Received: 27 September 2021 | Revised: 23 November 2023 | Accepted: 7 December 2023

DOI: 10.1111/jems.12573

ORIGINAL ARTICLE

Journal of Economics & Management Strategy WILEY

Do competitive bonuses ruin cooperation in heterogeneous teams?

E. Glenn Dutcher[1] | Regine Oexl[2] | Dmitry Ryvkin[3] | Timothy C. Salmon[4]

[1]Department of Economics, Ohio University, Athens, Ohio, USA

[2]Department of Economics, University of Innsbruck, Innsbruck, Austria

[3]School of Economics, Finance and Marketing, RMIT University, Melbourne, Victoria, Australia

[4]Department of Economics, Southern Methodist University, University Park, Texas, USA

Correspondence
E. Glenn Dutcher, Department of Economics, Ohio University, Athens, OH, USA.
Email: dutcherg@ohio.edu

Funding information
Austrian Science Fund, Grant/Award Numbers: P27912-G27, SFB F63

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, with unclear consequences overall, especially when team members' abilities differ. Using a lab experiment, we examine how increasing competitive incentives affects performance as team composition changes. We find that competitive incentives generally under-perform 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.

1 | INTRODUCTION

Many tasks are best handled by a team.[1] A team's success often relies on cooperation among team members, such as when one member helps another by sharing knowledge or takes on the part of a teammate's task (Lazear & 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 lead to free-riding. Alternatively, some believe competition among employees and appropriate rank-based rewards are the best way to motivate workers and drive the organization towards constant improvement.[3] Even though competitive mechanisms address free-riding concerns, they may diminish the willingness of individuals to help others and may even incentivize intrateam sabotage.[4] While the use of ranking mechanisms may currently be declining, which approach corporations should adopt is still debated. For example, Microsoft recently decided to scrap its rank-order system, but Yahoo! announced they were implementing one.[5] Given the relevant implications of these compensation issues, a more complete understanding of how they affect employee behavior is warranted. More generally, it is important to understand how to best address the dual incentive problem facing firms trying to incentivize cooperation while also incentivizing individual effort.

The dual incentive problem becomes even more complicated once the realistic assumption of worker heterogeneity is considered. Indeed, workers of varying abilities may respond to rank-based mechanisms differently, and their responses may also differ based on the composition of their team. This makes determining the effectiveness of payment schemes more difficult and also introduces a question of whether it is best to construct homogeneous or heterogeneous teams. The answer very likely depends on the nature of the task and the reward system.

This study assesses the validity of the opposing claims regarding the effect of competitive incentives on individual effort and helping behavior in a team production setting with heterogeneous agents, which has implications for the optimal structure of teams. Exploring these questions using observational data would be difficult because such data rarely contains information on effort and abilities and will omit information on behaviors involving help and sabotage. Further, the endogeneity of team construction and the design of the compensation mechanism would make it difficult to identify causal relationships. Therefore, we will investigate these questions through a controlled laboratory experiment guided by a theoretical model of the underlying incentives.

There are two schools of thought on how competitive incentives may affect output. Those who believe strongly in the importance of competitive incentives implicitly claim that such incentives lead to increases in individual effort that are more than offset by a reduction in help, leading to an increase in output. The other side claims the opposite: competitive incentives will ruin the willingness of coworkers to help each other and could even lead to sabotage, which will more than offset any (potential) gains in individual effort.[6] This viewpoint will usually come with an optimistic view of the ability of people to cooperate as a general human tendency.

The debates are not helped by the substantial prior evidence favoring the core arguments of both. For example, one of the most common results found in the contest literature is that competitive incentives drive individuals to exert substantially more effort than predicted in a standard model (see, e.g., a review by Dechenaux et al., 2015). These results lend credence to the use of competitive incentives in the field. On the other hand, there is also extensive literature showing that individuals are much better at cooperating than one would expect given the predictions from a standard model (see, e.g., Roth & Kagel, 1995). Because individuals contribute much more than expected in a public goods setting, teammates in an organization may also be able to solve cooperation problems and generate high output without the need for competitive incentives. However, these base studies on behavior in contests and public goods environments do not provide clear answers to the question of how 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 incentive mechanisms are present and agents vary in their ability.

A few prior studies examine more directly the effect of competition on cooperation. Buser and Dreber (2016) find that people tend to cooperate less after they have competed with each other, but similar results are found when the prize is allocated purely randomly rather than through competition. Drago and Garvey (1998) suggest that helping effort is reduced when incentives in promotion tournaments are strong. A similar finding is reported by Brown and Heywood (2009), based on a survey of finance industry employees. Hamilton et al. (2003) provide indirect evidence that high-ability workers were willing to help low-ability workers at a garment factory when the incentives were switched from individual- to team-based. In a similar study, Johnson and Salmon (2016) examine heterogeneity and sabotage, but sabotage occurs via a posttournament choice, not simultaneously with making the effort decision.

To the best of our knowledge, Danilov et al. (2019) is the only other paper analyzing help and sabotage in a team production setting involving a combination of cooperative and competitive incentives. They constructed a team production model using a variation of a Lazear-Rosen tournament with symmetric agents, and found a good agreement between observed behavior and their predictions in a carefully implemented experiment with a one-shot incentivized decision preceded by an extensive training phase. 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 competition and/or cooperation—and how those deviations develop over time in an organization-like setting with repeated interactions and feedback.[7] 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 to 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]

We present a theoretical model of decision-making for workers in a team where a combination of team-based and competitive incentives rewards output. Workers are heterogeneous in their productivity and can exert individual effort and 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. 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, orthogonally, increase the size of the competitive prize.

Our results largely support the comparative static predictions from the theory: when competitive incentives are added to a cooperative incentive framework, effort increases and helping behavior decreases. This is roughly in line with findings from Danilov et al. (2019). However, counter to Danilov et al. (2019), many of our findings are quite different from the theoretical point predictions, and, more importantly, we identify systematic behavioral deviations from the theory that mostly comport with the noncompetitive school of thought.

Specifically, we find that balanced teams with an equal number of agents of each ability level perform weakly better than any other combination of abilities, including homogeneous teams. Additionally, we find that the behavioral effects on total output do not support the competitive school as increasing the prize does not lead to more output than what was observed in the noncompetitive scheme and sometimes leads to less when moderate competitive incentives are used. Adding more high-ability types to the team also goes against the competitive school, given this leads to a decrease, not increase, in total output in excess of the corresponding theoretical prediction. Decomposing the individual effects of effort and help on output, relative to the noncompetitive scheme, medium-powered incentives lead to lower increases in effort (relative to predictions) and helping behavior is basically eliminated. Under high-powered competitive incentives, individual effort is actually below predicted levels; however, we also see that sabotage is not quite as high as predicted. The latter result indicates individuals' aversion to sabotage, leading to an illusion that helping behavior is higher than predicted under high-powered incentives. Finally, we find that increases in the number of high-ability agents on a team lead to a higher-than-predicted effort but lower-than-predicted help under competitive incentives. In sum, our results provide a strong case against using competitive incentives in teams, but they also do not support some of the most extreme claims made by the noncompetitive school.

These results have implications for the management of heterogeneous teams and team formation. That we do not find complete support or refutation of one school potentially explains why both schools still survive in modern corporate workplaces: There simply is not a one-size-fits-all recipe for how to balance competitive and noncompetitive incentives. Our results provide indications of where competitive incentives may be beneficial as well as where they might be problematic. Which approach is likely to be the most applicable will depend on what is important to a particular firm. If a firm is more reliant on individual effort, moderate and high-powered incentives do encourage higher effort and firms need not be too worried about the possibility of workers sabotaging each other, even when workers are competing for very high individual prizes. On the other hand, for a corporate environment that heavily relies on teammates helping each other, we find that competitive incentives could be disastrous. Moderate competitive incentives may not increase effort much and may virtually eliminate helping behavior. Thus, a firm that places a high value on helping behavior is likely best off relying on cooperative incentives. Finally, when possible, Balanced heterogeneous teams should be used.

