is rewarded by a combination of team-based and competitive incentives. Workers are heterogeneous in their productivity (in the experiment, we restrict heterogeneity to having only two types) and are able to exert individual effort as well as help or sabotage other team members. Our main interest is in understanding if the competitiveness of the setting leads to levels of effort and help/sabotage that differ from money-maximizing behavior and if these deviations support either of the two competing schools of thought on corporate compensation. The theoretical predictions provide a baseline for money-maximizing behavior, while the experiment is designed to isolate the relevant areas of the broader debate and allows causal identification of how incentives and team composition affect behavior. Specifically, in the experiment, we examine how behavior changes as we vary the proportion of high and low ability workers on a team and how behavior changes as we increase the strength of the competitive incentives.

Our results largely support the comparative static predictions from the theory: if incentives are given for cooperative behavior, cooperation increases and individual effort decreases, but when competitive incentives are used, cooperation decreases and individual effort increases. This is in line with findings from Danilov, Harbring and Irlenbusch (2019). However, counter to Danilov, Harbring and Irlenbusch (2019) our point predictions are rarely supported and we identify behavioral deviations from the theory that mostly comport with the noncompetitive school of thought. This has implications for heterogeneous teams and team formation. Specifically, we find that teams with an equal number of agents of each ability level perform weakly better than any other combination of different ability levels. Exploring why this result exists, we find that output is never above theoretical predictions. Instead, competitive incentives lead to larger negative deviations in total output than noncompetitive incentives. In other words, considering the effect of ratcheting up competition as leading to substitution of effort for help, we observe that, contrary to the predictions of the competitive school, this substitution is at most one-to-one, and in many cases it is lower.

Decomposing further the individual effects of effort and help on output, we show that in the pure revenue sharing scheme—in the absence of competitive incentives—subjects cooperate more than predicted by providing effort above the individually optimal levels. This offsets deviations in helping behavior that is a bit lower than predicted. With moderate competitive incentives, average effort is just a little above the theoretical predictions (but not as much as in the noncompetitive scheme) and helping behavior is virtually wiped out. Under strong competitive incentives, individual effort is below predicted, but we also see that sabotage is not quite as high as predicted. This highlights an additional regularity—that helping and sabotaging others is not as sensitive to incentives as is own effort—that is deserving of further attention.

Even though we find the competitive school is generally not supported, our experiment generates a complex set of findings, and potentially explains why both schools still survive: There simply is not a one-size-fits-all recipe for how to balance competitive and noncompetitive incentives. Understanding the strengths and weaknesses of both may help guide where one might or might not consider using each type of incentives in a team production setting.

The rest of the paper is organized as follows. Section 2 presents the theoretical model while Section 3 presents the experimental design and our parameter selection. The results are presented in Section 4, and Section 5 concludes.

2 Model

In this section we present a model that provides a set of predictions regarding how individuals will behave when competitive incentives are introduced into a team production setting. Consistent with our issues of interest, the model allows for heterogeneity in ability of team members and for team members to choose to devote their energy toward individual effort, helping another teammate or sabotaging another teammate. There are many different assumptions one can make in constructing such a model that will affect its predictions regarding effort, help and sabotage levels, and how helping behavior occurs between agents of various types.

Our goal is not to produce a general model which is calibrated on any specific setting. Rather, what we need from the model is a flexible and straightforward method of providing a set of baseline predictions regarding behavior in an environment which is amenable to conducting experiments. The model we present was constructed with this goal in mind, noting that our interest in the end will be mostly in examining the data for systematic patterns regarding how individuals alter their behavior as we increase the relative magnitude of competitive incentives and change the ability composition of the team.

Our model is a variation of several existing models of help and sabotage in teams employing homogeneous (Garvey and Swan, 1992; Danilov, Harbring and Irlenbusch, 2019) and heterogeneous (Kräkel, 2005*b*; Gürtler and Münster, 2013) agents.[9]

Consider a team consisting of $n \geq 2$ risk-neutral agents indexed by $i \in \{1, \ldots, n\}$ and characterized by (possibly heterogeneous) ability parameters $\gamma_i > 0$. Each agent i chooses effort $x_i \in \mathbb{R}_+$ associated with a strictly convex, increasing cost function $c(x_i)$. In addition, agent i chooses, for every agent $j \neq i$ in the team, the level of effort-modifying

[9]The main difference between these model and ours is in that they model the tournament component of the incentives à la Lazear and Rosen (1981) whereas we employ a lottery contest success function of Tullock (1980). Ultimately, both are a form of a noisy winner determination process. One advantage of our model is that it allows for a flexible closed-form solution for heterogeneous agents.

activity $k_{ij} \in \mathbb{R}$, where $k_{ij} > (<)0$ corresponds to agent i helping (sabotaging) agent j. Help and sabotage are associated with a strictly convex cost function $s(k_{ij})$, which is increasing (decreasing) for $k_{ij} > (<)0$. The output of agent i is given by

$$y_i = \gamma_i \max\{0, x_i + \sum_{j \neq i} k_{ji}\}, \tag{1}$$

Equation (1) ensures that output cannot be negative for any levels of sabotage. In the experiment, we choose parameters so that in equilibrium the constraints $x_i \geq 0$ and $y_i \geq 0$ are not binding.

Note that due to the specification of help in our model, a transfer of helping effort k_{ij} from agent i to agent j is augmented by j's productivity parameter γ_j. This implies that help directed to the types with the higher ability parameter is more effective than help directed to types with lower ability parameters, independently on the ability parameter of the type exerting the help. This is certainly one of many possible ways help can operate. Think, for example, of a team of lawyers, in which senior lawyers are responsible for strategic decisions while junior lawyers help them with mundane tasks, such as summarizing precedents. In this case, the junior lawyers helping the senior ones is the most efficient way to increase productivity of the firm.

Team incentives without competition

Every team member receives a piece rate r per unit of total team output $Y = \sum_{i=1}^{n} y_i$. For simplicity, suppose that effort and effort-modifying activities have the same cost, and both cost functions are quadratic: $c(x_i) = \frac{1}{2\alpha} x_i^2$ and $s(k_{ij}) = \frac{1}{2\alpha} k_{ij}^2$.[10] Note that k_{ij} can be positive or negative, but $s(k_{ij})$ is increasing in $|k_{ij}|$. This gives agent i's utility (payoff) in the form

$$\pi_i = rY - \frac{x_i^2}{2\alpha} - \sum_{j \neq i} \frac{k_{ij}^2}{2\alpha}.$$

[10] Parameter $\alpha > 0$ can be subsumed in γ_i and is redundant for modeling, but it will be helpful in calibrating the experiment. We decided to implement the same cost structures for effort and for help for expositional convenience and ease of understanding for subjects in the experiment.

Maximizing π_i with respect to x_i and k_{ij} (for each $j \neq i$), obtain

$$x_i^*(0) = r\alpha\gamma_i, \quad k_{ij}^*(0) = r\alpha\gamma_j.$$

These levels of effort and effort-modifying activities constitute the unique Nash equilibrium (NE) in dominant strategies.

Team incentives with competition for a bonus

We will now introduce an intra-team contest. We assume that there is a manager who imperfectly observes individual output levels y_i and rewards the agent whose output is perceived as the highest with a bonus $V \geq 0$. We model the winner determination process using the Tullock/lottery contest success function (CSF) whereby the probability for agent i's output to be perceived as the highest is $\frac{y_i}{\sum_{j=1}^n y_j}$. In this setting, agent i's expected payoff function is

$$\pi_i = \frac{V y_i}{\sum_{j=1}^n y_j} + rY - \frac{x_i^2}{2\alpha} - \sum_{j \neq i} \frac{k_{ij}^2}{2\alpha}. \tag{2}$$

In order to find the equilibrium, consider the system of first-order conditions for effort and help/sabotage levels, assuming interior solutions:[11]

$$\frac{V\gamma_i \sum_{m \neq i} y_m}{(\sum_{m=1}^n y_m)^2} + r\gamma_i = \frac{x_i}{\alpha}, \quad x_i \geq 0; \tag{3}$$

$$-\frac{V y_i \gamma_j}{(\sum_{m=1}^n y_m)^2} + r\gamma_j = \frac{k_{ij}}{\alpha}, \quad j \neq i. \tag{4}$$

Equations (3) and (4) can be manipulated to obtain a closed-form solution. Expressing

[11] In the experiment, we choose parameters so that the interior solution to first-order conditions indeed provides best responses in each case.

x_i from (3) and k_{ji} from (4), obtain individual outputs,

$$\begin{aligned} y_i &= \gamma_i \left(x_i + \sum_{j \neq i} k_{ji} \right) \\ &= \alpha\gamma_i \left(\frac{V(Y-y_i)}{Y^2} + r\gamma_i - \frac{V(Y-y_i)}{Y^2} + (n-1)r\gamma_i \right) = nr\alpha\gamma_i^2, \end{aligned} \tag{5}$$

and aggregate team output: $Y^* = \alpha rn \sum_{i=1}^n \gamma_i^2$. Plugging this expression and (5) into (3) and (4), obtain

$$x_i^*(V) = \gamma_i \left[r\alpha + \frac{V \sum_{m \neq i} \gamma_m^2}{rn(\sum_{m=1}^n \gamma_m^2)^2} \right], \quad k_{ij}^*(V) = \gamma_j \left[r\alpha - \frac{V\gamma_i^2}{rn(\sum_{m=1}^n \gamma_m^2)^2} \right]. \tag{6}$$

As seen from (6), while equilibrium effort increases with the introduction of the bonus, help decreases so as to exactly offset the impact of the increase in effort on aggregate output; the latter is independent of the bonus. For a sufficiently large V, help becomes negative, i.e., it turns into sabotage. Note that, other things being equal, this turning point is lower for higher-ability agents.

Note that the equilibrium predictions are not the social optimum. Instead, the choices maximizing welfare are above the equilibrium predictions, since contributions yield a payoff to all four group members, while the cost is only incurred by the contributor. This is true for both, effort and help.

3 Experimental Design and Predictions

3.1 Overview

Sessions were conducted in November 2017 and May 2018 at the Innsbruck Econ Lab. The experiment was computerized via z-Tree (Fischbacher, 2007), and recruitment of subjects took place via the recruitment system *hroot* (Bock, Baetge and Nicklisch, 2014). Once all subjects were checked in and seated at a computerized workstation, instructions were handed out and were read out loud.[12] After all questions were answered, the experiment began.

3.2 Treatments

In the experiment, we utilized two ability levels; high (H) and low (L). The ability of each subject was exogenously assigned at the beginning, and fixed throughout the session. Additionally, subjects were anonymously assigned into a fixed team with three others to make a team of four. In order to fully understand the effects of team composition, we utilized a between subjects design, where subjects were assigned into one of five potential team composition treatments—*HHHH*, *HHHL*, *HHLL*, *HLLL* or *LLLL*. In a given session, the group composition was fixed and each subject knew their own ability along with the ability of their three teammates.

The main part of the experiment consisted of three 8-round blocks for a total of 24 rounds. The first block included only a team incentive: all team members received the same payoff, proportional to the total team output; there was no contest incentive in this block. The second and third blocks added on the contest incentives, and we varied the size of the prize across these blocks, with one block using a relatively small prize, a low powered incentive, and the other a rather large prize, to represent a high powered incentive. The reason for using within-subject variation in the contest prize—and beginning all sessions with the no-contest baseline—is that we wish to understand how behavior changes with the *introduction* of contest incentives. Starting from a situation without contest incentives and then adding them in represents the change we are considering when a corporation implements a different system. To control for order effects in examining the different sizes of contest prizes, we varied the order of the prizes. In *order 1*, the low powered incentive was introduced in block 2 and the high powered in block 3; in *order 2*, these were reversed

[12] All instructions were neutrally framed. Instructions for one of the treatments are included in Appendix B.

with the high powered incentive in block 2 and the low powered incentive in block 3.[13] New instructions were handed out, and read out loud prior to each block, to introduce and explain the changes to the incentive scheme. It was common knowledge that one round per block was chosen randomly for payment.

In each round, all subjects made four simultaneous choices. They had to choose how many points to allocate to their own effort and how many to allocate to modifying their three other teammates' effort. Individual effort could be any integer from 0 to 150, while modifications ranged from -150 to 150.[14] Each choice entailed a cost, which was presented to subjects in a table in their instruction packet.[15] After all subjects made their choices, they were shown a results screen. On the results screen, subjects were reminded of their own choices and were shown the average help/sabotage in their team directed at all members, their own total effort (which is a combination of their own effort and effort modifying choices of their team members), their output after accounting for their ability, and the total team output. They were also reminded of the cost of each decision, the group payment from the team output, whether they won the prize or not (in blocks 2 and 3), and their total payoff in that round should it be chosen for payment. Following the main game, subjects' risk preferences were elicited using the "bomb" risk elicitation task (Crosetto and Filippin, 2013).

3.3 Parameters and Equilibrium Predictions

The goal of our experiments is to examine how behavior changes as we increase tournament incentives and change group composition in a team production environment. We have constructed a set of parameters which are intended to allow us to do just that. First, we

[13]The experiment was designed to obtain two sessions of each team composition with those two sessions using the different orderings. In the end, we conducted 11 sessions instead of 10 initially intended because one session configuration was accidentally used in two sessions.

[14]All amounts in the main part of the experiment were denominated in tokens. At the end of the session payoffs were translated into Euros at the exchange rate 100 tokens = €6.

[15]In order to ensure numeracy was not a concern, subjects had access to an on-screen calculator which calculated hypothetical payoffs given their own choices and hypothetical choices they entered for other members in their group. Sample screenshots are included in Appendix C.

Table 1: Equilibrium predictions

	$LLLL$	$HLLL$	$HHLL$	$HHHL$	$HHHH$
			$\boldsymbol{V} = 0$		
x_L	9.77	9.77	9.77	9.77	
x_H		19.53	19.53	19.53	19.53
k_{LL}	9.77	9.77	9.77		
k_{LH}		19.53	19.53	19.53	
k_{HL}		9.77	9.77	9.77	
k_{HH}			19.53	19.53	19.53
			$\boldsymbol{V} = 100$		
x_L	28.52	22.01	18.77	16.87	
x_H		31.78	31.53	30.18	28.91
k_{LL}	3.52	7.72	8.77		
k_{LH}		15.45	17.53	18.35	
k_{HL}		1.60	5.77	7.40	
k_{HH}			11.53	14.80	16.41
			$\boldsymbol{V} = 500$		
x_L	103.52	70.99	54.77	45.27	
x_H		80.76	79.53	72.79	66.41
k_{LL}	-21.48	-0.44	4.77		
k_{LH}		-0.88	9.53	13.61	
k_{HL}		-31.05	-10.23	-2.07	
k_{HH}			-20.47	-4.14	3.91
y_L	39.06	39.06	39.06	39.06	
y_H		156.25	156.25	156.25	156.25
Y	156.25	273.44	390.63	507.81	625.00

vary the size of the prize V between the values 0, 100 and 500 to represent no competitive incentives, moderate and then high incentives. Second, we vary the team composition by considering all possible configurations of L and H types, in four-person teams, as previously described. We set the ability of our workers to $\gamma_L = 1$ for low ability and $\gamma_H = 2$ for high ability workers, the piece rate to $r = 0.25$ and $\alpha = 39.0625$. This results in the cost functions $c(x_i) = 0.0128x_i^2$ and $s(k_{ij}) = 0.0128k_{ij}^2$. With these parameters, equations (5) and (6) lead to a set of equilibrium point predictions of behavior as displayed in Table 1.

There are a few key elements to note about these predictions. First, total group output, Y, increases with the number of H types in the group, yet, given a fixed group composition, it does not vary with V. Importantly, individual effort, x_i, does depend on V, as does help/sabotage, k_{ij}. These values move in opposite directions, and exactly offset each other in determining total output under our production function. This symmetric offset makes it easy to examine the exact channel through which V affects output and allows us to cleanly test the two schools of thought. Relatedly, these parameters also allow us to cleanly examine behavioral effects from different group compositions. That is, with a group of eight workers—four H types and four L types—divided into two groups of any configuration, the total output is predicted to be constant, 781.25, for any value of V (or 390.63 for each group, on average). That is, in equilibrium, all group assignments of the eight workers are equivalent to the firm. We may, however, find that they are not behaviorally equivalent.

A second element to note is that helping behavior is predicted to be directed mostly towards H types rather than L types, as both H and L types help H types more than L types in equilibrium.[16] This is due to the specification of help in our model. One should not mistake the prediction our model makes for a general claim—we take no stance on

[16]Prior studies examining helping behavior have found the opposite when output was determined by a minimum effort production function (Brandts et al., 2016) or when nonmonetary benefits of help are considered (Hamilton, Nickerson and Owan, 2003). In these studies, high ability workers tend to help the low ability ones.

what the "correct" relationship is; we rather want to understand how observed behavior changes as V and team composition change, relative to some predicted level of cooperative behavior and for that, all that is needed is to understand how helping behavior in general may change. Our environment was constructed to induce levels of help that were high enough to allow us to observe how they might change. Their relative sizes between types are of no importance to our research questions.

Finally, note that for $V > 0$, as the number of H types increases, the effort of each type is predicted to decrease, while help increases and sabotage is declining. This is because team production hinges more on the H types output and thus helping an H type leads to higher levels of team-based revenue while an increase in the number of H types also decreases the return to providing individual effort to win the contest. Additionally, when help is predicted to be positive, more help is predicted to be directed towards H types, than towards L types. However, when sabotage is predicted ($V = 500$), L types are predicted to receive less sabotage than H types.

Based on the values in Table 1—or equations (5) and (6)—we can construct a basic set of testable predictions. Even though such tests could be useful in generating a first estimate of behavior in this setting, our interest is not in simply demonstrating whether the model works, but rather in using these theoretical benchmarks to test for systematic behavioral *deviations* from the theory. We can then check if these deviations support one of the competing schools of thought where the competitive school would predict excess effort above the theoretical prediction when competitive incentives are used and reductions in help in these settings would not be so severe such that total output would be higher. The non-competitive school would predict the opposite. Because our main tests look at deviations from theory, any null results would imply conformance with theoretical predictions and lack of evidence in favor of either school.

4 Results

4.1 Data and Analysis Overview

We begin with an overview of the data. Table 7 in Appendix A presents the equilibrium predictions and observed averages in each treatment and reward scheme. While these summary statistics are useful for gaining an initial impression of the results, we will not conduct an exhaustive sequence of tests comparing the observed averages to theoretical predictions. As discussed above, our approach is instead to look for patterns in *deviations* from equilibrium predictions resulting from the introduction of competition and changes in the competitive environment, including group composition.[17]

The key pattern to our data which will underlie all of our results is that at no ($V = 0$) or moderate competitive incentives ($V = 100$), effort is greater than predicted by the model, while at high competitive incentives ($V = 500$), it is less than predicted. For helping behavior we find that it is never above the predictions, but at moderate incentives helping behavior is well below predicted. At high incentives though, helping behavior does not go as low as predicted, mostly because we observe much less sabotage than predicted. Effort and helping behavior lead to a tradeoff in total output such that total output is highest in the no competition case and lowest at intermediate competitive incentives. The composition of the group impacts these effects: As the number of H types goes up, effort increases but help decreases. These effects balance out to less of a hit on overall output when the competition of the setting varies, but it does not lead to a better outcome when competitive incentives are in place. These results suggest challenges for both schools of thought regarding the effect of competitive incentives but, most notably, the competitive school finds little support.

While the relationships demonstrated in the summary statistics are suggestive of

[17]For interested readers, Table 8 in Appendix A presents the observed averages with robust standard errors that can be used to perform comparisons to theory. Also included in the Appendix are Figures 5-8 which provide a graphical summary of the data.

our ultimate results, in the next several sections we present a more detailed regression analysis to deal with the interdependencies in the data and to allow us to explicitly study the deviations between the observed data and the theoretical predictions. We begin our analysis by looking into how team composition and the size of the prize affect total output. The overview of team composition leads naturally to questions about how the competitiveness of the setting (i.e., increases in V and increases in the number of H types) affects output. Following this, we present a deeper discussion of the explicit and implicit assumptions of the two schools of thought about these components. This leads to testing how competitiveness affects the individual components of output—effort and help—that concludes our analysis section.

4.2 Organization of Employees into Teams

If a manager has a fixed set of employees of various abilities, she can organize them into many different configurations. Using our experimental setting, the hypothetical exercise we consider is how a manager should allocate individuals to teams when she needs two teams of four members each and she has eight employees, four high and four low ability, to allocate between both teams. This implies that the manager can organize her employees into two teams that are Homogeneous (*HHHH* and *LLLL*), Balanced (*HHLL* and *HHLL*), or Asymmetric (*HHHL* and *HLLL*).[18] As stated in Section 3.3, our model predicts that total output of each team composition is constant for any group configuration and for any value of V which makes our analysis on output by group composition informative regarding behavioral deviations.

Figure 1 shows the predicted and observed outcomes for each team composition.[19]

[18] If the assumption that a manager has equal H and L types is relaxed, the only aspect that changes is how many team composition options a manager has. Importantly, if eight employees are present, a manager must have at least two of each type to have a choice on group composition. For instance, if a manager has two L types and six H types, the manager can either form one team of *HHHH* and one team of *HHLL* or two teams of *HHHL*. The empirical results based on the assumption of an equal number of each type are already informative on general behavior and thus we leave out these additional analysis for succinctness.

[19] For a more complete overview, Table 9 in Appendix A presents the averages for the differences between

From the Figure, there appears to be a difference between the Balanced team and the others. To test this, Table 2 displays the results of random effects GLS regressions. The dependent variable is total group output. The independent variables are binary variables equal to one if the group is classified as Homogeneous, and equal to one if the group is classified as Asymmetric, which implies the reference group is the Balanced group. The first column displays results for $V = 0$, the second for $V = 100$, and the third for $V = 500$.[20] There are 66 groups, and output was observed 8 times per group for each size of the prize, which leads to a total of 528 observations.

Table 2: Group output for Homogeneous, Balanced heterogeneous or Asymmetric heterogeneous team composition

	(1) V=0	(2) V=100 Dep. var.: Y	(3) V=500
Homogeneous	−67.23	−49.93	−31.58
	(67.51)	(57.11)	(65.36)
Asymmetric	−74.88	−37.60	-89.71^{*}
	(56.62)	(53.65)	(51.10)
Constant	468.48^{***}	363.53^{***}	414.29^{***}
	(50.49)	(42.25)	(41.23)
Observations	528	528	528
Number of groups	66	66	66

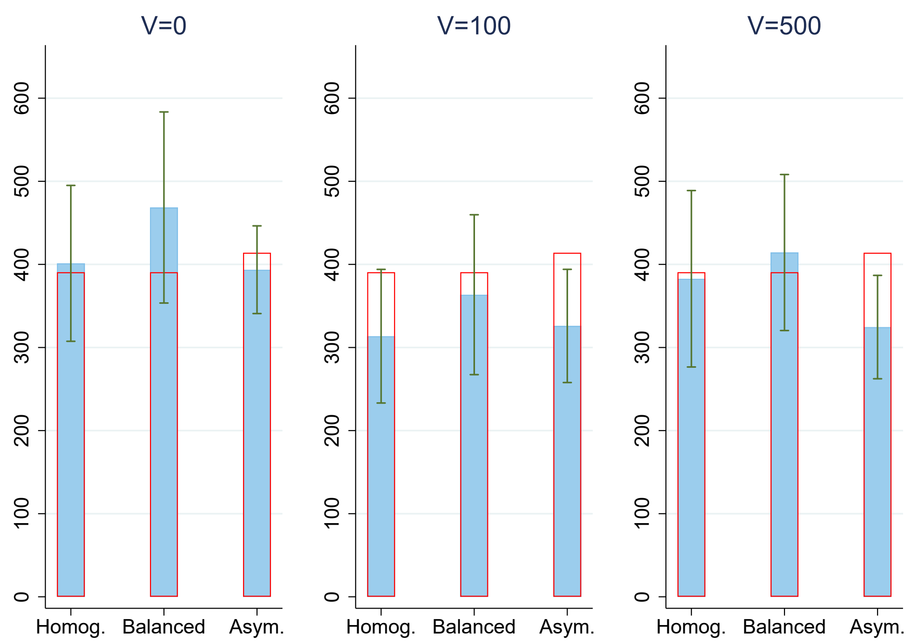
Robust standard errors clustered on group level in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. , The Balanced teams (*HHLL* and *HHLL*) are the baseline; "Homogeneous" is a dummy which equals 1 if the groups are *HHHH* or *LLLL*, and zero otherwise; "Asymmetric" is a dummy which equals 1 if the groups are *HHHL* or *HLLL* and zero otherwise.

The statistically insignificant estimate of the constant term implies that for the Balanced team, the observed total output does not differ from the predicted output for any

the equilibrium predictions and the observed averages by each type in each treatment and reward scheme (Tabe 7 shows the values for the equilibrium predictions and the observed averages). Table 8 presents even more dis-aggregated results, in that it shows the observed averages with robust standard errors that can be used to perform comparisons to theory. Also included in the Appendix are Figures 5-8 which provide a graphical summary of the data.

[20]For robustness, Tables 10 and 11 in Appendix A control for order effects and learning.

Figure 1: Average group output for Balanced, Asymmetric and Homogeneous team composition



Predicted levels are shown as empty boxes and average observed levels as filled boxes, with the error bars representing the 95% confidence intervals. We treat each group in one block as one independent observation, i.e., we have one observation per block in each group.
While also for the Asymmetric team composition, the predicted output is 390.63 per group on average, the displayed prediction bar is slightly higher in this graph to take into account that we have run by accident one additional session of *HHHL*.

value of V. The statistically insignificant estimate on Homogeneous implies it is no different than the Balanced team composition. However post-estimation Wald tests indicate when $V = 100$, output is less than the predicted level ($p < 0.01$). It is not significantly different from the predicted level for $V = 0$ ($p = 0.52$) or $V = 500$ ($p = 0.80$). For the Asymmetric team composition, deviations from output are significantly lower than in the Balanced group composition in the $V = 0$ and $V = 500$ cases. Post-estimation Wald tests also indicate that the observed output for the Asymmetric group is not significantly different from the predicted level for $V = 0$ ($p = 0.31$); however, output is lower than predicted for $V = 100$ ($p < 0.01$) and $V = 500$ ($p < 0.01$). This leads to the following result.

Result 1 (Group composition)
For $V = 0$, team composition does not lead to total output differences relative to equilibrium predictions; for $V > 0$, the Balanced composition fares weakly better than Homogeneous or Asymmetric team compositions.

This result may imply that a Balanced team composition leads to better overall in-group interactions in terms of allowing incentives to encourage reasonable effort while not impairing helping behavior as much as the other configurations.[21] That is, it leads to a (weakly) better overall output than unbalanced teams where all high ability types are in one group and all low ability types in another group, or where there are asymmetric numbers of each type in the group.

4.3 Competitiveness and Output

While the first result highlights that group organization can impact total output, this analysis also indicates that total output does not increasingly exceed equilibrium predic-

[21]It may also be of interest whether some team compositions produce more volatility in output than others. We examined the variance of aggregate output across treatments and found that output variance is somewhat lower in Balanced team composition as compared to Homogeneous and Asymmetric team composition, but the differences are not statistically significant.

tions for any group composition when moving from $V = 0$ to $V = 100$, or when moving from $V = 100$ to $V = 500$. If output is different from equilibrium, it is *lower*, not higher, for larger values of V. This provides our first insight against the competitive school of thought.

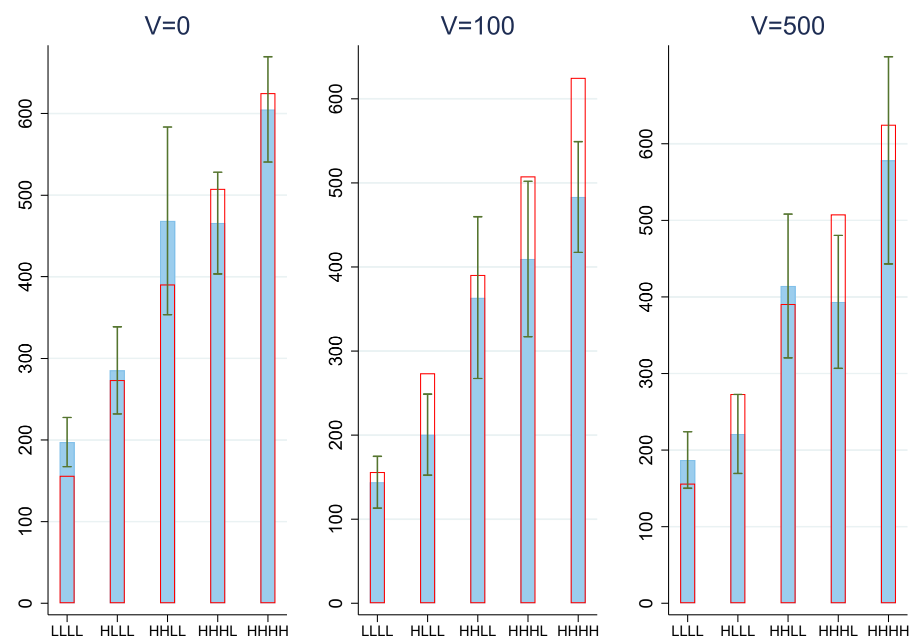
When looking into how competitive a setting is, we will focus on the size of V and the number of H types and assume that the competitiveness of the setting is increasing in the size of the prize and the number of H types in the group.[22] The initial result of output never being above equilibrium already serves as one strike against the competitive school. More generally, in line with the competitive school of thought, if a competitive scheme is beneficial, total output should be (weakly) below equilibrium when $V = 0$ (the setting devoid of competition), increasing above equilibrium when moving from $V = 0$ to $V = 100$ and then increasing further when moving to $V = 500$. Deviations from predictions should also be increasing in the number of H types in the group. To get a better sense of how the latter dimension of competitiveness affects output, Figure 2 compares the predicted to the average group output by group composition and size of the prize.

To examine behavioral effects stemming from increases in the competitiveness of the setting due to increases in V and the number of H types, column (1) of Table 3 displays the results of a random effects GLS regression, where the dependent variable is the difference in actual versus predicted total group output in every round.[23] We include binary variables for $V = 100$ and $V = 500$ and a variable accounting for the number of H types in the

[22]Note that our assumption that the competitiveness of the setting is increasing in the number of H types is premised on the often stated intuition that better opponents bring out the best among the competitors. An alternative way to frame this could be that competitiveness is an increasing function of the number of *own* types. For the H types, nothing changes with this assumption. For the L types, the assumption would be reversed. Similarly, an argument could be made, from our theoretical predictions, that as the number of H types increases, so does the incentive to help, which is met with a decreased incentive to provide effort. This may imply that the competitiveness works in the opposite direction in that competitiveness of the setting is decreasing in the number of H types. Either way of stating the influence of types on outcomes can be easily tested, but the potential for types to influence outcomes remains a key issue we are interested in.

[23]When running the regression in column (1) without controlling for the number of H types, we see that overall group output for $V = 0$ is not different than predicted. Yet, total output is lower than predicted when $V = 100$ ($p < 0.01$) and when $V = 500$ ($p = 0.041$), which do not support the competitive school of thought. Tables 12 and 13 in Appendix A control for order effects and learning.

Figure 2: Average group output



Predicted levels are shown as empty boxes and average observed levels as filled boxes, with the error bars representing the 95% confidence intervals. We treat each group in one block as one independent observation, i.e., we have one observation per block in each group.

team, running from 0 to 4. There are 66 groups, and output was observed 24 times per group, which leads to a total of 1,584 observations.

Table 3: Deviations from group output

	(1) Dep. var.: $Y - Y^{*}$
$V = 100$	-81.71^{***}
	(13.39)
$V = 500$	-47.96^{***}
	(17.85)
number H types	-24.01^{***}
	(7.86)
Constant	58.93^{***}
	(18.14)
Observations	1,584
Number of groups	66

Robust standard errors clustered on group level in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. 'Number H types' runs from 0 to 4.

From the regressions we see negative and significant estimates of the coefficients on $V = 100$ and $V = 500$ implying that deviations from predictions in both of these settings are below the deviations when $V = 0$, even after controlling for the number of H types in the group. A post-estimation Wald test confirms that the deviation is smaller in magnitude when $V = 500$ as compared to when $V = 100$ ($p = 0.018$). This confirms what is observed in the Figure and is counter to what is argued by supporters of the competitive school of thought.