The rest of the paper is organized as follows. Section 2 presents the theoretical model while Section 3 presents our experimental design and hypotheses. 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 the ability of team members and for team members to choose to devote their energy towards individual effort, helping another teammate or sabotaging another teammate. 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 that 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 (Danilov et al., 2019; Garvey & Swan, 1992) and heterogeneous (Gürtler & Münster, 2013; Kräkel, 2005) agents.[9]

Consider a team consisting of $n \geq 2$ risk-neutral agents indexed by $i \in \{1, ..., n\}$ and characterized by (possibly heterogeneous) ability parameters $\gamma_i > 0$. Each agent i chooses effort $x_i \in \mathrm{IR}_+$ 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 activity $k_{ij} \in \mathrm{IR}$, 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 in $|k_{ij}|$. The output of agent i is given by

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

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 a higher-ability type is more effective, regardless of the ability of the helper. We acknowledge that helping behavior can be modeled in a few different ways. We investigate a setting in which the effectiveness of helping behavior is due to the comparative advantage of the workers who perform tasks of varying difficulty. There are certainly settings, such as a production line with all workers performing the same task, where our model may not directly apply. In those cases, ability defines how efficiently effort is converted to output for this homogeneous task, and the effect of helping behavior is tied to the ability of the helper. However, the settings where we believe the two schools of thought are typically relevant involve instances where ability defines the level of difficulty of a task the worker can perform. As a simple example, let there be two tasks: an easy task 1 and a difficult task 2. Suppose a low-ability worker (an L type) has skills to perform task 1 only, and a high-ability worker (an H type) can perform both tasks. The workers are paid based on completed projects, where a completed project may require a mix of these tasks. If the projects are assigned according to the skills of the employees, H types will be assigned projects that are more difficult (e.g., require both tasks 1 and 2), while L types will be assigned only projects requiring task 1. Thus, helping behavior from an L type to an H type will take the form of the L type performing task 1 and freeing up the H type to perform task 2. Likewise, help from an H type to an L type will have the H type performing the easier task 1. The comparative advantage structure then implies that the effectiveness of helping behavior is tied to the ability of the one being helped. 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 the productivity of the firm.

2.1 | Team incentives without competition

Every team member receives a piece rate $r > 0$ 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 $|k_{ij}|$, as required. This gives agent i's utility (payoff) in the form

$$\pi_i = r\sum_{i=1}^{n} \gamma_i \max\left\{0, x_i + \sum_{j \neq i} k_{ji}\right\} - \frac{x_i^2}{2\alpha} - \sum_{j \neq i} \frac{k_{ij}^2}{2\alpha},$$

where the first sum represents the sum of agents' outputs, y_i. Maximizing π_i with respect to x_i and k_{ij} (for each $j \neq i$), obtain the corresponding optimal values

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

The zero argument here refers to the absence of a bonus. These levels of effort and effort-modifying activities constitute the unique Nash equilibrium in dominant strategies.

2.2 | Team incentives with competition for a bonus

We will now introduce an intrateam 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 $\frac{y_i}{\sum_{j=1}^{n} y_j}$ (customarily, it is assumed that the probability is $\frac{1}{n}$ if $\sum_{j=1}^{n} y_j = 0$). In this setting, agent i's expected payoff function is

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

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

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

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

Equations (3) and (4) can be manipulated to obtain a closed-form solution. Expressing x_i from (3) and k_{ji} from (4), obtain individual outputs,

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

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

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

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, that is, it turns into sabotage. Note that, other things being equal, this turning point is lower for higher-ability agents.

As expected, the equilibrium is not socially optimal. Instead, the choices maximizing welfare (aggregate payoffs) are above the equilibrium predictions, since contributions yield a benefit to all four group members, while the cost is only incurred by the contributor. This is true for both effort and help.

3 | EXPERIMENT DESIGN AND HYPOTHESES

3.1 | General procedures

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 et al., 2014). Average payment per subject was €22.97.

Once all subjects were checked in and seated at computerized workstations, instructions were handed out and read out loud.[12] After all questions were answered, the experiment began. At the end of the experiment, we elicited basic demographics.

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. To fully understand the effects of team composition, we utilized a between-subject design, where subjects were assigned to one of the five potential team composition treatments—*HHHH*, *HHHL*, *HHLL*, *HLLL*, or *LLLL*, see Table 1. 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 eight-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 bonus system. To control for order effects in examining the different sizes of contest prizes, we varied the order of the prizes between sessions. In *order 1*, the low-powered incentive was introduced in Block 2 and the high powered in Block 3; in *order 0*, these were reversed with the high-powered incentive in Block 2 and the low-powered incentive in Block 3. Having sessions with both ordering possibilities for how incentives are added allows us to identify whether any results we observe are due to the order in which incentives are implemented.[13] New instructions were handed out, and read out loud, before 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' efforts. 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 their total output which is a combination of their choice as well as the net modifications from others helping and sabotaging them. They did not see exactly what choices others made to generate that modification, but they could back out the net effect. They saw the average help/sabotage in their team directed at all members and the total team output. They were also

TABLE 1 Overview of treatments and demographic variables.

	LLLL	*HLLL*	*HHLL*	*HHHL*	*HHHH*
Number of individuals	48	48	48	72	48
Number of teams	12	12	12	18	12
Number of rounds	24	24	24	24	24
Percentage male	52.08	45.83	47.92	44.44	39.58
Average age	23.56	21.81	23.85	22.01	23.29
Average risk aversion	43.00	41.58	39.48	45.25	42.58
Average competitiveness	5.42	6.27	5.71	5.22	5.67

Note: Subjects played eight rounds for each value of V. “Risk aversion” runs from 0 to 100, with higher numbers indicating less risk aversion. “Competitive” answers the question “On a scale from 1 to 10, how much do you like being in competition with others?” and runs from 1 to 10, 10 being “very much.” A χ^2 test shows no difference in “percentage male” across treatments, and Kruskal–Wallis tests show no difference across the treatments for “risk aversion” and “competitiveness” ($p > .32$ for both variables, 12(18) observations per treatment), while “age” does differ between the treatments ($p < .01$).

reminded of the cost of each decision, the 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 & 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 that are intended to allow us to do just that. First, we 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 2; and based on these values, we can construct a basic set of testable predictions.

TABLE 2 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

Hypothesis 1 (Comparative statics from theory).

(A) As V increases, total output is unchanged. As the number of H types increases, total output rises.
(B) As V increases, effort increases. As the number of H types increases, effort is unchanged for $V = 0$, while it decreases when $V > 0$.
(C) As V increases, help decreases. As the number of H types increases, help is unchanged for $V = 0$, while it increases when $V > 0$.

It is predicted that total team output, Y, will not vary with the level of competitive incentives. Although this may seem unexpected, it is useful for our empirical exercise. This result obtains due to the offsetting impacts of V on help and effort. As V rises, effort rises but help falls. We have our parameters set up so that these changes offset each other leaving total output unchanged. This prediction allows for easily determining if effort rises faster than help declines or vice versa.

These comparative static predictions can be extracted out of a careful examination of Table 2. Regarding individual effort, we get the expected result that effort increases in the competitive incentives. We, however, get a somewhat counter-intuitive prediction that effort decreases with the number of H types. As the number of H types increases, so does the marginal value of helping others. This is because a unit of help increases the total output of those being helped by their productivity factor, which is, of course, higher for H types. Thus, if more help is directed towards other H types, it increases their individual output making them more competitive for the prize, with less effort from themselves. In turn, the marginal value of effort is lowered. Another source of the trade-off between effort and help is our model specification that separates the cost of own effort from the cost of helping, similar to that used in Danilov et al. (2019).

3.4 | Behavioral predictions and hypotheses

Section 3.3 provided the comparative statics predictions derived from a theory where agents maximize their monetary outcomes. Here we will explain how behavior could deviate from those predictions using the perspective of the competitive and noncompetitive schools of thought. We will present the hypotheses in the order we will examine the predictions in Section 4.

Should behavior differ from standard theory, our goal is to determine if the competitive or noncompetitive school of thought provides a better characterization of observed behavior. As an organizing structure, we will present our hypotheses from the perspective of the competitive school of thought.

An initial concern for an organization is sorting individuals into teams to maximize production. Given our setup, we can examine this issue by considering a group of eight workers (four of them H types and four L types) split into two four-person teams. There are three ways to create these teams, which we will refer to as *Homogeneous* (*HHHH* and *LLLL*), *Balanced heterogeneous* (*HHLL* and *HHLL*), or *Asymmetric heterogeneous* (*HHHL* and *HLLL*), where the latter two are made up of teams with heterogeneous types. Theoretically, group organization does not matter because the total output from each eight-person group is predicted to be 781.25 regardless of the group makeup (see Table 2). However, behaviorally, variation in the number of H types in each team can create differences in competitive pressure leading to variation in output. The competitive school of thought would argue that as the number of H types in a team increases, so does the competitive pressure.

It is intuitive that as L types are replaced with H types on a team, other H types may "rise to the challenge" by increasing their effort. This implies that increasing the proportion of lower-ability types on a team will be seen as a less competitive setting for the H types. The same logic can be applied to the L types, although the alternative viewpoint, in this case, is that they will be discouraged.[16] When considering how this will affect total output in different configurations of eight-person groups consisting of two four-person teams, it is useful to remember that H types' effort is converted into output at a higher rate, and thus a manager may be more interested in increasing the effort of the H types. Since there are three possible group configurations, we only need to consider two cases. First, let a manager transition from $HHHL + HLLL$ to $HHHH + LLLL$: As the H moves to the strong team, all other H's (and him/herself) feel more pressure, and so all will increase output more than output will decline in the weak team due to the departure of H. This effect is in line with the competitive school, and there may be an additional boost in effort because both teams become homogeneous, and hence L types can no longer be discouraged. Next, consider a transition from $HHLL + HHLL$ to $HHHL + HLLL$. The competitive school would say the boost in H's output in the (newly) stronger team will outweigh the loss in output in the (newly) weaker team. Thus, in both cases, the competitive school would predict that as teams become more homogeneous, the competitive pressure increases the overall output, and these groups would outperform more heterogeneous ones. The noncompetitive school of thought would then predict that mixed teams would maximize the benefits of cross-type helping.