We also find significant and negative estimate of the coefficient on the number of H types that implies as the competitiveness of the setting increases due to changes in the composition of the group, total deviations in output decrease. This is, once again, against the competitive school of thought.

Result 2 (Competitiveness and total output)

(A) Size of V: The difference between observed and predicted total output is decreasing when going from $V = 0$ to $V = 100$ and from $V = 0$ to $V = 500$, but is increasing when going from $V = 100$ to $V = 500$.

(B) Number of H types: The difference between observed total output and the equilibrium prediction is decreasing in the number of H types in the group.

This result shows that, given our production function, increases in the competitiveness of the setting does not translate into higher overall output.

4.4 Competitiveness: Effort and Help

Taken as a whole, the results thus far provide very little support for the competitive school of thought. Yet, although the analysis in the prior two sections provides some overview of how behavior may differ with different levels of competitiveness, it does not directly address the explicit and implied behavioral assumptions of the two schools of thought. The explicit argument of those advocating for the competitive school is that a competitive environment encourages individual effort. The experimental contest literature certainly supports this claim,[24] and shows that it usually is above what may be expected. Although not stated explicitly, the implied assumption is that these increases in effort are not offset by a larger decline in help. The noncompetitive school focuses on helping behavior and explicitly claims that a cooperative environment encourages help,[25] while competitive incentives discourage it and may even encourage sabotage. The implied assumption is that a decline in helping behavior due to more competitiveness is not met with even greater increase in individual effort. Even though these competing claims are not stated explicitly, the underlying notion behind both schools of thought is the degree of substitutability

[24]Overbidding is widely documented in experiments on lottery contests, see Sheremeta (2013) for a review. At the same time, bidding at or below equilibrium predictions is typical for experiments utilizing the Lazear-Rosen tournament framework (see, e.g., Bull, Schotter and Weigelt, 1987; Dutcher et al., 2015). However, the latter environment is not directly comparable to our setting.

[25]See, e.g., the review by Ledyard (1995).

between effort and help, as the competitiveness changes. While this substitutability is already on display in the rational-agent model that predicts that increases in individual effort are met with decreases in help,[26] the two schools claim *additional* behavioral effects exist.

To investigate these behavioral effects, we examine how an individual changes their choices as the competitiveness (i.e., increases in V and increases in the number of H types) changes, by comparing behavior against the theoretical prediction, for the two components of output—effort and help. This subsequent analysis can be especially useful when production functions differ from what we have specified, to the extent the behavioral responses of effort and help generalize to a broader set of production functions. These behavioral effects of competition on effort and help/sabotage are at the heart of both schools of thought.

4.4.1 Effort

To get a sense for what effort looks like for each type in each group configuration as the value of the prize changes, panels (a) and (b) in Figure 3 present average effort for the H and L types respectively, and show how this average effort compares to the equilibrium prediction. Deviations from equilibrium are common and deserve further empirical analysis.

Table 4 reports the results of random effects GLS regressions with standard errors clustered at the group level. Column (1) report the results pooling L and the H types. In order to understand if results differ by type, columns (2) and (3) report the results for the L and the H types respectively.[27] The dependent variable is the difference between an individual's observed effort choice and the theoretically predicted effort in a given period. The expectation from the competitive school of thought is that as the competitiveness

[26]The degree of substitutability is also shown to depend on type.

[27]See Tables 14, 15, and 16 in Appendix A for robustness checks controlling for order effects, risk aversion and learning, respectively.

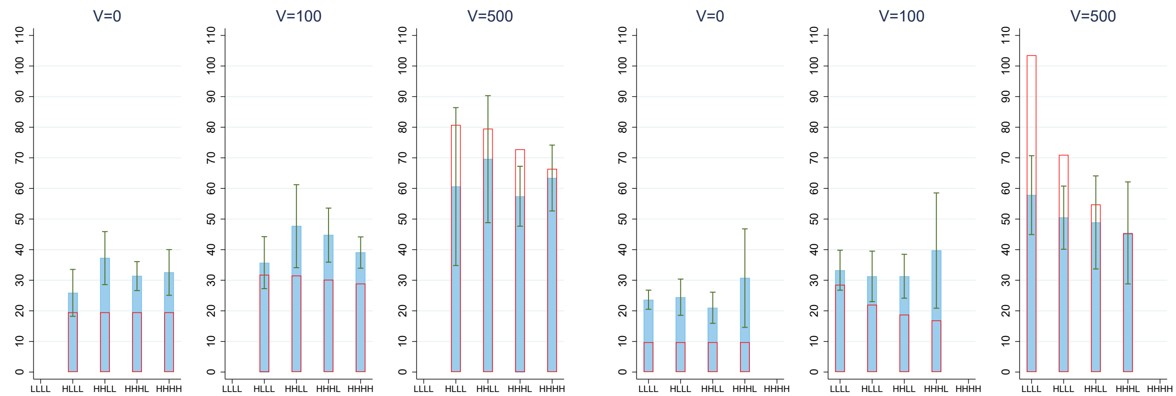
Table 4: Deviation of effort from equilibrium predictions

	(1)	(2)	(3)
	H & L	L	H
	Dep. variable: $x_{i,V} - x^*_{i,V}$		
$V = 100$	-6.79 ***	-9.04 ***	0.67
	(2.04)	(2.25)	(4.72)
V = 500	-50.24 ***	-55.63 ***	-36.59 ***
	(4.25)	(4.33)	(10.98)
Number H types	0.00	1.27	0.40
	(0.77)	(1.92)	(1.72)
$V = 100$ x Number H types	2.11**	4.14***	-0.36
	(0.95)	(1.31)	(1.65)
$V = 500$ x Number H types	9.30***	15.19***	4.42
	(1.59)	(2.51)	(3.39)
Constant	13.65***	13.22***	11.59**
	(1.48)	(1.63)	(5.03)
Observations	6,336	3,024	3,312
Number of indiv.	264	126	138
Number of groups	66	54	54

Robust standard errors clustered on group level in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. 'Number H types' runs from 0 to 4.

Figure 3: Average effort

(a) H types (b) L types



Predicted levels are shown as empty boxes and actual observed levels as filled boxes, with the standard error bars of the 95% confidence interval; we treat each group in one block as one independent observation, i.e. we have per group 1 observation per block—if groups are of heterogeneous types, we have the same group once in the graphs for H-types and once in the graphs for the L-types.

of the setting increases (by increases in V or the number of H types), so will positive deviations in effort. To explicitly test this, our first two explanatory variables are binary variables for $V = 100$ and $V = 500$, which implies the reference group is when $V = 0$. The third explanatory variable accounts for the number of H types in the group (ranging from 0 to 4).[28] Panels (a) and (b) in Figure 3 show that the competitive pressure of the prize does not appear to have the same impact on all group compositions, hence we also control for interaction effects between the size of the prize and the number of H types.

From column (1), we see that deviations in effort are indeed affected by the size of the prize and the number of H types. However columns (2) and (3) indicate the importance

[28] We have run additional regressions to test whether effort is higher than predicted for all values of V. We find that observed effort is higher than the predicted value when $V = 0$ for both types. In the $V = 0$ case, effort above the level predicted suggests individuals may be resolving some of the collective action problems as effort above the self-interested level could be consistent with individuals working towards the good of their group. Effort is also above equilibrium for both types when $V = 100$ ($p < 0.01$ for both types), which could either be due to individuals valuing the good of their group or from competitive desire to obtain the prize. Effort is below equilibrium at $V = 500$ ($p < 0.01$ for both types) suggesting that at this level of incentives, subjects are unwilling to contribute effort even up to the point of individual self-interest. That is, for both types, observed effort is significantly higher than predicted when $V = 0$ and $V = 100$, while it is significantly lower than predicted for $V = 500$.

of the type in this setting, given that the deviations are largest for the L types. From column (2), the results indicate that moving from $V = 0$ to $V = 100$ and from $V = 0$ to $V = 500$ leads to a reduction in (positive) deviations of effort for the L types. Likewise, when moving from $V = 100$ to $V = 500$, post-estimation Wald tests indicate lower levels of deviations ($p < 0.01$). The same result holds for H types when moving from $V = 0$ to $V = 500$ and from $V = 100$ to $V = 500$, but not when going from $V = 0$ to $V = 100$. These findings once again contradict the competitive school.

However, when looking at the interaction effects in column (2), it is seen that the composition of the group also impacts this result as the number of H types attenuates this effect. The positive effect is much stronger when $V = 500$, though this attenuation is only present for the L types. This leads to our next result:

Result 3 (Competitiveness and effort)

(A) Size of V: For both types, deviations of observed effort from the equilibrium prediction are weakly decreasing in V.

(B) Number of H types: Increasing the number of H *types in the group weakly increases deviations from the equilibrium prediction; it is most evident for the* L *types and is strongest when* $V = 500$.

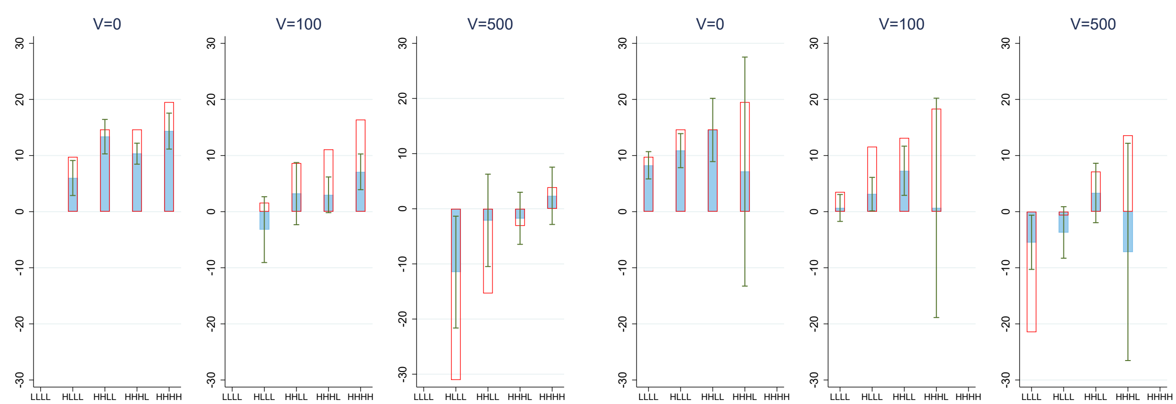
As before, the overall conclusion from our examination of the effort decisions is that we do not find much support for a behavioral effect claimed by the competitive school of thought, which predicts that increasing competitive pressure either through increasing the size of the prize or through increasing the number of H types drives up chosen effort to some hypercompetitive level above the equilibrium prediction. The one exception is the L types' response to an increase in the number of H types, as they appear to be increasing their effort relative to the equilibrium prediction as the number of H types increases. Examining panel (b) in Figure 3 makes it clear that this finding is due to predicted effort declining, while observed effort remains about the same. This effect is particularly pronounced in the $V = 500$ case. This suggests that the L types may simply be non-responsive to the configuration of their group, as they supply on average the same amount of effort regardless of the group configuration.

4.4.2 Help and Sabotage

Having established how competitive pressure alters effort provision, we now analyze how these same pressures affect help and sabotage behavior. Figure 4 presents an overview of the average help for each type, each group configuration and each value of V. Again, we show how these averages compare to equilibrium. It is immediately clear that deviations from equilibrium are universal, but are dependent upon the predicted sign of help. When help is predicted to be positive, the actual help is weakly lower. When help is predicted to be negative (sabotage), the actual help is weakly higher. We will come back to this difference between help and sabotage behavior, but first we will establish results concerning the competitiveness of the setting and deviations in help.

Figure 4: Av. help

(a) H types (b) L types



Predicted levels are shown as empty boxes and actual observed levels as filled boxes, with the standard error bars of the 95% confidence interval; we treat each group in one block as one independent observation, i.e. we have per group 1 observation per block—if groups are of heterogeneous types, we have the same group once in the graphs for H-types and once in the graphs for the L-types.

Table 5: Deviations of help and sabotage from equilibrium predictions

	(1) $H\&L$	(2) L	(3) H
	Dep. var.: $\overline{k_{ij,V} - k^*_{ij,V}}$		
$V = 100$	-2.64^{**}	-2.13^{*}	-2.24
	(1.10)	(1.24)	(2.85)
$V = 500$	12.67^{***}	14.38^{***}	24.62^{***}
	(1.88)	(1.99)	(4.85)
Number H types	-1.36^{***}	-2.63	-1.38^{*}
	(0.51)	(2.36)	(0.72)
$V = 100$ x number H types	-0.52	-1.50^{**}	-0.48
	(0.44)	(0.72)	(0.86)
$V = 500$ x number H types	-2.94^{***}	-8.88^{***}	-5.46^{***}
	(0.74)	(1.24)	(1.49)
Constant	-1.03	-0.41	-0.24
	(0.94)	(1.50)	(2.08)
Observations	6,336	3,024	3,312
Number of indiv.	264	126	138
Number of groups	66	54	54

Robust standard errors clustered on group level in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.
'Number H types' runs from 0 to 4.

Table 5 displays the results of random effects GLS regressions.[29] The dependent variable is the difference between average actual and predicted help, where a negative deviation implies less help than predicted (or more sabotage than predicted). For ease of comparison, the control variables are the same as those used for analyzing deviations in effort.[30]

The results indicate that when the prize increases from $V = 0$ to $V = 100$, deviations in help decrease with the strongest decline coming from L types. However, when the prize increases from $V = 0$ to $V = 500$, deviations in help increase for both types. Likewise, post-estimation tests indicate that when the prize increases from $V = 100$ to $V = 500$, deviations increase ($p < 0.1$ for both types). As noted previously, this result is driven by a decrease in help when $V = 100$, *and* less sabotage in the case of $V = 500$. That is, a higher prize leads to lower levels of help on average when help is predicted to be positive, but when the prediction is to sabotage, individuals refrain from sabotaging more than predicted. This is especially relevant in the $V = 500$ case for the H types where the highest level of sabotage is predicted, cf. Figure 4 (a).

As with effort, the interaction effects also indicate that group composition is impactful. The interaction effects indicate that this positive effect from $V = 500$ is reduced as the number of H types increases. The negative deviations in help when $V = 100$ is even further decreased as H types increase, but this holds only for L types. This leads to our final result.

Result 4 (Competitiveness and help)

[29] See Tables 17, 18, and 19 in Appendix A for robustness checks controlling for order effects, risk aversion and learning, respectively.

[30] As with effort, the magnitudes of deviations for each V serve as useful baselines, but do not address responses to the competitiveness of the setting. When not controlling for the number of H types in the group, for $V = 0$, deviations in help are negative, not positive, for both types, which may indicate that any sort of cooperative preferences are manifested through effort allocations. Deviations in help are also negative when $V = 100$ ($p < 0.01$ for both types), but are not different from zero when $V = 500$ for the L types ($p = 0.587$), and are greater than zero for the H types ($p = 0.017$). That is, observed help from H and L types is lower than predicted when $V = 0$ and $V = 100$; for $V = 500$, it is equal to the predicted value for the L types and greater than the predicted value for the H types.

(A) Size of V: For both types, the difference between observed and predicted help is decreasing in V when going from $V = 0$ to $V = 100$, while it is increasing in V when going from $V = 100$ to $V = 500$ and when going from $V = 0$ to $V = 500$.

(B) Number of H types: As the number of H types increase, deviations in help decline when $V = 500$ for both types, and also decline when $V = 100$ for the L types.

Table 6: Controlling in particular for sabotage

	(1)
	Dep. var.: $\overline{k_{ij,V} - k^*_{ij,V}}$
$V = 100$	-2.64^{**}
	(1.10)
$V = 500$	-11.02^{***}
	(4.04)
Number H types	-1.36^{***}
	(0.51)
$V = 100$ x number H types	-0.52
	(0.44)
$V = 500$ x number H types	3.27^{**}
	(1.32)
Sabotage	24.61^{***}
	(4.31)
Sabotage x number H types	-5.65^{***}
	(1.52)
Constant	-1.03
	(0.94)
Observations	19,008
Number of indiv.	264
Number of groups	66

Robust standard errors clustered on group level in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.
'Number H types' runs from 0 to 4.