Note that in our treatment where competitive incentives are not present ($V = 0$), the competitive school should not make a strong prediction about group dynamics, while the noncompetitive school should still predict that heterogeneous teams perform better. Because we are stating our hypothesis on group composition from the competitive school's perspective, we limit the prediction to the cases of $V > 0$.

Hypothesis 2 (Group composition). For $V > 0$, total output is higher in groups with homogeneous teams than in groups with heterogeneous teams.

In the following hypotheses, we focus on deviations from the theoretical predictions and how these deviations vary with our treatments rather than testing the simpler comparative statics derived from theory. This is because the theory is premised on agents maximizing monetary outcomes. If nonmonetary motives exist, as is argued by both schools of thought, testing for deviations from the theory is the best way to represent the nature of the predictions of the schools of thought and, in particular, to see if behavior is consistent with either of them. For instance, in the case of how output varies with V, the standard prediction is that output rises with V. Both schools may agree on the direction of this effect. What differentiates the two schools is their expectation of output rising more or less than the equilibrium prediction because of behavioral effects. Similarly, the noncompetitive school may find it hard to argue that output would decrease as the number of H types increases, given the substantial advantage they have in production. What the noncompetitive behavioral effect may indicate, however, is that output would not increase as much as predicted. We refer to these deviations as behavioral effects.

We now focus on how the two schools might classify behavioral effects if an organization is contemplating changes that make their workplace more competitive by either adding more H types or increasing individual prizes. It should be quite clear that the competitive school would predict that, holding team composition constant, an individual prize will lead to a behavioral increase in output, and this effect should strengthen as incentives increase (e.g., the size of V). The noncompetitive school would predict the opposite—that is, a behavioral decrease in output.

Likewise, the competitive school would argue that the competitiveness of the setting increases in the number of H types in the team, with the same prediction on the behavioral effects of this competitiveness. This leads to our next hypothesis.

Hypothesis 3 (Competitiveness and total output).

(A) The deviation of the observed total output from the equilibrium prediction, $Y - Y^*$, will be increasing in the value of V.
(B) The deviation of the observed total output from the equilibrium prediction $Y - Y^*$, will increase in the number of H types in the team.

Note that in formulating this hypothesis, we remain agnostic about the sign of the deviation, $Y - Y^*$. If it so happens that the deviation is negative, we should still be able to interpret an increasing, that is, a less negative deviation or a decreasing absolute deviation, as evidence in favor of the competitive school. This is because it will indicate that increasing competition mitigates whatever baseline behavioral effects suppressing output (relative to its equilibrium level) exist in the noncompetitive setting. The same caveat applies to the rest of the hypotheses in this section.

Turning to effort, because competitiveness is increasing in V and the number of H types, the competitive school would expect deviations in effort to rise as both V and the number of H types increase. The noncompetitive school is largely silent on the effects of competitiveness on effort.

Regarding helping behavior, the noncompetitive school would take the stance that increased competition will lead individuals to withhold help or increase sabotage. Thus, the noncompetitive school would predict that deviations in helping behavior would decrease as V rises and the number of H types increases. The competitive school is silent on helping behavior.

Hypothesis 4 (Competitiveness and effort).

(A) The deviation of the observed effort from the equilibrium prediction, $x_{i,V} - x^*_{i,V}$, will be increasing in the value of V, for both H and L types.
(B) The deviation of the observed effort from the equilibrium prediction, $x_{i,V} - x^*_{i,V}$, will be increasing in the number of H types in the team, for both H and L types.

Hypothesis 5 (Competitiveness and help).

(A) The deviation of the observed amount of help from the equilibrium prediction, $k_{ij,V} - k^*_{ij,V}$, will be decreasing in the value of V, for both H and L types.

(B) The deviation of the observed amount of help from the equilibrium prediction, $k_{ij,V} - k^*_{ij,V}$, will be decreasing in the number of H types in the team, for both H and L types.

The final two hypotheses can be classified as a test of whether individuals view competition as friendly or fierce. The competitive school would hope that the competition yields a mindset of "may the best (wo)man win," where increases in effort are the primary tool used. Because there is still a monetary incentive to help, helping behavior would still be present. The noncompetitive school takes a more jaded view that the competitive setting brings out the worst in people who will refuse to directly help their competitors or, worse, sabotage them to get ahead.

4 | RESULTS

4.1 | Data and analysis overview

We begin with an overview of the data. Table 3 presents the equilibrium predictions and observed averages of each key variable in each treatment and reward scheme.[17] The table also presents the results of two-sided Wilcoxon signed-rank tests, using one team as a unit of observation, comparing each of the variables to its theoretical prediction.

By examining these summary statistics, we can observe a few immediate patterns in the data. 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/sabotaging behavior we observe that help is generally less than the predicted amount at low or moderate incentives but above the predicted level at high incentives. This latter finding is not due to subjects providing much help in the high-incentive treatment, but rather, as we will see, it is due to subjects sabotaging much less than predicted. It is also useful to note that 100% of the subjects help at least one group member each period, independently on the size of V. Finally, the percentage of subjects who sabotage at least one member of their group is roughly constant within one block, and is higher for $V = 100$ than for $V = 0$, and higher for $V = 500$ than for $V = 100$.

Subjects' choices of effort and help lead to trade-offs in total output such that total output is higher than predicted with no competitive incentives and lower than predicted for intermediate and high competitive incentives. The composition of the team impacts these effects: As the number of H types goes up, effort increases, but help decreases. We will provide more detailed tests of these effects in the following sections.

4.2 | The level effects

We begin our formal tests by examining Hypothesis 1 regarding the base comparative statics predicted by the theory. We summarize our findings in our first result.

Result 1 (Theoretical predictions).

 Journal of Economics & Management Strategy –WILEY–

TABLE 3 Equilibrium predictions (NE) and observed averages (Obs.).

	LLLL			*HLLL*			*HHLL*			*HHHL*			*HHHH*		
	NE		**Obs.**	**NE**		**Obs.**	**NE**		**Obs.**	**NE**		**Obs.**	**NE**		**Obs.**
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
V = 100															
x_L	28.52		33.28	22.01		31.26	18.77	<	31.31**	16.87	<	39.69**			
x_H				31.78		35.74	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		409.50	625.00	>	483.21***
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

Note: Stars indicate significance levels of two-sided sign tests based on team averages, comparing the observed averages to the equilibrium predictions. To ease understanding of the direction of the deviations from the equilibrium prediction, we have added > and < signs when differences are significant. As can be seen from Table 1, we have 12 teams in all treatments, except for treatment *HHHL*, in which we have 18 teams. Standard errors (clustered by team) are omitted here for ease of exposition; they are provided in Table A1 in Appendix A.

***$p < .01$; **$p < .05$; *$p < .1$.

(A) *Individual effort is increasing in V; it is also increasing in the number of H types on a team in the absence of competitive incentives and for moderate incentives, but not for high incentives.*

(B) *Help is decreasing in V and increasing with the number of H types on a team.*

(C) *Total team output is not significantly different comparing $V = 0$ and 500, but it is lower for $V = 100$ than for $V = 0$ for teams with at least two H types. Total output increases in the number of H types for each V.*

TABLE 4 The level effects.

	(1) $x_{i,V}$	(2) $\overline{k_{ij,V}}$	(3) Y
$V = 100$	11.96***	−4.39***	−21.59
	(2.90)	(0.98)	(24.44)
$V = 500$	34.00***	−9.28***	3.59
	(4.17)	(1.31)	(24.83)
Number of H types	2.03***	0.88**	96.54***
	(0.72)	(0.37)	(7.73)
$(V = 100) \times$ (number of H types)	−0.05	0.12	−9.74
	(0.92)	(0.32)	(7.92)
$(V = 500) \times$ (number of H types)	−1.33	0.49	−4.46
	(1.50)	(0.49)	(12.68)
Constant	14.74***	12.68***	251.98***
	(5.24)	(4.28)	(22.64)
Controls	Yes	Yes	Yes
Observations	6336	6336	1584
Number of individuals	264	264	264
Number of teams	66	66	66

Note: Robust standard errors clustered by team in parentheses. "Number of H types" runs from 0 to 4. Controls include "Gender," "Risk Aversion," "Competitiveness," "Order," and "Period" in columns (1) and (2), and only the latter two in column (3).

***$p < .01$; **$p < .05$; *$p < .1$.

Support for these claims is based on Table 4 as well as postestimation Wald tests. This table reports regressions examining the three main variables using random effects GLS with standard errors clustered at the team level. All regressions include a set of standard controls we will use in other regressions as well. Our explanatory variables are binary variables for $V = 100$ and 500 (with $V = 0$ as the reference group), a variable that accounts for the number of H types in the team (ranging from 0 to 4), and their interactions. Because we will use the same right-hand side variables in our regressions throughout the paper, we will also rely on the same set of Wald tests. For ease of exposition, Table 5 numbers the Wald tests. When we refer to the use of a Wald test, we will refer to its number from this Table.