Overall, the results mostly support the idea that competitiveness leads to a behavioral "anti-help" mindset. However, we also observe the opposite effect—aversion to sabotage—when sabotage is predicted. High levels of sabotage are predicted, on average, for the H

types when $V = 500$, and in the *HLLL* setting for the H types and the *LLLL* setting for the L types. In all of these cases individuals are largely refraining from sabotaging each other, cf. Figure 4. Thus, even though competitiveness may lead to lower levels of help, this aversion implies that the main drawback from competitiveness, i.e., sabotage, may be more limited than predicted, thus blunting the worst of the impacts predicted by the noncompetitive school of thought.

The implication is that the aversion to sabotage is a level effect that needs to be controlled for. The question the above analysis cannot answer is how the competitiveness of the setting influences sabotage behavior after controlling for this level effect. To investigate this in more detail, we have re-run the above regression for help, controlling for instances when help is predicted to turn into sabotage (see Table 6); the dummy variable 'Sabotage' is a binary variable equal to one if predicted help is negative. Furthermore, to understand how the competitiveness of the setting affects sabotage behavior, we include an interaction effect with the number of H types in a group.

As can be seen by the positive effect from the *Sabotage* variable, individuals do not sabotage as much as predicted, and this drives most of the significant effect on $V = 500$ in Table 5. However, the negative sign on the interaction term of sabotage with the number of H types indicates that once again, group composition is an important component. Specifically, after controlling for the general aversion to sabotage,the competitiveness of the setting erodes this positive effect.[31] This goes against the competitive school of thought.

5 Conclusions

In this paper, our goal was to empirically examine if competitive pressures in a group production setting led to consistent behavioral responses not predicted by standard the-

[31] This is consistent for both the H and the L types. See Table 20 in Appendix A for the regressions for the two types separated.

ory. Our hypotheses were built around two schools of thought: the "competitive" school, which argues that competitive pressures lead to additional increases in effort, and the "noncompetitive" school, which argues that competitive pressures lead to toxic environments, where the decrease in help (or increase in sabotage) is not compensated by an increase in effort.

Our main take-away is that we find little to no support for the competitive school of thought in our setting. Pure team incentives appear sufficient for encouraging team production, and we find little need to augment those incentives with a competitive prize. The only aspect of competitive incentives that appears to work better than expected is the relatively low level of observed sabotage compared to what is predicted when the prize is high. However, the competitiveness of the setting did not lead to much higher levels of effort, as predicted by the competitive school, but rather individuals chose less sabotage *and* less effort than predicted.[32]

Our findings lead to several broader points. First, in a team setting where output is reliant upon at least a moderate degree of help, managers should avoid introducing competitive incentives. Giving additional (non-zero) prizes is costly, and we show that it does not lead to increases in output. If team production is not reliant upon helping behavior, then competitive incentives may work.

Second, we provide results regarding how changing the composition of a team—in terms of having more or fewer high ability types in it—affects behavior. We generally find that neither high nor low ability types change their effort much as we change the composition of the team. This suggests that people are less sensitive to team composition than expected. In regard to helping behavior, the prediction in our environment is that team members should take advantage of the efficiency enhancing nature of helping high productivity team members, and so, as their number increases, helping behavior should

[32]An additional aspect one might be concerned about under competitive incentives is the inequality induced, given that the H types are predicted to win much more often than the L types. We actually find that the L types win more than expected, indicating that while they still end up receiving the bonus payment much less often than the H types, the induced inequality is not as strong as predicted.

rise. We find that it generally does, though again not as much as predicted. There may be a behavioral effect here of the increased competition from too many high productivity types harming helping behavior. For total output, we find that balanced heterogeneous teams comprised of an equal number of high and low ability workers perform weakly better than any other team composition under competitive incentives. When there is only one high ability type, competition does not incentivize workers much because its outcome is very predictable. When there are too many high ability types, they compete intensely and help is underprovided. The team organization with balanced groups combines the best of the two worlds: There is enough help from low ability types to go around, and moderate competition between high ability types sustains effort at a high level.

Third, responses to incentives are not symmetric. There is a much greater willingness of individuals to respond to incentives by changing their effort than by changing their helping behavior. The lower willingness to modify others' effort through help leads to lower levels of net output when the highest amount of help is predicted to generate a high output (i.e., when there is no prize). However it works in a manager's favor when the highest amount of sabotage is predicted, at a high prize. This result may go in line with why organizations spend so much time building team cohesion through corporate retreats, rather than through incentivizing helping behavior. More work can be done in this domain to better understand these effects.

Finally, our results have implications for organizations choosing how much to reward individual effort and help in a group production setting. The empirical differences we observe at "extreme" incentives imply they do not really work. When effort is supposed to be the highest—at the highest prize value—it is lower than predicted. When help is supposed to be the highest—at the lowest prize value—it is lower than predicted. These findings have implications for theoretical assumptions regarding the effect of incentives on these components of output. Again, more work is needed to reveal what these functional forms should look like to approximate reality.

On balance, in our setting competitive incentives neither excessively hurt nor help. As we note though, different production functions which give rise to different values of help and individual effort could yield different conclusions. Thus, our results regarding the underlying effects on effort and helping behavior should be considered in light of any alternative weighting for those elements of team production.

References

Alchian, Armen A., and Harold Demsetz. 1972. "Production, information costs, and economic organization." *The American Economic Review*, 62(5): 777–795.

Balafoutas, Loukas, Florian Lindner, and Matthias Sutter. 2012. "Sabotage in Tournaments: Evidence from a Natural Experiment." *Kyklos*, 65(4): 425–441.

Bock, Olaf, Ingmar Baetge, and Andreas Nicklisch. 2014. "hroot: Hamburg Registration and Organization Online Tool." *European Economic Review*, 71(C): 117–120.

Brandts, Jordi, David J Cooper, Enrique Fatas, and Shi Qi. 2016. "Stand by Me - Experiments on Help and Commitment in Coordination Games." *Management Science*, 62: 2916–2936.

Brown, Michelle, and John Heywood. 2009. "Helpless in Finance: The Cost of Helping Effort Among Bank Employees." *Journal of Labor Research*, 30(2): 176–195.

Bull, Clive, Andrew Schotter, and Keith Weigelt. 1987. "Tournaments and Piece Rates: An Experimental Study." *Journal of Political Economy*, 95(1): 1–33.

Buser, Thomas, and Anna Dreber. 2016. "The flipside of comparative payment schemes." *Management Science*, 62(9): 2626–2638.

Carpenter, Jeffrey, Peter Hans Matthews, and John Schirm. 2010. "Tournaments and Office Politics: Evidence from a Real Effort Experiment." *American Economic Review*, 100(1): 504–17.

Charness, Gary, and David I. Levine. 2004. "Sabotage! Survey Evidence on When it is Acceptable." *Working Paper*.

Charness, Gary, David Masclet, and Marie Claire Villeval. 2014. "The Dark Side of Competition for Status." *Management Science*, 60(1): 38–55.

Chen, Kong-Pin. 2003. “Sabotage in Promotion Tournaments.” *Journal of Law, Economics and Organization*, 19(1): 119–140.

Chowdhury, Subhasish M., and Oliver Gürtler. 2015. “Sabotage in contests: a survey.” *Public Choice*, 164(1-2): 135–155.

Crosetto, Paolo, and Antonio Filippin. 2013. “The “bomb” risk elicitation task.” *Journal of Risk and Uncertainty*, 47(1): 31–65.

Danilov, Anastasia, Christine Harbring, and Bernd Irlenbusch. 2019. “Helping under a combination of team and tournament incentives.” *Journal of Economic Behavior & Organization*, 162: 120–135.

Dechenaux, Emmanuel, Dan Kovenock, and Roman M. Sheremeta. 2015. “A survey of experimental research on contests, all-pay auctions and tournaments.” *Experimental Economics*, 18(4): 609–669.

del Corral, Julio, Juan Prieto-Rodriguez, and Rob Simmons. 2010. “The Effect of Incentives on Sabotage: The Case of Spanish Football.” *Journal of Sports Economics*, 11(3): 243–260.

Deutscher, Christian, Bernd Frick, Oliver Gürtler, and Joachim Prinz. 2013. “Sabotage in Tournaments with Heterogeneous Contestants: Empirical Evidence from the Soccer Pitch.” *The Scandinavian Journal of Economics*, 115(4): 1138–1157.

Dickinson, David L., and R. Mark Isaac. 1998. “Absolute and relative rewards for individuals in team production.” *Managerial and Decision Economics*, 19(4-5): 299–310.

Drago, Robert, and Gerald T. Garvey. 1998. “Incentives for Helping on the Job: Theory and Evidence.” *Journal of Labor Economics*, 16(1): pp. 1–25.

Dutcher, E Glenn, Loukas Balafoutas, Florian Lindner, Dmitry Ryvkin, and Matthias Sutter. 2015. "Strive to be first or avoid being last: An experiment on relative performance incentives." *Games and Economic Behavior*, 94: 39–56.

Dye, RONALD A. 1984. "The Trouble with Tournaments." *Economic Inquiry*, 22(1): 147–149.

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

Garicano, Luis, and Ignacio Palacios-Huerta. 2005. "Sabotage in Tournaments: Making the Beautiful Game a Bit Less Beautiful." C.E.P.R. Discussion Papers CEPR Discussion Papers 5231.

Garvey, Gerald T., and Peter L. Swan. 1992. "Managerial objectives, capital structure, and the provision of worker incentives." *Journal of Labor Economics*, 10(4): 357–379.

Gürtler, Oliver, and Johannes Münster. 2013. "Rational self-sabotage." *Mathematical Social Sciences*, 65(1): 1–4.

Gürtler, Oliver, Johannes Münster, and Petra Nieken. 2013. "Information Policy in Tournaments with Sabotage." *The Scandinavian Journal of Economics*, 115(3): 932–966.

Hamilton, Barton H., Jack A. Nickerson, and Hideo Owan. 2003. "Team Incentives and Worker Heterogeneity: An Empirical Analysis of the Impact of Teams on Productivity and Participation." *Journal of Political Economy*, 111(3): 465–497.

Harbring, Christine, and Bernd Irlenbusch. 2008. "How many winners are good to have? On tournaments with sabotage." *Journal of Economic Behavior & Organization*, 65(3-4): 682–702.

Harbring, Christine, and Bernd Irlenbusch. 2011. "Sabotage in Tournaments: Evidence from a Laboratory Experiment." *Management Science*, 57(4): 611–627.

Harbring, Christine, Bernd Irlenbusch, Matthias Krakel, and Reinhard Selten. 2007. "Sabotage in Corporate Contests - An Experimental Analysis." *International Journal of the Economics of Business*, 14(3): 367–392.

Heap, Shaun P Hargreaves, Abhijit Ramalingam, and Brock V Stoddard. 2016. "Endowment inequality in public goods games: A re-examination." *Economics Letters*, 146: 4–7.

Hollinger, R.C., and J.P. Clark. 1983. *Theft by employees.* Lexington Books.

Johnson, David, and Timothy C. Salmon. 2016. "Sabotage versus Discouragement: Which Dominates Post Promotion Tournament Behavior?" *Southern Economic Journal*, 82(3): 673–696.

Kräkel, Matthias. 2005*a*. "Helping an Sabotaging in Tournaments." *International Game Theory Review*, 07(02): 211–228.

Kräkel, Matthias. 2005*b*. "Helping and sabotaging in tournaments." *International Game Theory Review*, 7(02): 211–228.

Lawler, Edward E., Susan Albers Mohrman, and George Benson. 2001. *Organizing for high performance: Employee involvement, TQM, reengineering, and knowledge management in the Fortune 1000: The CEO report.* Jossey-Bass.

Lazear, Edward P. 1989. "Pay Equality and Industrial Politics." *Journal of Political Economy*, 97(3): 561–80.

Lazear, Edward P., and Kathryn L. Shaw. 2007. "Personnel Economics: The Economist's View of Human Resources." *Journal of Economic Perspectives*, 21(4): 91–114.

Lazear, Edward P., and Sherwin Rosen. 1981. “Rank-Order Tournaments as Optimum Labor Contracts.” *European Journal of Political Economy.*

Ledyard, John O. 1995. “Public Goods: A Survey of Experimental Research.” 111–194.

Münster, Johannes. 2007. “Selection Tournaments, Sabotage, and Participation.” *Journal of Economics and Management Strategy*, 16: 943–970.

Nalebuff, Barry J., and Joseph E. Stiglitz. 1983. “Prizes and Incentives: Towards a General Theory of Compensation and Competition.” *The Bell Journal of Economics*, 14(1): pp. 21–43.

Sheremeta, Roman M. 2013. “Overbidding and heterogeneous behavior in contest experiments.” *Journal of Economic Surveys*, 27(3): 491–514.

Shubik, Martin. 1954. “Does the Fittest Necessarily Survive.” *in book*, in M. Shubik (ed.), Readings in Game Theory and Related Behavior(Doubleday, New York).

Tullock, Gordon. 1980. “Toward a Theory of the Rent Seeking Society.” *in book*, 97–112.

Vandegrift, Donald, and Abdullah Yavas. 2010. “An Experimental Test of Sabotage in Tournaments.” *Journal of Institutional and Theoretical Economics*, 166(2): 259–285.

A More Figures and Results

Table 7: Equilibrium predictions (Eq.) and observed averages (Obs.)

	LLLL		*HLLL*		*HHLL*		*HHHL*		*HHHH*	
	Eq.	Obs.	*Eq.*	Obs.	*Eq.*	Obs.	*Eq.*	Obs.	*Eq.*	Obs.
					$\boldsymbol{V} = 0$					
x_L	*9.77*	23.64	*9.77*	24.46	*9.77*	21.01	*9.77*	30.69		
x_H			*19.53*	25.88	*19.53*	37.23	*19.53*	31.36	*19.53*	35.54
k_{LL}	*9.77*	8.26	*9.77*	8.93	*9.77*	9.28				
k_{LH}			*19.53*	14.77	*19.53*	17.20	*19.53*	7.14		
k_{HL}			*9.77*	6.00	*9.77*	9.94	*9.77*	7.34		
k_{HH}					*19.53*	20.21	*19.53*	11.83	*19.53*	14.36
y_L	*39.06*	49.36	*39.06*	48.31	*39.06*	50.24	*39.06*	53.18		
y_H			*156.25*	140.38	*156.25*	184.00	*156.25*	137.54	*156.25*	151.26
Y	*156.25*	197.46	*273.44*	285.30	*390.63*	468.48	*507.81*	465.79	*625.00*	605.04
					$\boldsymbol{V} = 100$					
x_L	*28.52*	33.28	*22.01*	31.64	*18.77*	31.31	*16.87*	39.69		
x_H			*31.78*	35.78	*31.53*	47.67	*30.18*	44.74	*28.91*	39.05
k_{LL}	*3.52*	0.66	*7.72*	2.54	*8.77*	4.65				
k_{LH}			*15.45*	4.29	*17.53*	8.61	*18.35*	0.68		
k_{HL}			*1.60*	-3.20	*5.77*	2.32	*7.40*	1.91		
k_{HH}					*11.53*	5.00	*14.80*	3.55	*16.41*	7.10
y_L	*39.06*	36.00	*39.06*	34.44	*39.06*	41.79	*39.06*	48.58		
y_H			*156.25*	97.25	*156.25*	139.98	*156.25*	120.31	*156.25*	120.80
Y	*156.25*	144.00	*273.44*	200.58	*390.63*	363.53	*507.81*	409.5	*625.00*	483.21
					$\boldsymbol{V} = 500$					
x_L	*103.52*	57.81	*70.99*	50.44	*54.77*	48.89	*45.27*	45.47		
x_H			*80.76*	60.60	*79.53*	69.57	*72.79*	57.44	*66.41*	63.40
k_{LL}	*-21.48*	-5.46	*-0.44*	-3.80	*4.77*	0.81				
k_{LH}			*-0.88*	-3.51	*9.53*	4.59	*13.61*	-7.17		
k_{HL}			*-31.05*	-11.48	*-10.23*	-1.94	*-2.07*	-2.00		
k_{HH}					*-20.47*	-2.42	*-4.14*	-1.59	*3.91*	2.35
y_L	*39.06*	46.77	*39.06*	39.83	*39.06*	51.39	*39.06*	47.64		
y_H			*156.25*	101.58	*156.25*	155.76	*156.25*	115.31	*156.25*	144.58
Y	*156.25*	187.09	*273.44*	221.08	*390.63*	414.29	*507.81*	393.58	*625.00*	578.33