The basic theoretical predictions are that individual effort increases with V, is independent of the number of H types on the team for $V = 0$, and decreases with the number of H types for $V > 0$. The regression in column (1) of Table 4 shows us that the first prediction holds for *LLLL* groups ($p < .01$ for the comparison of coefficients on $V = 100$ and 500). We conducted Wald tests 1–3 for other team compositions. In all cases, $p < .01$. For the effect of the number of H types, column (1) results imply that effort is increasing when $V = 0$. For $V > 0$, we used Wald tests 4 and 5. The first test produces $p = .014$ and the second $p = .65$, implying that effort increases in H types for $V = 100$ but remains constant for $V = 500$. These results are also illustrated in Figure 1.

The results for helping behavior are based on the regression in column (2) of Table 4. As predicted, help declines as V rises for all team compositions ($p < .01$ from Wald tests 1 to 3). Help also increases in the number of H types, which is predicted for $V > 0$ but not for $V = 0$ ($p < .02$ from Wald tests 4 and 5). These results for average help are also illustrated in Figure 2.

Finally, column (3) of Table 4 examines total team output. The prediction is that team output is constant in V, for each team composition and increases in the number of H types for any V. The insignificant coefficients on $V = 100$ and 500 show that total team output does not vary with the prize for *LLLL* teams. However, using Wald test 1, we find that team output is higher for $V = 100$ than for $V = 0$ for all teams with at least two H types ($p = .12$ for $n_H = 1$ and $p < .03$ for $n_H = 2, 3, 4$). In contrast, Wald test 2 shows there are no significant differences in output between $V = 500$ and 0 for any number of H types ($p > .73$ for $n_H = 1, 2, 3, 4$). Wald tests 4 and 5 also confirm that output increases in the number of H types for each V ($p < .01$ for $V = 100$ and 500). The results for average output are illustrated in Figure 3.

TABLE 5 Wald tests used.

Wald test number	Wald test
1	$\beta_{V=100} + \beta_{(V=100)\times(\text{number of } H \text{ types})} \times n_H = 0$
2	$\beta_{V=500} + \beta_{(V=500)\times(\text{number of } H \text{ types})} \times n_H = 0$
3	$\beta_{V=100} + \beta_{(V=100)\times(\text{number of } H \text{ types})} \times n_H =$
	$\beta_{V=500} + \beta_{(V=500)\times(\text{number of } H \text{ types})} \times n_H$
4	$\beta_{\text{number of } H \text{ types}} + \beta_{(V=100)\times(\text{number of } H \text{ types})} = 0$
5	$\beta_{\text{number of } H \text{ types}} + \beta_{(V=500)\times(\text{number of } H \text{ types})} = 0$

Note: $n_H = 0, 1, 2, 3, 4$ is the number of H types in the team.

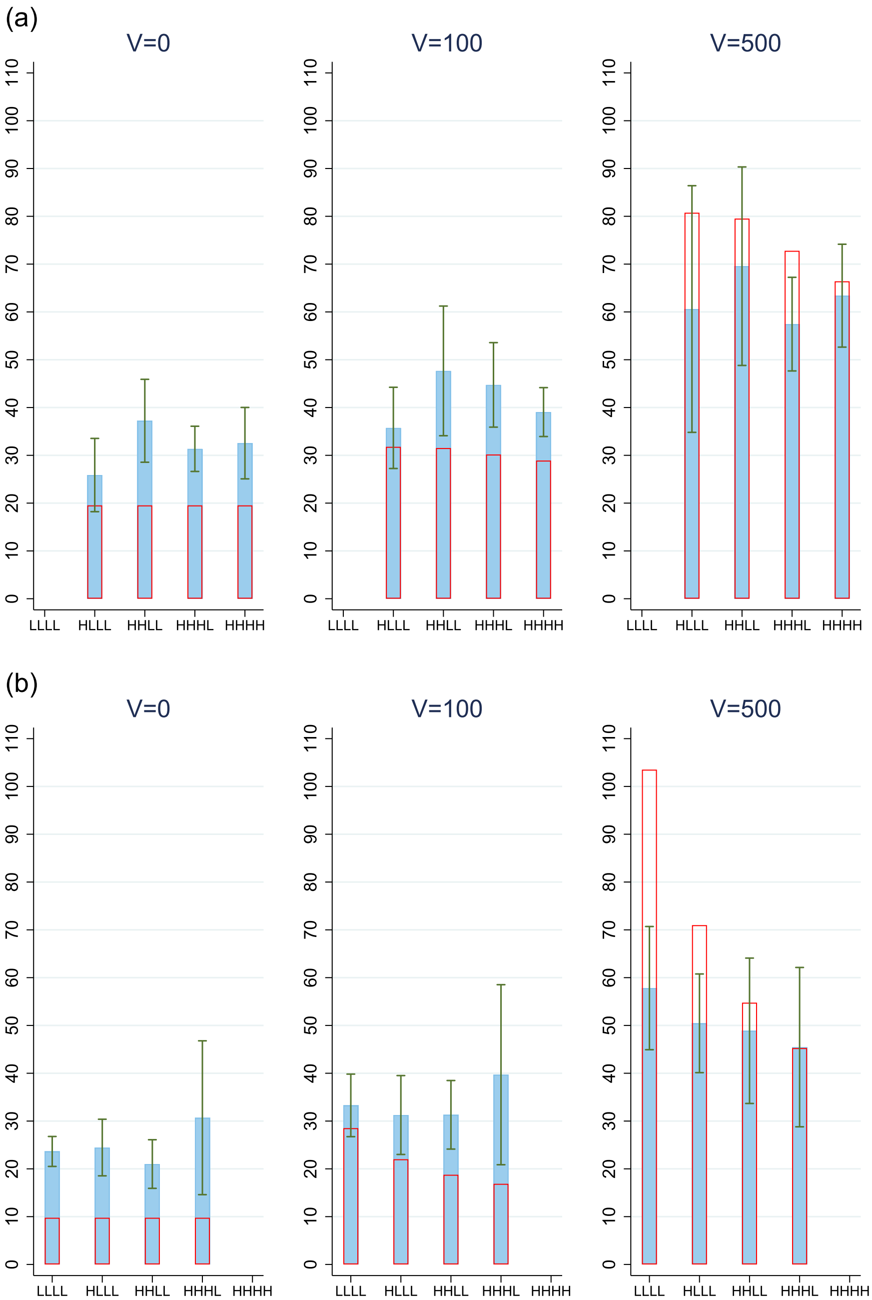

FIGURE 1 Average effort. Predicted levels are shown as empty boxes and actual observed levels as filled boxes, with the standard error bars of the 95% confidence interval; we treat each team in one block as one observation. If teams have heterogeneous types, we have the same team once in the graphs for H types and once in the graphs for the L types. (a) H types and (b) L types. [Color figure can be viewed at wileyonlinelibrary.com]

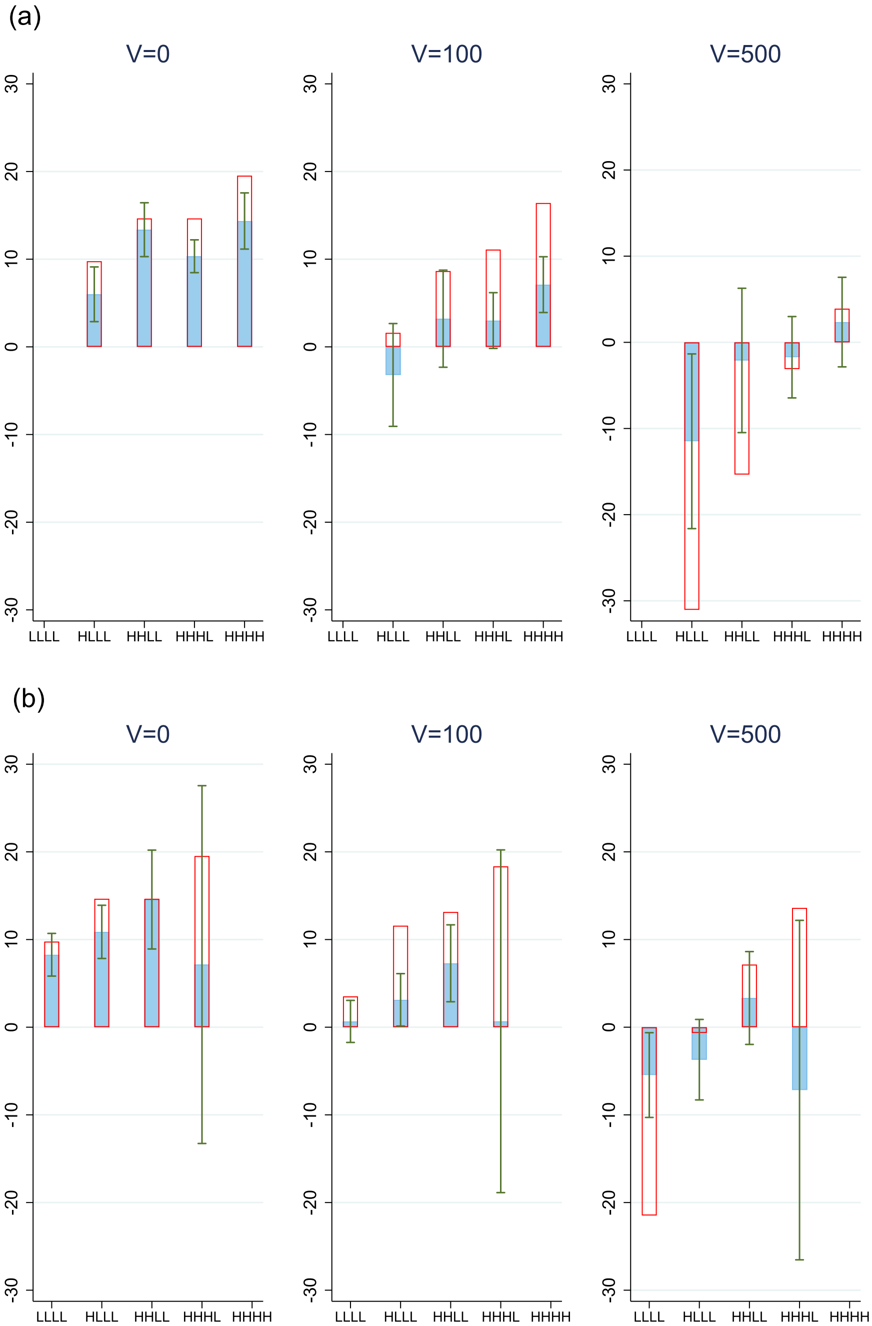

FIGURE 2 Average help. Predicted levels are shown as empty boxes and actual observed levels as filled boxes, with the standard error bars of the 95% confidence interval; we treat each team in one block as one observation. If teams have heterogeneous types, we have the same team once in the graphs for H types and once in the graphs for the L types. (a) H types and (b) L types. [Color figure can be viewed at wileyonlinelibrary.com]

4.3 | Organization of employees into teams

Having established some deviations from the theoretical predictions, we now test if these deviations can be organized by behavioral notions from the two schools of thought. We begin with a practical application.

Of particular concern to a manager is how she may divide a fixed set of employees into teams when they vary in ability. 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. This implies that the manager can organize her employees into two teams that are Homogeneous (*HHHH* and *LLLL*), Balanced heterogeneous (*HHLL* and *HHLL*), or Asymmetric heterogeneous (*HHHL* and *HLLL*).[18] As noted above, in our model total eight-person group output is predicted to be constant for any group composition and any value of V. This allows us to easily identify deviations from the theory. As Hypothesis 2 points out, should behavioral deviations occur, the competitive school of thought predicts homogeneous teams to produce the highest level of output for $V > 0$, and the noncompetitive school may expect teams

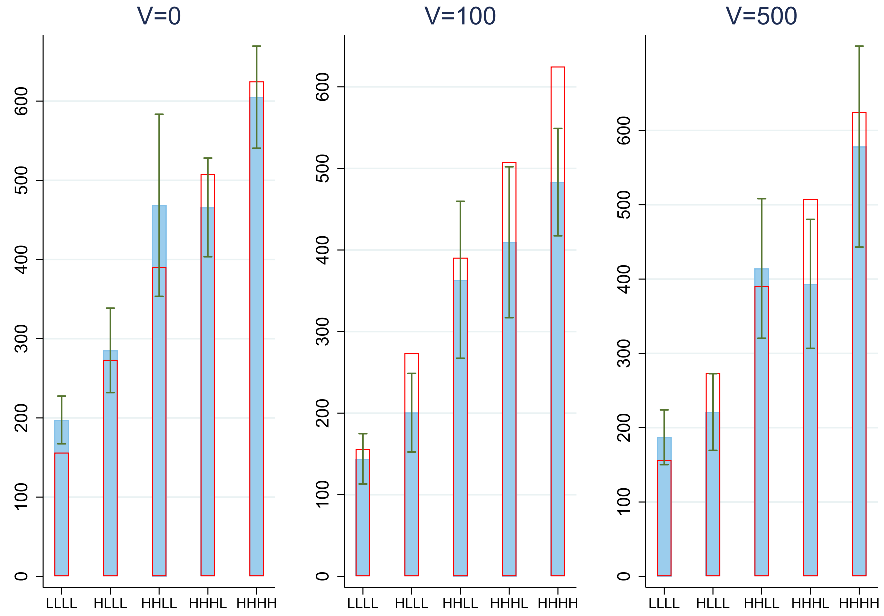

FIGURE 3 Average team output, Y. Predicted levels are shown as empty boxes and average observed levels as filled boxes, with the error bars representing the 95% confidence intervals. We treat each team in one block as one observation. [Color figure can be viewed at wileyonlinelibrary.com]

that are mixed to perform better. Even though our hypothesis stated from the perspective of the competitive school of thought focuses on $V > 0$, we also include analysis when $V = 0$ to better understand how the group composition may affect behavior absent a competitive incentive. Our result shows that we do not find support for Hypothesis 2.

> **Result 2** (Group composition). *For each value of V, output in the Balanced heterogeneous group is higher than in the Asymmetric heterogeneous group, and is the same or higher than that of the Homogeneous group. The Asymmetric heterogeneous group leads to the same or lower output than the Homogeneous group.*

Support for this result can be seen in Figure 4, which shows the predicted and observed total group output for each kind of group composition, and Table 6, which reports the corresponding averages with robust standard errors. From the figure, there appears to be a difference between the Balanced group and the others. By running a number of t tests comparing the total group output between the three group compositions, we can detect these differences are statistically significant for the different values of V. We find that the Balanced heterogeneous group's output is different from the Asymmetric heterogeneous group's output for each level of V ($V = 0$, $p < .01$; $V = 100$, $p = .05$; $V = 500$, $p = .01$), but when compared with the Homogeneous group, the difference is only (marginally) significant for $V = 0$ ($V = 0$, $p = .095$; $V = 100$, $p = .15$; $V = 500$, $p = .47$). The Asymmetric heterogeneous group produces lower output than the Homogeneous group only when $V = 500$ ($V = 0$, $p = .30$; $V = 100$, $p = .77$; $V = 500$, $p = .06$).

More generally, a composition with Balanced heterogeneous teams leads to better overall in-team interactions, allowing incentives to encourage reasonable effort while not impairing helping behavior as much as the other configurations. The effect observed when $V = 0$ supports this claim. With a minor exception, we also consistently find that the Homogeneous group composition does not generate output that is substantially higher than others when $V > 0$, implying the competitive school of thought is not supported. Because of the large variance, we do not find statistical differences between Homogeneous groups and Balanced heterogeneous groups for $V > 0$, even though there are large differences in the average output. This provides further support that we need a more careful examination of deviations from theory and the underlying behavioral mechanisms that may lead to these deviations.

4.4 | Deviations from predictions

Motivated by Figures 1–3, we now take a closer look at the deviations from theory.[19] The following regressions explicitly use the deviations between the observed data and the theoretical predictions to more clearly determine if the deviations are in line with our behavioral hypotheses. We begin our analysis by examining how team composition and

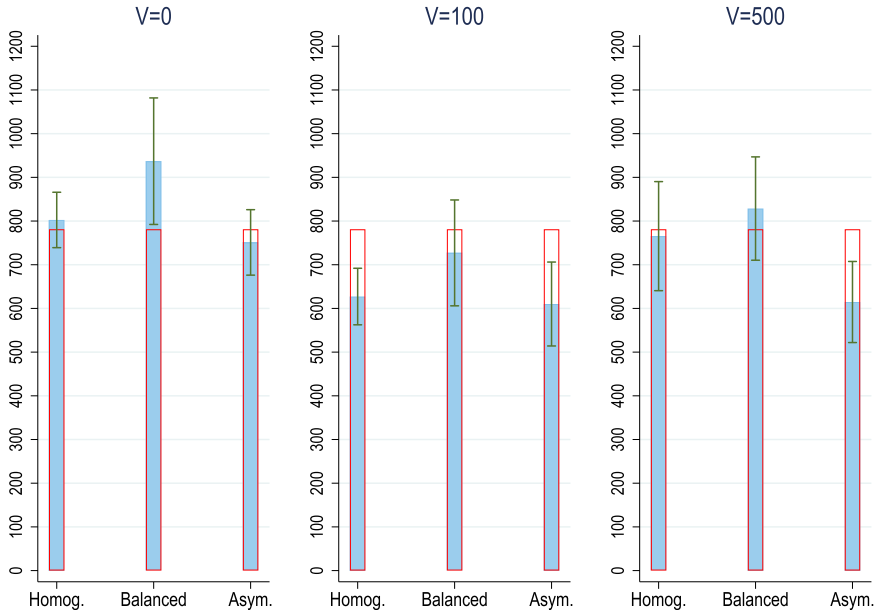

FIGURE 4 Average group output for Homogeneous, Balanced heterogeneous, or Asymmetric heterogeneous group compositions. Predicted levels are shown as empty boxes and average observed levels as filled boxes, with the error bars representing 95% confidence intervals. We treat each team in one block as one independent observation. Standard errors for eight-person groups are calculated by combining standard errors for the corresponding four-person teams. [Color figure can be viewed at wileyonlinelibrary.com]

TABLE 6 Group output for Homogeneous, Balanced heterogeneous, and Asymmetric heterogeneous team composition.