Table 8: Overview of observed averages and clustered standard errors

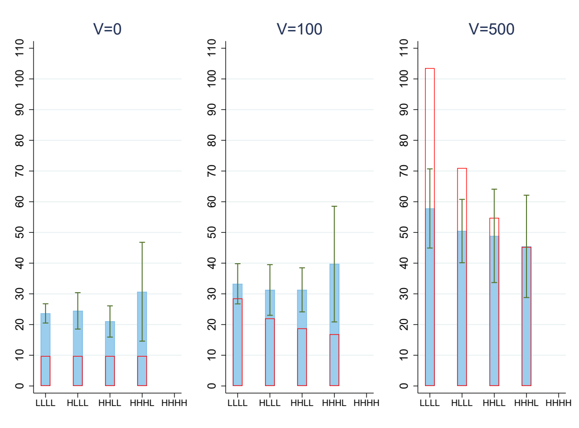
	LLLL	*HLLL*	*HHLL*	*HHHL*	HHHH
			$\boldsymbol{V}=0$		
x_L	23.64	24.46	21.01	30.69	
	(1.42)	(2.69)	(2.31)	(7.63)	
x_H		25.88	37.23	31.36	32.54
		(3.48)	(3.94)	(2.24)	(3.39)
k_{LL}	8.26	8.93	9.28		
	(1.11)	(1.10)	(1.72)		
k_{LH}		14.77	17.20	7.14	
		(2.29)	(3.53)	(9.67)	
k_{HL}		6.00	9.94	7.34	
		(1.42)	(1.22)	(0.67)	
k_{HH}			20.21	11.83	14.36
			(3.63)	(1.25)	(1.46)
y_L	49.36	48.31	50.24	53.18	
	(3.42)	(4.81)	(4.48)	(7.54)	
y_H		140.38	184.00	137.54	151.26
		(13.10)	(25.58)	(11.08)	(7.32)
Y	197.46	285.30	468.48	465.79	605.04
	(13.70)	(24.25)	(52.24)	(29.54)	(29.29)
			$\boldsymbol{V}=100$		
x_L	33.28	31.26	31.31	39.69	
	(2.97)	(3.75)	(3.26)	(8.93)	
x_H		35.74	47.67	44.74	39.05
		(3.86)	(6.16)	(4.19)	(2.32)
k_{LL}	0.66	2.54	4.65		
	(1.09)	(1.28)	(1.11)		
k_{LH}		4.29	8.61	0.68	
		(1.79)	(2.61)	(9.27)	
k_{HL}		-3.20	2.32	1.91	
		(2.66)	(2.25)	(1.23)	
k_{HH}			5.00	3.55	7.10
			(3.65)	(1.71)	(1.45)
y_L	36.00	34.44	41.79	48.58	
	(3.50)	(5.08)	(5.42)	(10.23)	
y_H		97.25	139.98	120.31	120.80
		(11.67)	(21.09)	(13.95)	(7.48)
Y	144.00	200.58	363.53	409.50	483.21
	(14.00)	(21.89)	(43.70)	(43.83)	(29.91)
			$\boldsymbol{V}=500$		
x_L	57.81	50.44	48.89	45.47	
	(5.85)	(4.69)	(6.91)	(7.90)	
x_H		60.60	69.57	57.44	63.40
		(11.72)	(9.43)	(4.64)	(4.89)
k_{LL}	-5.46	-3.80	0.81		
	(2.20)	(2.04)	(1.07)		
k_{LH}		-3.51	4.59	-7.17	
		(2.39)	(3.30)	(9.18)	
k_{HL}		-11.48	-1.94	-2.00	
		(4.61)	(3.31)	(1.96)	
k_{HH}			-2.42	-1.59	2.35
			(5.27)	(2.42)	(2.36)
y_L	46.77	39.83	51.39	47.64	
	(4.18)	(4.68)	(8.56)	(8.73)	
y_H		101.58	155.76	115.31	144.58
		(21.04)	(22.77)	(13.48)	(15.36)
Y	187.09	221.08	414.29	393.58	578.33
	(16.74)	(23.44)	(42.66)	(41.13)	(61.45)
numb. indiv.	48	48	48	72	48
numb. groups	12	12	12	18	12
numb. rounds	8	8	8	8	8

Stand. errors are clustered on the group level and displayed in parenthesis.

Table 9: Difference between observed averages and equilibrium predictions

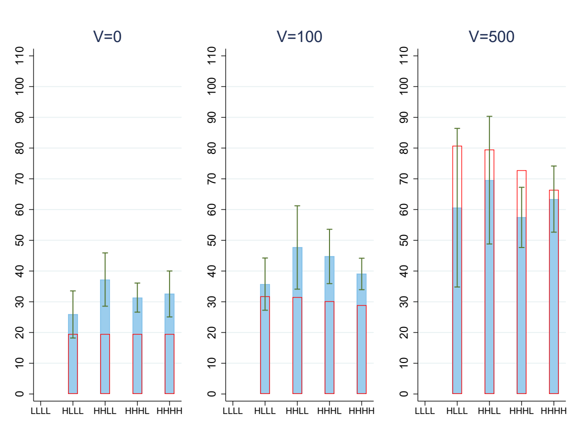
	$LLLL$	$LLLH$	$LLHH$	$LHHH$	$HHHH$
			$\boldsymbol{V} = 0$		
x_L	13.87	14.70	11.24	23.17	
x_H		6.34	17.70	12.95	13.01
k_{LL}	-1.50	-0.84	-0.49		
k_{LH}		-4.76	-2.33	-19.33	
k_{HL}		-3.77	0.17	-3.03	
k_{HH}			0.68	-6.46	-5.18
y_L	10.30	9.25	11.18	14.34	
y_H		-15.88	27.75	-19.01	-4.99
Y	41.21	11.86	77.85	-42.68	-19.96
			$\boldsymbol{V} = 100$		
x_L	4.76	9.25	12.54	25.58	
x_H		3.96	16.14	17.82	10.14
k_{LL}	-2.86	-5.18	-4.11		
k_{LH}		-11.16	-8.92	-24.90	
k_{HL}		-4.81	-3.45	-5.72	
k_{HH}			-6.53	-11.88	-9.31
y_L	-3.06	-4.62	2.72	11.47	
y_H		-59.00	-16.27	-38.84	-35.45
Y	-12.25	-72.85	-27.09	-105.05	-141.79
			$\boldsymbol{V} = 500$		
x_L	-45.71	-20.55	-5.88	0.24	
x_H		-20.15	-9.96	-10.37	-3.01
k_{LL}	16.02	-3.37	-3.96		
k_{LH}		-2.63	-4.94	-29.74	
k_{HL}		19.57	8.29	-2.11	
k_{HH}			18.05	-0.08	-1.56
y_L	7.71	0.77	12.32	4.64	
y_H		-54.67	-0.49	-49.23	-11.67
Y	30.84	-52.35	23.67	-143.05	-46.67

Figure 5: Average effort, L-types



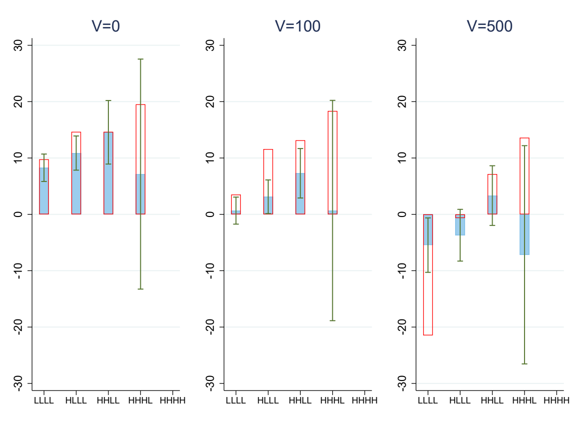
Predicted levels are shown as empty boxes and actual observed levels as filled boxes, with the standard error bars of the 95% confidence interval; we treat each group in one block as one independent observation, i.e. we have per group 1 observation per block—if groups are of heterogeneous types, we have the same group once in the graphs for H-types and once in the graphs for the L-types.

Figure 6: Average effort, H-types



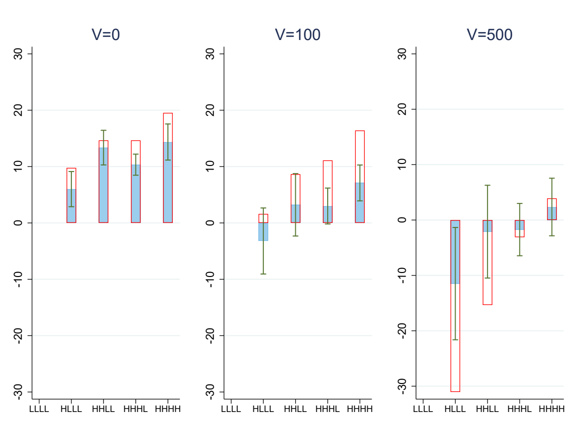
Predicted levels are shown as empty boxes and actual observed levels as filled boxes, with the standard error bars of the 95% confidence interval; we treat each group in one block as one independent observation, i.e. we have per group 1 observation per block—if groups are of heterogeneous types, we have the same group once in the graphs for H-types and once in the graphs for the L-types.

Figure 7: Average help, L-types



Predicted levels are shown as empty boxes and actual observed levels as filled boxes, with the standard error bars of the 95% confidence interval; we treat each group in one block as one independent observation, i.e. we have per group 1 observation per block—if groups are of heterogeneous types, we have the same group once in the graphs for H-types and once in the graphs for the L-types.

Figure 8: Average help, H-types



Predicted levels are shown as empty boxes and actual observed levels as filled boxes, with the standard error bars of the 95% confidence interval; we treat each group in one block as one independent observation, i.e. we have per group 1 observation per block—if groups are of heterogeneous types, we have the same group once in the graphs for H-types and once in the graphs for the L-types.

Table 10: Group composition—Robustness check: controlling for order effects

	(1) V=0	(2) V=100 Order 1	(3) V=500	(4) V=0	(5) V=100 Order 2	(6) V=500
	Dep. var.: $Y - Y^{*}$					
Homogeneous	-50.41	-70.76	-106.11	-84.04	-29.09	42.96
	(93.98)	(73.63)	(78.69)	(51.35)	(53.68)	(62.36)
Asymmetric	-50.52	-35.04	-118.59^{*}	-163.17^{***}	-121.63^{**}	-129.95^{**}
	(93.62)	(73.35)	(70.78)	(57.09)	(58.75)	(59.37)
Constant	64.18	16.14	73.64	91.52^{**}	-70.34	-26.32
	(91.24)	(67.87)	(62.74)	(44.60)	(44.91)	(46.11)
Observations	288	288	288	240	240	240
Number of groups	36	36	36	30	30	30

Robust standard errors clustered on group level in parentheses; *** = p<0.01, ** = p<0.05, * = p<0.1.

Table 11: Group composition—Robustness check 2: controlling for learning

	(1) V=0	(2) V=100 First half of periods	(3) V=500	(4) V=0	(5) V=100 Second half of periods	(6) V=500
	Dep. var.: $Y - Y^{*}$					
Homogeneous	-34.80	-29.12	-30.05	-99.65	-70.72	-33.10
	(46.87)	(50.45)	(53.40)	(67.35)	(45.97)	(55.42)
Asymmetric	-74.34^{*}	-49.74	-131.70^{***}	-122.29^{*}	-72.32	-94.58^{*}
	(44.86)	(54.89)	(49.11)	(69.66)	(48.29)	(52.79)
Constant	42.93	-29.46	45.97	112.77^{*}	-24.73	1.35
	(40.74)	(46.77)	(39.57)	(65.18)	(38.98)	(45.65)
Observations	264	264	264	264	264	264
Number of groups	66	66	66	66	66	66

Robust standard errors clustered on group level in parentheses; *** = p<0.01, ** = p<0.05, * = p<0.1.

Table 12: Total output—Robustness check: controlling for order effects

	(1)	(2)	(3)	(4)
	$H7$	$H8\&H9$	$H7$	$H8\&H9$
	Order 1		Order 2	
	Dep. variable: $Y - Y^{*}$			
V = 100	-47.08^{***}	-47.08^{***}	-123.26^{***}	-123.26^{***}
	(17.22)	(17.23)	(18.60)	(18.61)
V = 500	-43.14	-43.14	-53.75^{**}	-53.75^{**}
	(26.58)	(26.60)	(23.37)	(23.38)
Number H types		-21.65^{**}		-29.41^{**}
		(10.38)		(12.71)
Constant	22.12	69.04^{**}	-7.37	51.45^{**}
	(20.15)	(26.86)	(22.63)	(24.37)
Observations	864	864	720	720
Number of groups	36	36	30	30

Robust standard errors clustered on group level in parentheses; *** = p<0.01,** = p<0.05,* = p<0.1.; 'V = 100' is a dummy equal to 1 if V = 100 and zero otherwise; similarly for 'V = 500'. 'Number H types' runs from 0 to 4.

Table 13: Total output—Robustness check: controlling for learning

	(1)	(2)	(3)	(4)
	$H7$	$H8\&H9$	$H7$	$H8\&H9$
	First half of periods		Second half of periods	
	Dep. variable: $Y - Y^{*}$			
V = 100	-59.15^{***}	-59.15^{***}	-104.27^{***}	-104.27^{***}
	(15.14)	(15.15)	(16.37)	(16.38)
V = 500	-21.31	-21.31	-74.62^{***}	-74.62^{***}
	(20.35)	(20.37)	(19.66)	(19.67)
Number H types		-20.38^{**}		-27.64^{***}
		(8.78)		(7.64)
Constant	-3.51	39.11^{**}	20.94	78.74^{***}
	(14.49)	(17.19)	(18.23)	(21.54)
Observations	792	792	792	792
Number of groups	66	66	66	66

Robust standard errors clustered on group level in parentheses; *** = p<0.01,** = p<0.05,* = p<0.1.; 'V = 100' is a dummy equal to 1 if V = 100 and zero otherwise; similarly for 'V = 500'. 'Number H types' runs from 0 to 4.

Table 14: Effort—Robustness check: controlling for order effects

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	L	H	L	H	L	H	L	H
	Order 1				Order 2			
	Dep. variable: $x_{i,V} - x^*_{i,V}$							
V = 100	-3.08^{*}	3.13	-3.08^{*}	3.13	-6.07^{**}	-5.02^{*}	-6.07^{**}	-5.02^{*}
	(1.65)	(2.11)	(1.65)	(2.11)	(2.57)	(3.04)	(2.57)	(3.04)
V = 500	-38.83^{***}	-19.20^{***}	-38.83^{***}	-19.20^{***}	-39.15^{***}	-28.68^{***}	-39.15^{***}	-28.68^{***}
	(4.81)	(4.22)	(4.81)	(4.22)	(5.74)	(4.13)	(5.74)	(4.13)
Number H types			5.28^{**}	-0.43			10.76^{**}	3.93^{*}
			(2.36)	(2.62)			(4.24)	(2.29)
Constant	16.16^{***}	14.64^{***}	9.92^{***}	15.93^{*}	12.90^{***}	10.37^{***}	2.15	-1.43
	(1.78)	(1.97)	(2.98)	(8.40)	(2.38)	(2.71)	(3.68)	(6.87)
Observations	1,584	1,872	1,584	1,872	1,440	1,440	1,440	1,440
Number of indiv.	66	78	66	78	60	60	60	60
Number of groups	30	30	30	30	24	24	24	24

Robust standard errors clustered on group level in parentheses; *** = p<0.01 ** = p<0.05 * = p<0.1. 'V = 100' is a dummy equal to 1 if V = 100 and zero otherwise; similarly for 'V = 500'. 'Number H types' runs from 0 to 4.