	$V = 0$	$V = 100$	$V = 500$
Homogeneous	802.50	627.21	765.43
	(32.34)	(33.02)	(63.68)
Asymmetric heterogeneous	751.09	610.08	614.67
	(38.22)	(48.99)	(47.34)
Balanced heterogeneous	936.96	727.06	828.58
	(73.87)	(61.81)	(60.32)

Note: The Balanced heterogeneous group includes teams *HHLL*; the Asymmetric heterogeneous group includes teams *HHHL* and *HLLL*, and the Homogeneous group includes teams *HHHH* and *LLLL*.

the size of the prize affect total output. Following this, we test how competitiveness affects the individual components of output—effort, and help.

As a reminder, the competitive school of thought argues that as an incentive scheme becomes more competitive, it will generate larger positive behavioral deviations from theory in output and effort, while the noncompetitive school predicts less help and more sabotage compared with theory. In this section, we study these deviations in detail and test Hypotheses 3–5. As discussed in Section 3.4, we use the deviations of the observed values from the equilibrium as our dependent variables, which may be positive or negative. Competitiveness measures are the size of the prize and the number of H types in the group. We interpret a result as supporting the competitive school if the coefficient estimate on a competitiveness variable is positive in the case of effort or output.[20]

4.4.1 | Competitiveness and output

While Result 1 highlights that the basic comparative statics of the theory hold, Result 2 indicates that the team composition may affect total output. According to the simple pairwise comparisons seen in Table 3 (variable Y), when a competitive prize is present, team output is generally equal to or below equilibrium predictions, while when there is no prize, it is equal to or above equilibrium. This provides our second piece of evidence against the competitive school of thought.

Result 3 (Competitiveness and total output).

TABLE 7 Deviations in team output.

	(1) Dependent variable: $Y - Y^*$
$V = 100$	−21.59
	(24.44)
$V = 500$	3.59
	(24.83)
Number of H types	−20.64***
	(7.73)
$(V = 100) \times$ (Number of H types)	−9.74
	(7.92)
$(V = 500) \times$ (Number of H types)	−4.46
	(12.68)
Constant	95.73***
	(22.64)
Controls	Yes
Observations	1584
Number of teams	66

Note: Robust standard errors clustered on team level in parentheses. "Number of H types" runs from 0 to 4. Controls include "Order" and "Period."
***$p < .01$; **$p < .05$; *$p < .1$.

(A) Size of V: *For total output, the deviations are the same or larger for* $V = 0$ *than for* $V = 100$, *and there is no difference between* $V = 0$ *and* 500. *The deviation is the same or larger for* $V = 500$ *than for* $V = 100$.
(B) Number of H *types: For total output, the deviation is decreasing in the number of* H *types in the team.*

Support for this result is based on Table 7 and postestimation Wald tests. The table displays the results of a random effects GLS regression, where the dependent variable is the deviation from the actual to the predicted total team output in every round. The explanatory variables and additional controls are the same as in column (3) of Table 4.

The statistically insignificant estimate of the coefficient on $V = 100$ in Table 7 implies that deviations from predictions in this setting are not significantly different from the deviation when $V = 0$ in *LLLL* teams. There is also no significant effect when $V = 500$. A postestimation Wald test also shows that the two coefficients are not significantly different from each other ($p = .14$) in *LLLL* teams. We conducted Wald tests 1–3 for other team compositions. For $V = 100$, $p = .12$ for *HLLL* teams; for all other team compositions, $p < .03$. For $V = 500$, $p > .73$ in all cases. This implies that when the prize is low ($V = 100$), and the team is comprised of more than one H-agent (i.e., for *HHLL*, *HHHL*, and *HHHH* teams), deviations from predictions are more negative than relative to the baseline. On the other hand, a large prize ($V = 500$) does not lead to larger deviations than the baseline.

The statistically significant negative estimate on the number of H types shows that deviations from theory decrease in the number of H types on a team for $V = 0$. For $V > 0$, we again used Wald tests 4 and 5. All tests produce $p < .07$, implying that deviations in output are even more negative as the number of H types increases for both values of V.

In sum, these results provide very little support for the competitive school of thought and our Hypotheses 3(A) and 3(B).

4.4.2 | Competitiveness: Effort and help

Even though the analysis up to this point provides an overview of how aggregate behavior may depend on competitiveness, we have yet to address the core behavioral assumptions of the two schools where the competitive school focuses on deviations in effort while the noncompetitive school focuses on deviations in help/sabotage.

To investigate these behavioral effects, we examine how the competitiveness (i.e., increases in V and increases in the number of H types on the team) affects deviations from the theoretical predictions, now focusing separately on 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.

Effort

We begin with a summary result, largely going against Hypothesis 4.

Result 4 (Competitiveness and effort).

(A) Size of V: *For effort, deviations are becoming more negative in V when going from $V = 100$ to 500 or from $V = 0$ to 500. This holds for both types: H and L.*
(B) Number of H types: *For effort, increasing the number of H types in the team mitigates the effect of V.*

This result can already be seen from panels (a) and (b) of Figure 1. To check if these deviations align with our hypotheses, Table 8 reports the results of random effects GLS regressions with standard errors clustered at the team level. The dependent variable is the deviation of an individual's observed effort choice to the theoretically predicted effort in a given period. The controls are the same as those used in column (1) of Table 4. Column (1) reports the results pooling the L and H types. To understand if results differ by type, columns (2) and (3) report the results separately for the L and H types.

From the regressions, deviations in effort are shown to be affected by the size of the prize. The regressions in Table 8 indicate that deviations in effort are lower when $V = 500$ relative to $V = 0$ for *LLLL* teams. Likewise, postestimation Wald tests indicate deviations are more negative when $V = 500$ relative to $V = 100$ ($p < .01$ in all three specifications holding group composition constant).

TABLE 8 Deviation of effort from equilibrium predictions.

	(1) **H and L** **Dependent variable: $x_{i,V} - x^*_{i,V}$**	**(2)** **L**	**(3)** **H**
$V = 100$	−4.49	−4.80	1.45
	(2.82)	(3.33)	(5.82)
$V = 500$	−47.95***	−51.50***	−35.74***
	(4.41)	(5.05)	(10.64)
Number of H types	−0.37	1.23	0.52
	(0.72)	(1.83)	(1.75)
$(V = 100) \times$ (Number of H types)	2.08**	4.03***	−0.36
	(0.94)	(1.31)	(1.64)
$(V = 500) \times$ (Number of H types)	9.33***	15.30***	4.42
	(1.60)	(2.50)	(3.40)
Constant	5.68	3.74	2.87
	(5.57)	(7.58)	(8.49)
Controls	Yes	Yes	Yes
Observations	6336	3024	3312
Number of individuals	264	126	138
Number of teams	66	54	54

Note: Robust standard errors clustered on team level in parentheses. "Number of H types" runs from 0 to 4. Controls include "Order," "Period," "Gender," "Risk Aversion," and "Competitiveness."

***$p < .01$; **$p < .05$; *$p < .1$.

We again conducted Wald tests 1–3 for other team compositions. When considering the pooled sample, $p > .15$ for $V = 100$, and $p < .01$ for $V = 500$, for all team compositions that contain at least one H type. The deviation is more negative when going from $V = 100$ to 500 ($p < .01$ for all tests in the pooled sample).

We also observe no effect on the number of H types for $V = 0$. However, from the first column, we see a positive interaction between the size of the prize and the number of H types. Columns (2) and (3) further decompose this effect and show that it is due to changes in the behavior of L types. When a prize is present, L types respond by increasing their effort as the number of H types increases, and this response is stronger when $V = 500$. No interaction effects are found for the H types.

For $V > 0$, we used Wald tests 4 and 5. Both tests produce $p < .05$, implying that deviations in effort increase in the number of H types for $V > 0$ relative to the baseline.

As before, the overall conclusion from our examination of the deviations in 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. The one exception is the L types' response to an increase in the number of H types, as their response to the number of H types is in line with the competitive school. However, examining panel (b) in Figure 1 makes it clear that this finding is due to predicted effort declining, while observed effort remains about the same (the effect is particularly pronounced in the $V = 500$ case). This suggests that the L types may simply be nonresponsive to their team's configuration.

Help and sabotage

Having established how competitive pressure alters effort provision, we now analyze how these same pressures affect help and sabotage behavior, as outlined in Hypothesis 5 that predicts less help and more sabotage with increased competitiveness. This final result indicates mixed support for the competitive school.

Result 5 (Competitiveness and help).

(A) Size of V*: For help/sabotage, the deviations are larger for* $V = 500$ *than for* $V = 100$ *or* 0. *This holds for both types*.

(B) *Number of H types*: Deviations in help decline as the number of H types increase. This effect is strongest when $V = 500$.

From Figure 2 it is immediately clear that deviations from equilibrium are always present and depend upon the predicted sign of help. When help is predicted to be positive, the observed help is weakly lower. When help is predicted to be negative (i.e., sabotage is predicted), the observed help is weakly higher. We will return to this difference between help and sabotage behavior, but first, we will establish results concerning the competitiveness of the setting and deviations in help.

Table 9 displays the results of random effects GLS regressions. The dependent variable is the deviation of average actual help from predicted help, where a negative deviation implies less help than predicted (or more sabotage than predicted). The right-hand-side variables are the same as those used for analyzing deviations in effort.

The results indicate that when the prize increases from $V = 0$ to 100, deviations in help decrease slightly, but the effect is not statistically significant for the *LLLL* teams (see the regressions in Table 4). However, when the prize increases from $V = 0$ to 500, deviations in help increase. Likewise, Wald tests indicate that when the prize increases from $V = 100$ to 500, deviations increase ($p < .01$).

We again use Wald tests 1–3 for other team compositions. When $V = 500$, $p < .06$ for all team compositions, implying deviations are greater than the baseline. However, when $V = 100$, the effect is only significant at the 5% level for *HHHL* and *HHHH*, while for *HLLL* and *HHLL* $p > .12$. This indicates that, relative to the baseline, the negative deviations when $V = 100$ is only observed when the team is mostly or wholly high ability. The difference between $V = 100$ and 500 is significant for all team compositions ($p < .001$ in all cases). This implies that deviations are greater when $V = 500$ than when $V = 100$, regardless of the makeup of the team.

As noted previously, these results are mainly driven by the fact that a higher prize is predicted to lead to substantial sabotage, but individuals refrain from it. This is especially relevant in the $V = 500$ case for the H types where the highest level of sabotage is predicted, compare Figure 2a for the *HLLL* and *HHLL* teams.

As the number of H types increases, deviations in help decrease for $V = 0$, and, as with effort, the interaction effects are also impactful. We use Wald tests 4 and 5 as before to assess the effect of the number of H types for different values

TABLE 9 Deviations of help and sabotage from equilibrium predictions.

	(1) *H* and *L*	(2) *L*	(3) *H*
	Dependent variable: $\overline{k_{ij,V} - k^*_{ij,V}}$		
$V = 100$	−0.36	0.31	−0.26
	(1.35)	(1.62)	(3.12)
$V = 500$	14.94***	16.75***	26.78***
	(2.00)	(2.24)	(5.06)
Number of *H* types	−1.30***	−2.88	−1.40*
	(0.49)	(2.27)	(0.74)
$(V = 100) \times$ (Number of *H* types)	−0.56	−1.57**	−0.48
	(0.44)	(0.71)	(0.86)
$(V = 500) \times$ (Number of *H* types)	−2.91***	−8.82***	−5.46***
	(0.73)	(1.24)	(1.47)
Constant	6.58	12.43	1.95
	(6.05)	(10.07)	(4.20)
Controls	Yes	Yes	Yes
Observations	6336	3024	3312
Number of individuals	264	126	138
Number of teams	66	54	54

Note: Robust standard errors clustered on team level in parentheses. "Number of *H* types" runs from 0 to 4. Controls include "Order," "Period," "Gender," "Risk Aversion," and "Competitiveness."

***$p < .01$; **$p < .05$; *$p < .1$.

of $V > 0$. The deviations in help when $V = 100$ decrease as the number of *H* types increases, but this holds only for *L* types. The strongest effect is observed when $V = 500$ ($p < .01$ from postestimation Wald tests for comparisons between the interaction when $V = 100$ and 500 in all three specifications), which is present for both types. This implies that the deviations are more positive in this prize structure than in others.

Overall, the results mostly support the idea that competitiveness leads to a behavioral "antihelp" 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 $V = 500$, in the *HLLL* setting for the *H* types and in the *LLLL* setting for the *L* types. In all of these cases, individuals are largely refraining from sabotaging each other, compare Figure 2. Thus, even though competitiveness may lead to lower levels of help, the main drawback from competitiveness, that is, sabotage, may also be more limited than predicted, thus blunting the worst of the impacts predicted by the noncompetitive school of thought.

One possible implication is that this aversion to sabotage leads to a structural change in behavior that needs to be controlled for. To investigate this in more detail, we have rerun the above regression for help, controlling for instances when help is predicted to turn into sabotage (see Table 10); the dummy variable "Sabotage" is a binary variable equal to one if predicted help is negative, and instead of looking at average help or sabotage behavior, we consider each help/sabotage decision at the individual level. Therefore, the number of observations is three times the number of observations in column (1) in Table 9. 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 team.

The positive coefficient estimate on the *Sabotage* variable shows that individuals do not sabotage as much as predicted, which drives most of the significant effect on $V = 500$ in Table 9. However, the negative sign on the interaction term of *Sabotage* with the number of *H* types indicates that team composition is an important component. Specifically, after controlling for the general aversion to sabotage, the competitiveness of the setting erodes this positive effect, which goes against the competitive school of thought.

TABLE 10 Controlling for instances when sabotage is predicted.

	(1) Dependent variable: $k_{ij,V} - k^*_{ij,V}$
$V = 100$	−0.29
	(1.33)
$V = 500$	−8.52**
	(4.05)
Number of H types	−1.30***
	(0.49)
($V = 100$) × (Number of H types)	−0.56
	(0.44)
($V = 500$) × (Number of H types)	3.23**
	(1.30)
Sabotage	24.40***
	(4.26)
Sabotage × (Number of H types)	−5.50***
	(1.50)
Constant	5.05
	(4.59)
Controls	Yes
Observations	19,008
Number of individuals	264
Number of teams	66

Note: Robust standard errors clustered on team level in parentheses. "Number of H types" runs from 0 to 4. "Sabotage" is a binary variable equal to one if predicted help is negative, and zero otherwise. Controls include "Order," "Period," "Gender," "Risk Aversion," and "Competitiveness."

***$p < .01$; **$p < .05$; *$p < .1$.

5 | CONCLUSIONS

In this paper, our goal was to empirically examine if competitive pressures in a team production setting led to consistent behavioral responses not predicted by standard theory. 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 takeaway 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 with 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.[21]

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. Paying additional bonuses is costly, and we find 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 as much as we change the composition of the team. This suggests that people are less sensitive to team composition than expected. 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 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: low-ability types still help at a decent level, 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 efforts than by changing their helping behavior. The lower willingness to modify others' efforts 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 team 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; in particular, about the shape of agents' costs. Again, more work is needed to reveal what these functional forms should look like to approximate reality.

ACKNOWLEDGMENTS

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.

ORCID

Regine Oexl http://orcid.org/0000-0001-5596-8746
Dmitry Ryvkin http://orcid.org/0000-0001-9314-5441

ENDNOTES

[1] The number of organizations utilizing teamwork has been growing since the 1980s (Lazear & 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 et al., 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 et al. (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 & Levine, 2004; Hollinger & 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 (Gürtler et al., 2013; Münster, 2007). For summary articles, see Chowdhury and Gürtler (2015) and Dechenaux et al. (2015), Chap. 6.1. 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 trade-off between efficiency and sabotage, and models "negative sabotage," that is, helping behavior. Empirical research that

tests theories of sabotage typically relies on data from field or laboratory experiments (Balafoutas et al., 2012; Carpenter et al., 2010; del Corral et al., 2010; Deutscher et al., 2013; Garicano & Palacios-Huerta, 2005; Harbring & Irlenbusch, 2008, 2011); by and large, these empirical studies confirm the above theories.

5 "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.

6 See, for example, "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.

7 As a referee has pointed out, our results, therefore, may not hold in one-shot interactions.

8 The theoretical literature found diverging results on the effect of heterogeneity on behavior in tournaments with sabotage. For example, Chen (2003) and Münster (2007) find that high-ability workers are more likely to attract sabotage. By contrast, Kräkel (2005) finds that an underdog is more likely to choose help than a favorite, whereas a favorite is likelier to sabotage than an underdog. These basic predictions were supported in subsequent experiments (Charness et al., 2014; Harbring et al., 2007; Vandegrift & Yavas, 2010); yet, in all these papers, helping behavior is usually ignored or treated as "negative sabotage," which disregards behavioral aspects.

9 The main difference between these models and ours is that they model the tournament component of the incentives à la Lazear and Rosen (1981) whereas we employ a lottery CSF 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.