Table 15: Effort—Robustness check: controlling for risk aversion

	(1)	(2)	(3)	(4)
	L	H	L	H
	Dep. var.: $(x_{i,V} - x^*_{i,V})$			
V = 100	-4.50^{***}	-0.41	-4.50^{***}	-0.41
	(1.49)	(1.86)	(1.50)	(1.86)
V = 500	-38.99^{***}	-23.32^{***}	-38.99^{***}	-23.32^{***}
	(3.68)	(2.98)	(3.68)	(2.98)
Number H types			7.71^{***}	1.87
			(2.00)	(1.94)
Risk Aversion	-0.39^{*}	-0.29^{***}	-0.38^{**}	-0.29^{***}
	(0.20)	(0.10)	(0.17)	(0.10)
Constant	-1.44	0.32	-9.85	-5.43
	(7.96)	(4.62)	(7.88)	(8.16)
Observations	3,024	3,312	3,024	3,312
Number of indiv.	126	138	126	138
Number of groups	54	54	54	54

Robust standard errors clustered on group level in parentheses; *** = p<0.01 ** = p<0.05 * = p<0.1. ; risk aversion runs from -100 to 0 with higher numbers indicating a higher risk aversion. 'V = 100' is a dummy equal to 1 if V = 100 and zero otherwise; similarly for 'V = 500'. 'Number H types' runs from 0 to 4.

Table 16: Effort—Robustness check: controlling for learning

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	L	H	L	H	L	H	L	H
	First half of the periods				Second half of the periods			
	Dep. variable: $x_{i,V} - x^*_{i,V}$							
V = 100	-5.98^{***}	0.02	-5.98^{***}	0.02	-3.03^{**}	-0.85	-3.03^{**}	-0.85
	(2.01)	(2.17)	(2.01)	(2.17)	(1.47)	(2.05)	(1.47)	(2.05)
V = 500	-41.19^{***}	-22.74^{***}	-41.19^{***}	-22.74^{***}	-36.79^{***}	-23.90^{***}	-36.79^{***}	-23.90^{***}
	(4.16)	(3.10)	(4.16)	(3.10)	(3.58)	(3.30)	(3.58)	(3.30)
Number H types			8.73^{***}	2.71			6.70^{***}	0.79
			(2.17)	(1.95)			(2.31)	(1.95)
Constant	16.61^{***}	13.01^{***}	7.04^{***}	4.87	12.61^{***}	12.56^{***}	5.27^{**}	10.19
	(1.43)	(1.61)	(2.53)	(5.76)	(1.75)	(1.96)	(2.35)	(6.33)
Observations	1,512	1,656	1,512	1,656	1,512	1,656	1,512	1,656
Number of indiv.	126	138	126	138	126	138	126	138
Number of groups	54	54	54	54	54	54	54	54

Robust standard errors clustered on group level in parentheses;*** = p<0.01 ** = p<0.05 * = p<0.1. 'V = 100' is a dummy equal to 1 if V = 100 and zero otherwise; similarly for 'V = 500'. 'Number H types' runs from 0 to 4.

Table 17: Help/sabotage—Robustness check: controlling for order effects

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	L	H	L	H	L	H	L	H
	Order 1				Order 2			
	Dep. variable: $k_{ij,V} - k^*_{ij,V}$							
V = 100	-3.52^{***}	-2.24^{**}	-3.52^{***}	-2.24^{**}	-4.05^{***}	-5.59^{***}	-4.05^{***}	-5.59^{***}
	(1.15)	(1.00)	(1.15)	(1.00)	(1.34)	(1.65)	(1.34)	(1.65)
V = 500	2.35	8.31^{***}	2.35	8.31^{***}	7.18^{**}	8.14^{***}	7.18^{**}	8.14^{***}
	(2.79)	(2.08)	(2.80)	(2.09)	(2.96)	(2.43)	(2.96)	(2.44)
Number H types			-2.74^{**}	-4.01^{***}			-11.06^{**}	-2.72^{*}
			(1.10)	(1.03)			(5.03)	(1.41)
Constant	-1.09	-4.92^{***}	2.15^{**}	7.11^{**}	-5.70^{*}	-3.71^{***}	5.36	4.46
	(1.20)	(0.92)	(1.09)	(3.07)	(3.11)	(1.13)	(3.35)	(4.00)
Observations	1,584	1,872	1,584	1,872	1,440	1,440	1,440	1,440
Number of indiv.	66	78	66	78	60	60	60	60
Number of groups	30	30	30	30	24	24	24	24

Robust standard errors clustered on group level in parentheses; *** = p<0.01 ** = p<0.05 * = p<0.1. 'V = 100' is a dummy equal to 1 if V = 100 and zero otherwise; similarly for 'V = 500'. 'Number H types' runs from 0 to 4.

Table 18: Help/sabotage—Robustness check: controlling for risk aversion

	(1) L	(2) H	(3) L	(4) H
	Dep. var.: $(k_{ij,V} - k^*_{ij,V})$			
V = 100	-3.77^{***}	-3.69^{***}	-3.77^{***}	-3.69^{***}
	(0.87)	(0.94)	(0.87)	(0.94)
V = 500	4.65^{**}	8.23^{***}	4.65^{**}	8.23^{***}
	(2.05)	(1.57)	(2.05)	(1.57)
Number H types			-6.08^{***}	-3.39^{***}
			(2.09)	(0.87)
Risk Aversion	−0.27	−0.06	−0.27	−0.07
	(0.26)	(0.05)	(0.24)	(0.04)
Constant	7.83	−1.77	14.47	8.64^{***}
	(9.52)	(2.17)	(10.72)	(3.34)
Observations	3,024	3,312	3,024	3,312
Number of indiv.	126	138	126	138
Number of groups	54	54	54	54

Robust standard errors clustered on group level in parentheses; *** = p<0.01 ** = p<0.05 * = p<0.1; risk aversion runs from -100 to 0 with higher numbers indicating a higher risk aversion. 'V = 100' is a dummy equal to 1 if V = 100 and zero otherwise; similarly for 'V = 500'. 'Number H types' runs from 0 to 4.

Table 19: Help/sabotage—Robustness check: controlling for learning

	(1) L	(2) H	(3) L	(4) H	(5) L	(6) H	(7) L	(8) H
	First half of the periods				Second half of the periods			
	Dep. variable: $k_{ij,V} - k^*_{ij,V}$							
V = 100	-3.25^{***}	-2.20^{**}	-3.25^{***}	-2.20^{**}	-4.30^{***}	-5.19^{***}	-4.30^{***}	-5.19^{***}
	(0.76)	(0.94)	(0.76)	(0.94)	(1.15)	(1.11)	(1.15)	(1.11)
V = 500	5.88^{***}	10.84^{***}	5.88^{***}	10.84^{***}	3.41^{*}	5.63^{***}	3.41^{*}	5.63^{***}
	(2.16)	(1.50)	(2.16)	(1.50)	(2.05)	(1.80)	(2.05)	(1.80)
Number H types			-6.63^{***}	-3.20^{***}			-5.55^{**}	-3.53^{***}
			(2.23)	(0.86)			(2.29)	(0.90)
Constant	-3.65^{**}	-5.44^{***}	3.61^{**}	4.17^{*}	-2.92^{*}	-3.35^{***}	3.16^{*}	7.24^{***}
	(1.55)	(0.71)	(1.49)	(2.42)	(1.69)	(0.82)	(1.68)	(2.71)
Observations	1,512	1,656	1,512	1,656	1,512	1,656	1,512	1,656
Number of indiv.	126	138	126	138	126	138	126	138
Number of groups	54	54	54	54	54	54	54	54

Robust standard errors clustered on group level in parentheses;*** = p<0.01,** = p<0.05,* = p<0.1.; 'V = 100' is a dummy equal to 1 if V = 100 and zero otherwise; similarly for 'V = 500'. 'Number H types' runs from 0 to 4.

Table 20: Controlling in particular for sabotage for the two types

	(1) L	(2) H	(3) L	(4) H
$V = 100$	-2.13^{*}	-2.24	-3.77^{***}	-3.69^{***}
	(1.24)	(2.85)	(0.87)	(0.94)
$V = 500$	10.39	27.50^{***}	-4.40^{**}	3.86^{**}
	(7.84)	(5.70)	(1.82)	(1.95)
Number H types	-2.63	-1.38^{*}	-3.52	-1.63^{**}
	(2.36)	(0.72)	(2.26)	(0.70)
$V = 100$ x number H types	-1.50^{**}	-0.48		
	(0.72)	(0.86)		
$V = 500$ x number H types	-6.50^{*}	-5.97^{***}		
	(3.65)	(1.58)		
Sabotage	6.73		20.70^{***}	22.90^{***}
	(7.91)		(2.28)	(5.19)
Sabotage x number H types	-10.89^{**}	-0.84	-16.65^{***}	-6.57^{***}
	(4.50)	(0.93)	(2.79)	(1.89)
Constant	-0.41	-0.24	0.57	0.48
	(1.50)	(2.08)	(1.40)	(2.04)
Observations	9,072	9,936	9,072	9,936
Number of indiv.	126	138	126	138
Number of groups	54	54	54	54

Robust standard errors clustered on group level in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.
'Number H types' runs from 0 to 4.

B Instructions

B.1 Instructions subjects received first

Dear participants,

welcome to today's experiment.

Please read the instructions for the experiment carefully. For a better understanding, in the following we will only use male designations. Those should be understood gender neutral. All statements in the instructions are true, and all participants receive exactly the same instructions. The experiment as well as the data analysis is anonymous.

We ask you to not talk to other participants and to use only the resources and devices that are provided by the conductors of the experiment. Please switch off all electronic devices. In addition, at the computer you are only allowed to use features that are necessary for the experiment. If you do not comply with these rules, you won't be paid in this experiment and you are not allowed to participate in any further experiments.

Your earnings in the experiment depend on your decisions and potentially the decisions of others. The currency used in the experiment is Tokens. Tokens will be converted to Euros at a rate of 100 Tokens to 6 Euro. You have already received a Euro 9.00 participation fee. Your earnings from the experiment will be incorporated into your participation fee. At the end of today's experiment, you will be paid in private and in cash.

The experiment will last around 90 Minutes. It consists of two parts; both parts are completely independent from each other. That is, your payment for part x only depends on decisions that you take in part x, and does not depend on decisions you take in the other part of the experiment.

At the beginning of each part you receive the corresponding instructions. We will read the instructions out loud and will give you time for questions. If you have a question, please raise your hand. Your question will then be answered privately. Thank you for your attention and for participating in this experiment.

B.2 Instructions subjects received second

The first part of the experiment consists of 24 periods, divided in three blocks of 8 periods; Blocks 1, 2 and 3. In each period you will be asked to make a set of decisions. At the beginning of each block, you will receive a new set of instructions. At the end of Part 1, we will randomly choose one period from each block to determine your earnings from Part 1. Because you do not know which periods will be chosen when you are making your decisions, you should make decisions in each period as if it were to be paid.

Remember that you were given 9 Euro show up fee at the beginning of the Experiment. Any gains or losses incurred in this part of the experiment will be offset against this amount.

Block 1

Matching

In Block 1, you will be matched with three other participants to make a group of four. You will stay in the same group for all 8 periods. To ensure anonymity, you and the three others in your group will be labeled by the computer program as member A, B, C and D. Each group member's label will be the same for all 8 periods.

Types

At the beginning of Block 1, each participant will be randomly assigned a type. You will be either an H or L type. You will keep this assignment for all 8 periods as well. You will know what type you are, and you will know the types of the other members in your group. More on the role of the types in a moment.

Decisions – Overview

Each round, you and your other group members will be working to generate a group output. The group output will determine how much each of you will earn. "Working" means that each of you will choose how many effort points to invest into an **individual output**. The sum of all individual outputs will determine your group output.

For your individual output you choose effort points on the range of 0 to 150 (a whole number). Each point you choose will be costly to you. You can find a table of costs on the separate sheet. These costs are denoted in Token. Notice that the first point will cost you 0.01 Token, 10 points cost you 1.28 Token, and 100 points 128 Token. This indicates that your per point-cost of effort is increasing with the total effort. These effort point costs will be subtracted from your earnings in each round.

If you are an L type, then your **total individual output** will be equal to your total effort points. If you are an H type, your total individual output will be equal to twice your total effort points.

All group members will choose simultaneously their own effort points.

The number of Token each group member receives from the total group output is equal to the total group output multiplied by 0.25. That is, for each point you and your group member's generate towards total group output, you and each of your group members receive 0.25 Token. For instance, if the total group output was 162 points, then you and every member of your group would receive 162*0.25=40.5 Token. To determine your net earnings you would then have to subtract off the cost of your chosen effort. For instance, if your effort was 23 points, your cost would be 6.72 Token. This would result in total earnings of 40.5-6.72 = 33.78 Token. Please turn to your screen and I will go through an example of how these decisions look like on the screen.

[read out loud] On the screen, you see a brief reminder of your task and a box where you will be able to type in the number of effort points you wish to choose for your individual output. You can choose any number of points between 0 and 150. Please type in 20. On the bottom of the screen, there is a calculator to calculate the costs. This calculator automatically updates the costs of your choices when you press the "calculate" button. These costs are the same as those in the Cost Table. If you have entered a choice of 20, and you press the "calculate" button, you will notice that it shows you the cost is 5.12 Token, the number that corresponds to a cost of a choice of 20 on your Cost Table. Notice that what you earn from your 20 points of effort you chose for your individual output is $20 * 0.25 = 5$ *Token if you an L type and* $2 * 20 * 0.25 = 10$ *Token if you are an H type. Please turn your attention back to the instructions and we will describe the next task in the experiment.*

Alterations to Efforts of Other Group Members

In addition to making your own effort choice, in each round you will also be able to affect the effort of your group members. On the screen you can choose additional effort points towards increasing or decreasing the effort of your group members. You will be able to modify the effort of others by increasing or decreasing their effort by up to 150 effort points per group member. Each of these efforts again means costs to you as shown in the table.

This means you will have a total of 4 decisions to make per period. You will choose how many effort points to exert towards your own effort. Then you decide for each of your group members regarding whether and how much you want to alter their effort. Regardless of your type, each effort point you choose towards raising or lowering the effort of others changes their effort by one point.

When determining each group members total individual output, we will first add their own effort points with all of the effort points others have chosen to increase or decrease the effort. If that individual is an L type, their total individual output is the same as their total effort. If that individual is an H type, their total individual output will be twice this sum. Note that this means that you can alter the total individual output of an H type by two points per 1 point of effort you chose. Each point of effort you choose to alter t the individual output of an L type alters their total effort by only 1 point. Similarly, your group member's choices affect your total individual output.

For instance, if you chose 10 effort points for your effort and your group members modified your effort by 5, -2, and 19 points, your modified individual effort would be 10+5-2+19=32. If you are an L type, your total individual output would be 32. If you are an H type, your total individual output would be 32*2 = 64. Similar calculations also hold for your group members who are H or L types. It is possible that your total individual output will be negative. In this case, the computer will assign you a total individual output of zero so that you will never have a negative total individual output.

Each effort point is costly, independently on whether you chose it for your own effort or to affecting the efforts of others. All of your efforts determine your total cost (in Token). For instance, if you altered (increased or decreased) the effort of each team member by 10 points and chose an individual effort of 10 points, the total cost to you would be 1.28+1.28+1.28+1.28=5.12 Token: the cost of an effort level of 10 is equal to 1.28 (as in the cost Table).

It is important to note that effort costs are treated separately for each decision. If for instance, consider the following example: you chose to reduce the effort of one group member by 20 (by choosing -20), leave the one of another group member unchanged (by choosing 0) and increased the one of the third group member by 10 (by choosing 10). In addition, you chose an own effort of 10 points. Then you would still be expending a total of 40 effort points (20+10+10). But, according to the cost table, your cost for these decisions would be 5.12+0+1.28+1.28=7.68 Token, which are the costs associated with effort points of 20, 0, 10 and 10. Please turn to your screen, on which I will walk you through such a decision situation.

On your screen, you are given a brief reminder of your task and an input box for each of your group members. The points chosen for each of these input boxes determine the amount you wish to alter each group member's effort. You can choose any number between -150 and 150 (negative 150 and positive 150) points. Notice beside each input box is each group members' label (A, B, C or D) and their type (H or L).

Please type in -18 in the first box, 10 in the second and 23 in the third. On the bottom of the screen, there is a calculator for you to use and calculate costs. This calculator automatically updates with the costs of your choices when you press the "calculate" button. These costs are taken from the cost table you've been given. If you have chosen -18, 10 and 23 points and you press the "calculate" button, you see the following cost for these choices: 12.2 Token (4.15+1.28 +6.77). These are the same as the sum of costs of alterations of 18, 10 and 23 on your cost table.

You will also notice that the on-screen calculator asks you to enter a hypothetical amount you believe your group members will contribute to the group output. This is purely for you to be able to understand how payoffs work. Anything entered here has no impact on your actual payoff or the decisions of others. Below this, you can see how the group output and total payoffs change as you change your choices of your effort points and modify the efforts of others. Please turn your attention back to the instructions and we will go through the feedback you will receive after a round is over.

Feedback

After each round, you will be shown a results screen which will show you your four decisions. In addition, you see the following information: your total individual effort and total individual output after the modifications from your group members, the average modification of efforts from your group, the total group output, the costs associated with your choices, and your payoff, should that round be chosen for payment.

Payoff Example

We will now explain you by means of an example how your payoff is calculated. Let's assume that the total group output from your group was 162 points. The gain you would receive from this total group output is 162*0.25=40.5 Token. Let's also assume you chose 20 points for your own effort and chose to alter the efforts of your group members by -18, 10 and 23 points, respectively. This would result in a total cost of 5.12+4.15+1.28+6.77 = 17.32 Token. Thus, in this example, the Token gained in this period would be your gains minus your total costs: 40.5-17.32=23.18.

Summary

At the beginning of the first block, you will be randomly assigned a type, L or H, and will be grouped with 3 other people to make a group of 4. You will keep this type and this group for the entire 8 periods of block 1. Each period you must choose the number of points for your own individual output and modifications to each of your group member's individual outputs. The costs of these decisions will be deducted from your gains from the total group output.

Are there any questions?

If not, please turn to your screen. There you will be shown your type and the types of your other 3 group members. After reviewing this information, please click on the "continue" button, and the first round of Block 1 will begin. If you are finished making decisions on a screen, you must click on the "continue" button to advance. The program only advances if everyone has clicked on the continue button for a given portion so, please pay attention *to the screen and click the "Continue" button if you are finished making decisions on that screen.*

B.3 Instructions subjects received third

Block 2

Block 2 is similar to Block 1 except for one change. You and your group members will still take decisions for 8 periods. You have the same four choices of your own effort and modifications to your group members' efforts as in Block 1. The costs and gains from the group output are as previously defined. Also, you are in the same group as before, and your types are the same as in Block 1.

In Block 2, however, there will be a bonus awarded to one of the group members, to encourage higher effort. The bonus will be awarded using a lottery, where your probability of winning is increasing in your total individual output, and is decreasing in the total individual output of the others.

If you win the bonus, you get 100 [500] Token. Only one group member can win the bonus per period. Your probability of winning is determined as

$$\text{Chance of winning} = \frac{\text{Your total individual output (TIO)}}{\text{TIO of A + TIO of B + TIO of C + TIO of D}}$$

As an example, suppose that your total individual output was 30 points and that the other members of your group had total individual outputs of 21, 52 and 9 points. Your chance of winning is thus 30/(30+21+52+9)=0.27, or 27% (rounded). Likewise, the chance of each of your group members to win the bonus is 19%, 46% and 8% for the group members who had 21, 52 and 9 points respectively. It is easy to see that increases in total individual output lead to a greater chance in winning the prize.

To see how the likelihoods work, imagine the percentages represent the number of balls each group member has in a common container. If someone randomly selected one of these balls to determine the winner, the chance of winning can now be thought of as the likelihood your own ball is drawn. Group member C, who has 46 balls, has a much higher chance of their ball being drawn than group member D, who only has 8 balls.

Let's go through another example. If your group members' total indiv. outcomes had remained the same, but you had a total individual output of 40 (instead of 30), your chance of winning would increase from 27% to 33% (40/(40+21+52+9)=0.33 or 33%). Since your chance of winning went up, your group members' chances of winning went down to 17%, 43% and 7% (for the group members who had total individual outputs of 21, 52 and 9 points respectively).

Likewise, the chances will also change if your total individual output had remained the same, but the total individual output of one of your group members had changed (because they chose a different effort or their total individual output was modified by you or other group members).

For this example, assume that your total individual output was again 30 and the total individual output of two of your group members was still 21 and 52, but the fourth group member had an increase in his total individual output to 18 (instead of 9). Now, instead of you having a 27% chance of winning you would have a chance of winning of 23% (30/(30+21+52+18) = 0.23 or 23%). Similarly, your group members would have a 16%, 40% and 14% chance of winning respectively. Notice that the group member whose total individual output is higher, now has a larger chance of winning, while all other group members have a smaller chance of winning. Similarly, your chances of winning the bonus can increase if the total individual output of another member of your group goes down.

Continue with the same example. **Suppose** your total individual output is 30, the total indiv. outputs of two other group members are 21 and 9, but the total individual output of the group member who previously had 52 decreases to 30. Then your chance of winning is 33% (30/(30+21+9+30)=0.33 or 33%).To assist you in your decision, the calculator on the screen will now also show you how your chance of winning changes as you change your choices.

End of a Period

At the end of each period, you will see the same information as in Block 1, except now, you will also be told if you won the bonus or not.

Do you have questions?

B.4 Instructions subjects received fourth

Block 3

The instructions for Block 3 are very similar to those from Block 2. Specifically, you and your group members will still make decisions for 8 periods where you have the same four choices: your own effort and the modifications to your group members' efforts. The costs and the and gains from the group output are as previously defined. You are in the same group as before and the types are the same as before. The only change is that in Block 3, the size of the bonus has increased to 500 [100] Token. Everything else stays as in Block 2. Do you have questions?

B.5 Instructions subjects received fifth

On your computer screen you will see a square composed of 100 numbered boxes, like shown below.

Behind one of these boxes hides a mine; all the other 99 boxes are free from mines. You do not know where this mine lies. You only know that the mine can be in any place

1	2	3	4	5	6	7	8	9	10
11	12	13	14	15	16	17	18	19	20
21	22	23	24	25	26	27	28	29	30
31	32	33	34	35	36	37	38	39	40
41	42	43	44	45	46	47	48	49	50
51	52	53	54	55	56	57	58	59	60
61	62	63	64	65	66	67	68	69	70
71	72	73	74	75	76	77	78	79	80
81	82	83	84	85	86	87	88	89	90
91	92	93	94	95	96	97	98	99	100

with equal probability. Your task is to decide how many boxes to collect. Boxes will be collected in numerical order, starting with number 1. So you will be asked to choose a number between 1 and 100. At the end of the experiment we will randomly determine the number of the box containing the mine. If you happen to have harvested the box where the mine is located – i.e. if your chosen number is greater than or equal to the drawn number – you will earn zero. If the mine is located in a box that you did not harvest – i.e. if your chosen number is smaller than the drawn number – you will earn an amount equivalent to the number you have chosen.

Table 21: Cost Table

Your choice of effort	costs	Your choice of effort	costs	Your choice of effort	costs
0	0.00	51	33.29	101	130.57
1	0.01	52	34.61	102	133.17
2	0.05	53	35.96	103	135.80
3	0.12	54	37.32	104	138.44
4	0.20	55	38.72	105	141.12
5	0.32	56	40.14	106	143.82
6	0.46	57	41.59	107	146.55
7	0.63	58	43.06	108	149.30
8	0.82	59	44.56	109	152.08
9	1.04	60	46.08	110	154.88
10	1.28	61	47.63	111	157.71
11	1.55	62	49.20	112	160.56
12	1.84	63	50.80	113	163.44
13	2.16	64	52.43	114	166.35
14	2.51	65	54.08	115	169.28
15	2.88	66	55.76	116	172.24
16	3.28	67	57.46	117	175.22
17	3.70	68	59.19	118	178.23
18	4.15	69	60.94	119	181.26
19	4.62	70	62.72	120	184.32
20	5.12	71	64.52	121	187.40
21	5.64	72	66.36	122	190.52
22	6.20	73	68.21	123	193.65
23	6.77	74	70.09	124	196.81
24	7.37	75	72.00	125	200.00
25	8.00	76	73.93	126	203.21
26	8.65	77	75.89	127	206.45
27	9.33	78	77.88	128	209.72
28	10.04	79	79.88	129	213.00
29	10.76	80	81.92	130	216.32
30	11.52	81	83.98	131	219.66
31	12.30	82	86.07	132	223.03
32	13.11	83	88.18	133	226.42
33	13.94	84	90.32	134	229.84
34	14.80	85	92.48	135	233.28
35	15.68	86	94.67	136	236.75
36	16.59	87	96.88	137	240.24
37	17.52	88	99.12	138	243.76
38	18.48	89	101.39	139	247.31
39	19.47	90	103.68	140	250.88
40	20.48	91	106.00	141	254.48
41	21.52	92	108.34	142	258.10
42	22.58	93	110.71	143	261.75
43	23.67	94	113.10	144	265.42
44	24.78	95	115.52	145	269.12
45	25.92	96	117.96	146	272.84
46	27.08	97	120.44	147	276.60
47	28.28	98	122.93	148	280.37
48	29.49	99	125.45	149	284.17
49	30.73	100	128.00	150	288.00
50	32.00				

C Screenshots

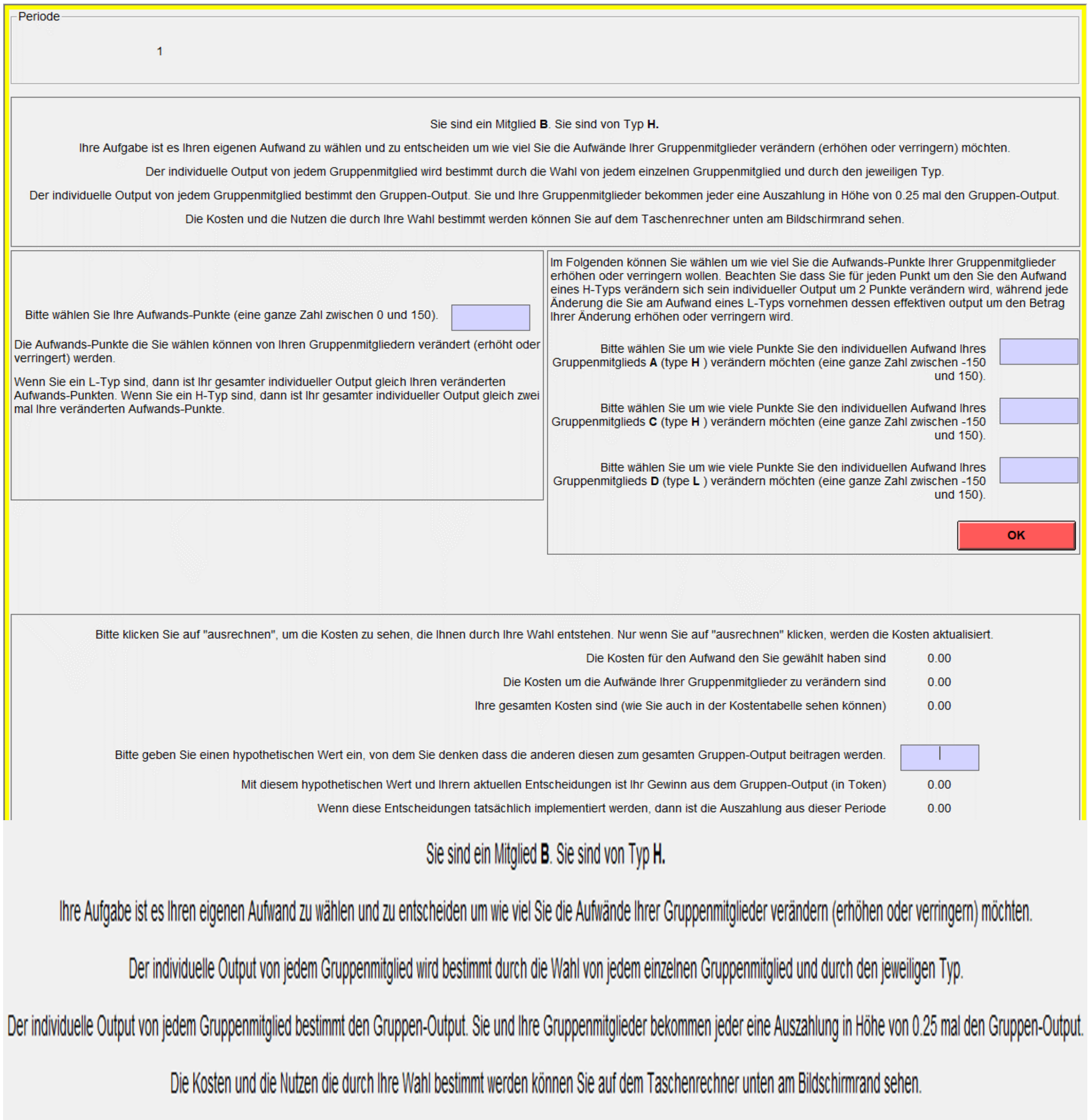
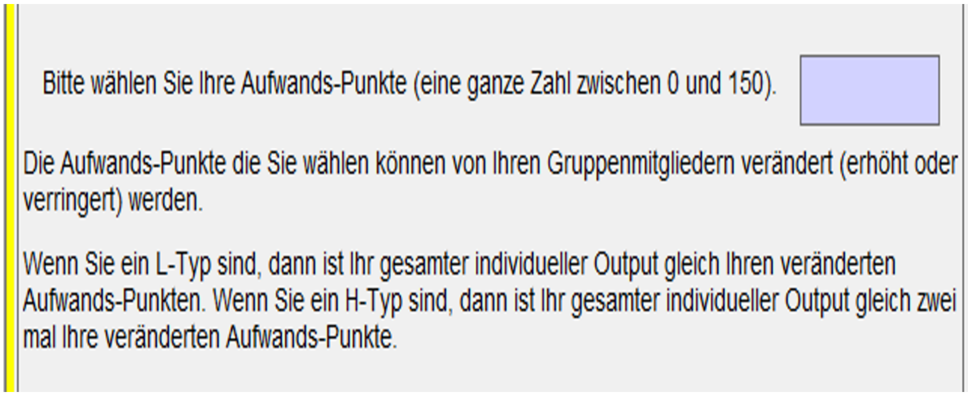


Figure 10: Full screen

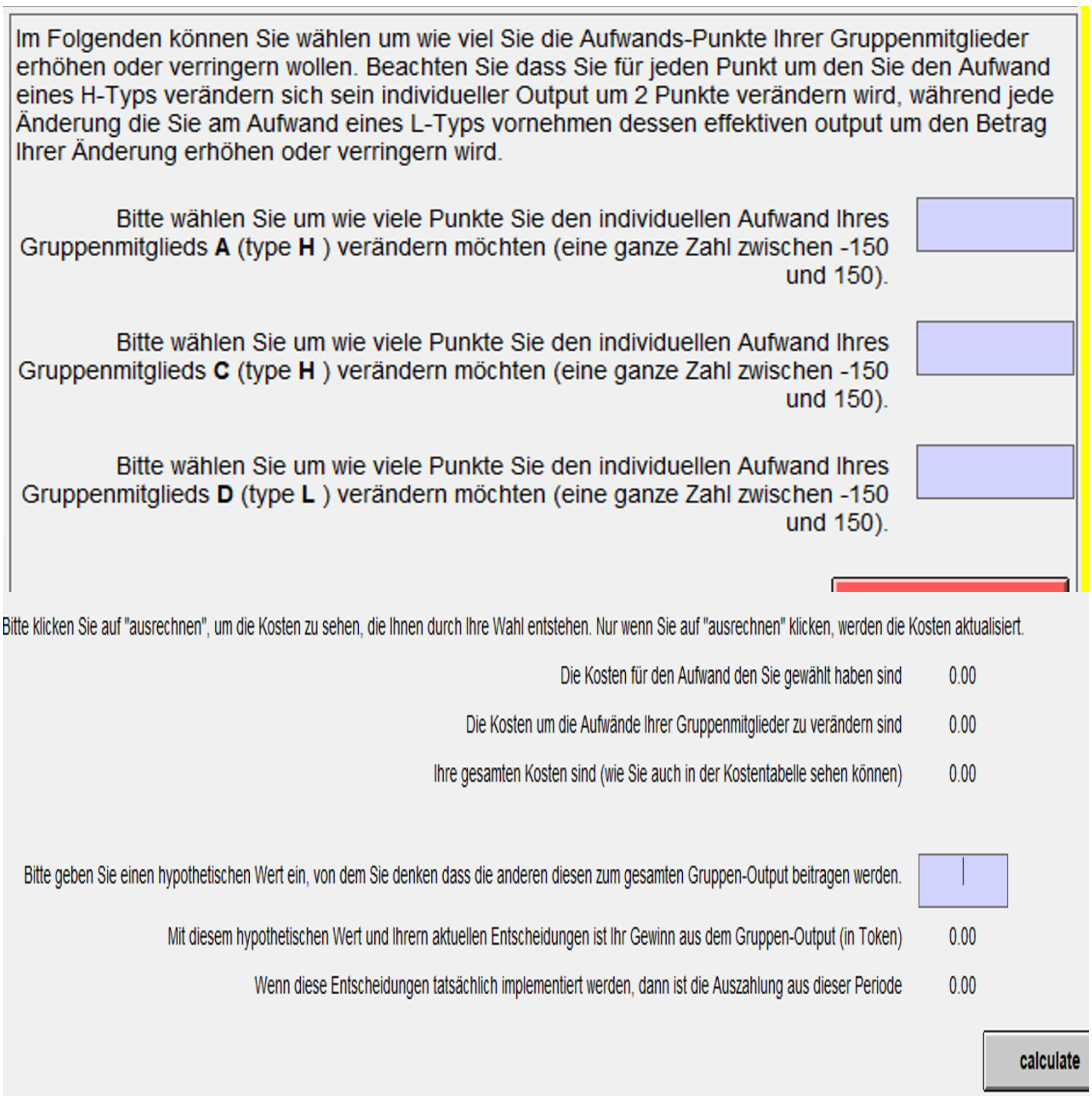


Bitte wählen Sie Ihre Aufwands-Punkte (eine ganze Zahl zwischen 0 und 150).

Die Aufwands-Punkte die Sie wählen können von Ihren Gruppenmitgliedern verändert (erhöht oder verringert) werden.

Wenn Sie ein L-Typ sind, dann ist Ihr gesamter individueller Output gleich Ihren veränderten Aufwands-Punkten. Wenn Sie ein H-Typ sind, dann ist Ihr gesamter individueller Output gleich zwei mal Ihre veränderten Aufwands-Punkte.

Figure 11: Full screen



Im Folgenden können Sie wählen um wie viel Sie die Aufwands-Punkte Ihrer Gruppenmitglieder erhöhen oder verringern wollen. Beachten Sie dass Sie für jeden Punkt um den Sie den Aufwand eines H-Typs verändern sich sein individueller Output um 2 Punkte verändern wird, während jede Änderung die Sie am Aufwand eines L-Typs vornehmen dessen effektiven output um den Betrag Ihrer Änderung erhöhen oder verringern wird.

Bitte wählen Sie um wie viele Punkte Sie den individuellen Aufwand Ihres Gruppenmitglieds **A** (type **H**) verändern möchten (eine ganze Zahl zwischen -150 und 150).

Bitte wählen Sie um wie viele Punkte Sie den individuellen Aufwand Ihres Gruppenmitglieds **C** (type **H**) verändern möchten (eine ganze Zahl zwischen -150 und 150).

Bitte wählen Sie um wie viele Punkte Sie den individuellen Aufwand Ihres Gruppenmitglieds **D** (type **L**) verändern möchten (eine ganze Zahl zwischen -150 und 150).

Bitte klicken Sie auf "ausrechnen", um die Kosten zu sehen, die Ihnen durch Ihre Wahl entstehen. Nur wenn Sie auf "ausrechnen" klicken, werden die Kosten aktualisiert.

Die Kosten für den Aufwand den Sie gewählt haben sind 0.00

Die Kosten um die Aufwände Ihrer Gruppenmitglieder zu verändern sind 0.00

Ihre gesamten Kosten sind (wie Sie auch in der Kostentabelle sehen können) 0.00

Bitte geben Sie einen hypothetischen Wert ein, von dem Sie denken dass die anderen diesen zum gesamten Gruppen-Output beitragen werden.

Mit diesem hypothetischen Wert und Ihrern aktuellen Entscheidungen ist Ihr Gewinn aus dem Gruppen-Output (in Token) 0.00

Wenn diese Entscheidungen tatsächlich implementiert werden, dann ist die Auszahlung aus dieser Periode 0.00

calculate

Figure 13: Full screen

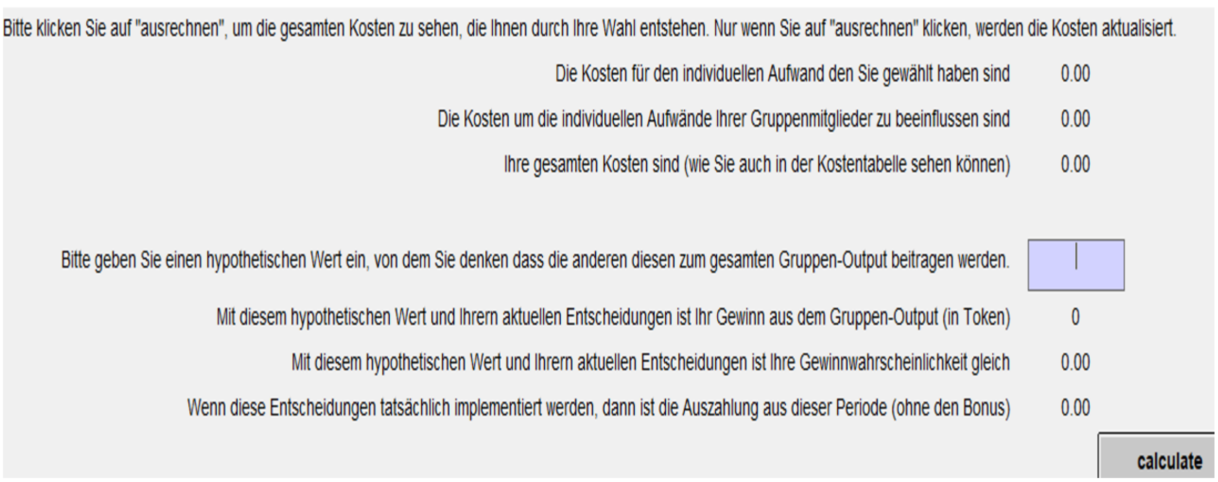


Figure 14: Full screen

University of Innsbruck - Working Papers in Economics and Statistics
Recent Papers can be accessed on the following webpage:

https://www.uibk.ac.at/eeecon/wopec/

2021-26 **Glenn E. Dutcher, Regine Oexl, Dmitry Ryvkin, Tim Salmon:** Competitive versus cooperative incentives in team production with heterogeneous agents

2021-25 **Anita Gantner, Regine Oexl:** Respecting Entitlements in Legislative Bargaining - A Matter of Preference or Necessity?

2021-24 **Silvia Angerer, E. Glenn Dutcher, Daniela Glätzle-Rützler, Philipp Lergetporer, Matthias Sutter:** The formation of risk preferences throughsmall-scale events

2021-23 **Stjepan Srhoj, Dejan Kovač, Jacob N. Shapiro, Randall K. Filer:** The Impact of Delay: Evidence from Formal Out-of-Court Restructuring

2021-22 **Octavio Fernández-Amador, Joseph F. Francois, Doris A. Oberdabernig, Patrick Tomberger:** Energy footprints and the international trade network: A new dataset. Is the European Union doing it better?

2021-21 **Felix Holzmeister, Jürgen Huber, Michael Kirchler, Rene Schwaiger:** Nudging Debtors to Pay Their Debt: Two Randomized Controlled Trials

2021-20 **Daniel Müller, Elisabeth Gsottbauer:** Why Do People Demand Rent Control?

2021-19 **Alexandra Baier, Loukas Balafoutas, Tarek Jaber-Lopez:** Ostracism and Theft in Heterogeneous Groups

2021-18 **Zvonimir Bašić, Parampreet C. Bindra, Daniela Glätzle-Rützler, Angelo Romano, Matthias Sutter, Claudia Zoller:** The roots of cooperation

2021-17 **Silvia Angerer, Jana Bolvashenkova, Daniela Glätzle-Rützler, Philipp Lergetporer, Matthias Sutter:** Children's patience and school-track choices several years later: Linking experimental and field data

2021-16 **Daniel Gründler, Eric Mayer, Johann Scharler:** Monetary Policy Announcements, Information Schocks, and Exchange Rate Dynamics

2021-15 **Sebastian Bachler, Felix Holzmeister, Michael Razen, Matthias Stefan:** The Impact of Presentation Format and Choice Architecture on Portfolio Allocations: Experimental Evidence

2021-14 **Jeppe Christoffersen, Felix Holzmeister, Thomas Plenborg:** What is Risk to Managers?

2021-13 **Silvia Angerer, Daniela Glätzle-Rützler, Christian Waibel:** Trust in health care credence goods: Experimental evidence on framing andsubject pool effects

2021-12 **Rene Schwaiger, Laura Hueber:** Do MTurkers Exhibit Myopic Loss Aversion?

2021-11 **Felix Holzmeister, Christoph Huber, Stefan Palan:** A Critical Perspective on the Conceptualization of Risk in Behavioral and Experimental Finance

2021-10 **Michael Razen, Alexander Kupfer:** Can increased tax transparency curb corporate tax avoidance?

2021-09 **Changxia Ke, Florian Morath, Anthony Newell, Lionel Page:** Too big to prevail: The paradox of power in coalition formation

2021-08 **Marco Haan, Pim Heijnen, Martin Obradovits:** Competition with List Prices

2021-07 **Martin Dufwenberg, Olof Johansson-Stenman, Michael Kirchler, Florian Lindner, Rene Schwaiger:** Mean Markets or Kind Commerce?

2021-06 **Christoph Huber, Jürgen Huber, and Michael Kirchler:** Volatility Shocks and Investment Behavior

2021-05 **Max Breitenlechner, Georgios Georgiadis, Ben Schumann:** What goes around comes around: How large are spillbacks from US monetary policy?

2021-04 **Utz Weitzel, Michael Kirchler:** The Banker's Oath And Financial Advice

2021-03 **Martin Holmen, Felix Holzmeister, Michael Kirchler, Matthias Stefan, Erik Wengström:** Economic Preferences and Personality Traits Among Finance Professionals and the General Population

2021-02 **Christian König-Kersting:** On the Robustness of Social Norm Elicitation

2021-01 **Laura Hueber, Rene Schwaiger:** Debiasing Through Experience Sampling: The Case of Myopic Loss Aversion.

2020-34 **Kai A. Konrad, Florian Morath:** The Volunteer's Dilemma in Finite Populations

2020-33 **Katharina Momsen, Markus Ohndorf:** Expressive Voting vs. Self-Serving Ignorance

2020-32 **Silvia Angerer, Daniela Glätzle-Rützler, Christian Waibel:** Monitoring institutions in health care markets: Experimental evidence

2020-31 **Jana Friedrichsen, Katharina Momsen, Stefano Piasenti:** Ignorance, Intention and Stochastic Outcomes

2020-30 **Esther Blanco, Alexandra Baier, Felix Holzmeister, Tarek Jaber-Lopez, Natalie Struwe:** Substitution of social concerns under the Covid-19 pandemic

2020-29 **Andreas Hackethal, Michael Kirchler, Christine Laudenbach, Michael Razen, Annika Weber:** On the (ir)relevance of monetary incentives in risk preference elicitation experiments

2020-28 **Andrej Gill, Matthias Heinz, Heiner Schumacher, Matthias Sutter:** Trustworthiness in the Financial Industry

2020-27 **Matthias Sutter, Michael Weyland, Anna Untertrifaller, Manuel Froitzheim:** Financial literacy, risk and time preferences - Results from a randomized educational intervention

2020-26 **Rene Schwaiger, Jürgen Huber, Michael Kirchler, Daniel Kleinlercher, Utz Weitzel:** Unequal Opportunities, Social Groups, and Redistribution

2020-25 **Roman Inderst, Martin Obradovits:** Competitive Strategies when Consumers are Relative Thinkers: Implications for Pricing, Promotions, and Product Choice

2020-24 **Martin Obradovits, Philipp Plaickner:** Price-Directed Search and Collusion

2020-23 **Helena Fornwagner, Oliver P. Hauser:** Climate action for (my) children

2020-22 **Esther Blanco, Natalie Struwe, James M. Walker:** Incentivizing public good provision through outsider transfers: experimental evidence on sharing rules and additionality requirements

2020-21 **Loukas Balafoutas, Helena Fornwagner, Rudolf Kerschbamer, Matthias Sutter, Maryna Tverdostup:** Diagnostic Uncertainty and Insurance Coverage in Credence Goods Markets

2020-20 **Anna Ulrichshofer, Markus Walzl:** Customer Disputes, Misconduct, and Reputation Building in the Market for Financial Advice

2020-19 **Anna Ulrichshofer, Markus Walzl:** Social Comparison and Optimal Contracts in the Competition for Managerial Talent

2020-18 **Martin Obradovits, Philipp Plaickner:** Searching for Treatment

2020-17 **Jun Honda:** The Gender-Punishment Gap revisited

2020-16 **Jun Honda:** The Relation between Rankings and Risk-Taking in the Labor Market for Financial Advice

2020-15 **Christina Bannier, Eberhard Feess,Natalie Packham, Markus Walzl:** Differentiation and Risk-Aversion in Imperfectly Competitive Labor Markets

2020-14 **Felix Holzmeister, Rudolf Kerschbamer:** oTree: The Equality Equivalence Test

2020-13 **Parampreet Christopher Bindra, Graeme Pearce:** The effect of priming on fraud: Evidence from a natural field experiment

2020-12 **Alessandro De Chiara, Marco A. Schwarz:** A Dynamic Theory of Regulatory Capture

2020-11 **Christoph Huber, Jürgen Huber, Michael Kirchler:** Market shocks and professionals' investment behavior - Evidence from the COVID-19 crash

2020-10 **Elisabeth Gsottbauer, Daniel Müller, Samuel Müller, Stefan T. Trautmann, Galina Zudenkova:** Social class and (un)ethical behavior: Causal versus correlational evidence

2020-09 **Parampreet Christopher Bindra, Rudolf Kerschbamer, Daniel Neururer, Matthias Sutter:** Reveal it or conceal it: On the value of second opinions in a low-entry-barriers credence goods market

2020-08 **Robert Steiger, Eva Posch, Gottfried Tappeiner, Janette Walde:** Effects of climate change on tourism demand considering individual seasonal preferences

2020-07 **Fang Liu, Alexander Rasch, Marco A. Schwarz, Christian Waibel:** The role of diagnostic ability in markets for expert services

2020-06 **Matthias Stefan, Jürgen Huber, Michael Kirchler, Matthias Sutter, Markus Walzl:** Monetary and Social Incentives in Multi-Tasking: The Ranking Substitution Effect

2020-05 **Michael Razen, Jürgen Huber, Laura Hueber, Michael Kirchler, Matthias Stefan:** Financial Literacy, Economic Preferences, and Adolescents' Field Behavior

2020-04 **Christian König-Kersting, Johannes Lohse, Anna Louisa Merkel:** Active and Passive Risk-Taking

2020-03 **Christoph Huber, Jürgen Huber:** Bad bankers no more? Truth-telling and (dis)honesty in the finance industry

2020-02 **Dietmar Fehr, Daniel Müller, Marcel Preuss:** Social Mobility Perceptions and Inequality Acceptance

2020-01 **Loukas Balafoutas, Rudolf Kerschbamer:** Credence goods in the literature: What the past fifteen years have taught us about fraud, incentives, and the role of institutions

University of Innsbruck

Working Papers in Economics and Statistics

2021-26

Glenn E. Dutcher, Regine Oexl, Dmitry Ryvkin, Tim Salmon

Competitive versus cooperative incentives in team production with heterogeneous agents

Abstract
A debate among practicing managers is whether to use cooperative or competitive incentives for team production. While competitive incentives may drive individual effort higher, they may also lead to less help and more sabotage; an issue exacerbated when team members' abilities are varied. Using a lab experiment, we examine how increasing competitive incentives affects performance as team composition changes. We find that competitive incentives generally underperform noncompetitive incentives and a larger bonus does not generate enough effort to compensate for a loss in help. Our results help understand better how to balance out individual versus team rewards and how firms could structure teams when employees have heterogeneous abilities.

ISSN 1993-4378 (Print)
ISSN 1993-6885 (Online)