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.

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

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

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, the session payoffs were translated into Euros at the exchange rate of 100 Tokens = €6.

15 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.

16 At least anecdotally, the addition of a better competitor is often believed to have a positive effect on others' efforts. Take, for instance, the "speeding up" Upton Sinclair wrote about in *The Jungle*, where a foreman would put a "pace man" for other workers to try to catch. The use of a "rabbit"—another runner—in distance running events also fulfills the same role. These rabbits run just a bit faster than the leaders for a portion of the race (usually the beginning) resulting in a faster pace than could be achieved without the rabbit. Such a positive "superstar effect" was identified empirically, for example, by Hill (2014a) in track and field, Hill (2014b) in sprint, and Yamane and Hayashi (2015) in swimming. The alternative point of view is that there is a negative superstar effect, which was observed, for example, in golf (Brown, 2011) and chess (Bilen & Matros, 2021).

17 See also Figures A1–A4 in Appendix A for an overview of how the variables changed over time.

18 If the assumption that a manager has equal numbers of 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 analyses for succinctness.

19 The direction of our results in this section is robust to the order of play and learning. These robustness checks are available from the authors upon request.

20 As explained in Section 3.4, this holds even if the deviations are negative. A negative deviation in the baseline (the noncompetitive setting) indicates behavioral factors are suppressing the corresponding variable. A larger (i.e., a positive or less negative) deviation in the competitive setting would imply that competitiveness leads to a behavioral boost in the variable relative to the baseline.

21 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.

REFERENCES

Alchian, A. A., & Demsetz, H. (1972). Production, information costs, and economic organization. *American Economic Review*, *62*(5), 777–795.

Balafoutas, L., Lindner, F., & Sutter, M. (2012). Sabotage in tournaments: Evidence from a natural experiment. *Kyklos*, *65*(4), 425–441.

Bilen, E., & Matros, A. (2021). *The queen's gambit: Explaining the superstar effect using evidence from chess* [Working Paper]. https://ernbilen.github.io/pdfs/Superstar_Effect.pdf.

Bock, O., Baetge, I., & Nicklisch, A. (2014). hroot: Hamburg registration and organization online tool. *European Economic Review*, 71(C), 117–120.

Brown, J. (2011). Quitters never win: The (adverse) incentive effects of competing with superstars. *Journal of Political Economy*, *119*(5), 982–1013.

Brown, M., & Heywood, J. (2009). Helpless in finance: The cost of helping effort among bank employees. *Journal of Labor Research*, *30*(2), 176–195.

Buser, T., & Dreber, A. (2016). The flipside of comparative payment schemes. *Management Science*, *62*(9), 2626–2638.

Carpenter, J., Matthews, P. H., & Schirm, J. (2010). Tournaments and office politics: Evidence from a real effort experiment. *American Economic Review*, *100*(1), 504–517.

Charness, G., & Levine, D. I. (2004). *Sabotage! Survey evidence on when it is acceptable* [Working Paper]. https://escholarship.org/content/qt5qw3h8fj/qt5qw3h8fj_noSplash_feb21480d25a896f5caf476c7787e01a.pdf?t=krno8w.

Charness, G., Masclet, D., & Villeval, M. C. (2014). The dark side of competition for status. *Management Science*, *60*(1), 38–55.

Chen, K.-P. (2003). Sabotage in promotion tournaments. *Journal of Law, Economics and Organization*, *19*(1), 119–140.

Chowdhury, S. M., & Gürtler, O. (2015). Sabotage in contests: A survey. *Public Choice*, *164*(1–2), 135–155.

Crosetto, P., & Filippin, A. (2013). The "bomb" risk elicitation task. *Journal of Risk and Uncertainty*, *47*(1), 31–65.

Danilov, A., Harbring, C., & Irlenbusch, B. (2019). Helping under a combination of team and tournament incentives. *Journal of Economic Behavior & Organization*, *162*, 120–135.

Dechenaux, E., Kovenock, D., & Sheremeta, R. M. (2015). A survey of experimental research on contests, all-pay auctions and tournaments. *Experimental Economics*, *18*(4), 609–669.

del Corral, J., Prieto-Rodriguez, J., & Simmons, R. (2010). The effect of incentives on sabotage: The case of Spanish Football. *Journal of Sports Economics*, *11*(3), 243–260.

Deutscher, C., Frick, B., Gürtler, O., & Prinz, J. (2013). Sabotage in tournaments with heterogeneous contestants: Empirical evidence from the soccer pitch. *The Scandinavian Journal of Economics*, *115*(4), 1138–1157.

Dickinson, D. L., & Isaac, R. M. (1998). Absolute and relative rewards for individuals in team production. *Managerial and Decision Economics*, *19*(4–5), 299–310.

Drago, R., & Garvey, G. T. (1998). Incentives for helping on the job: Theory and evidence. *Journal of Labor Economics*, *16*(1), 1–25.

Dye, R. A. (1984). The trouble with tournaments. *Economic Inquiry*, *22*(1), 147–149.

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

Garicano, L., & Palacios-Huerta, I. (2005). *Sabotage in tournaments: Making the beautiful game a bit less beautiful* [Working Paper]. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=831964.

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

Gürtler, O., & Münster, J. (2013). Rational self-sabotage. *Mathematical Social Sciences*, *65*(1), 1–4.

Gürtler, O., Münster, J., & Nieken, P. (2013). Information policy in tournaments with sabotage. *The Scandinavian Journal of Economics*, *115*(3), 932–966.

Hamilton, B. H., Nickerson, J. A., & Owan, H. (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, C., & Irlenbusch, B. (2008). How many winners are good to have? On tournaments with sabotage. *Journal of Economic Behavior & Organization*, *65*(3–4), 682–702.

Harbring, C., & Irlenbusch, B. (2011). Sabotage in tournaments: Evidence from a laboratory experiment. *Management Science*, *57*(4), 611–627.

Harbring, C., Irlenbusch, B., Krakel, M., & Selten, R. (2007). Sabotage in corporate contests—An experimental analysis. *International Journal of the Economics of Business*, *14*(3), 367–392.

Hill, B. (2014a). The heat is on: Tournament structure, peer effects, and performance. *Journal of Sports Economics*, *15*(4), 315–337.

Hill, B. (2014b). The superstar effect in 100-meter tournaments. *International Journal of Sport Finance*, *9*(2), 111.

Hollinger, R., & Clark, J. (1983). *Theft by employees*. Lexington Books.

Johnson, D., & Salmon, T. C. (2016). Sabotage versus discouragement: Which dominates post promotion tournament behavior? *Southern Economic Journal*, *82*(3), 673–696.

Kräkel, M. (2005). Helping and sabotaging in tournaments. *International Game Theory Review*, *7*(2), 211–228.

Lawler, E. E., Mohrman, S. A., & Benson, G. (2001). *Organizing for high performance: Employee involvement TQM, reengineering, and knowledge management in the Fortune 1000: The CEO report*. Jossey-Bass.

Lazear, E. P. (1989). Pay equality and industrial politics. *Journal of Political Economy*, *97*(3), 561–580.

Lazear, E. P., & Rosen, S. (1981). Rank-order tournaments as optimum labor contracts. *Journal of Political Economy*, *89*(5), 841–864.

Lazear, E. P., & Shaw, K. L. (2007). Personnel economics: The economist's view of human resources. *Journal of Economic Perspectives*, *21*(4), 91–114.
Münster, J. (2007). Selection tournaments, sabotage, and participation. *Journal of Economics and Management Strategy*, *16*, 943–970.
Nalebuff, B. J., & Stiglitz, J. E. (1983). Prizes and incentives: Towards a general theory of compensation and competition. *The Bell Journal of Economics*, *14*(1), 21–43.
Roth, A. E., & Kagel, J. H. (1995). *The handbook of experimental economics* (Vol. 1). Princeton university press.
Shubik, M. (1954). Does the fittest necessarily survive. In M. Shubik (Ed.), *Readings in game theory and related behavior*. Doubleday.
Tullock, G. (1980). Efficient rent seeking. In J. M. Buchanan, R. D. Tollison, & G. Tollison (Eds.), *Toward a theory of the rent-seeking society* (pp. 97–112). Texas A&M University Press.
Vandegrift, D., & Yavas, A. (2010). An experimental test of sabotage in tournaments. *Journal of Institutional and Theoretical Economics*, *166*(2), 259–285.
Yamane, S., & Hayashi, R. (2015). Peer effects among swimmers. *The Scandinavian Journal of Economics*, *117*(4), 1230–1255.

How to cite this article: Glenn Dutcher, E., Oexl, R., Ryvkin, D., & Salmon, T. C. (2024). Do competitive bonuses ruin cooperation in heterogeneous teams? *Journal of Economics & Management Strategy*, 1–35. https://doi.org/10.1111/jems.12573

APPENDIX A: MORE FIGURES AND RESULTS

See Figures A1–A4 and Tables A1 and A2.

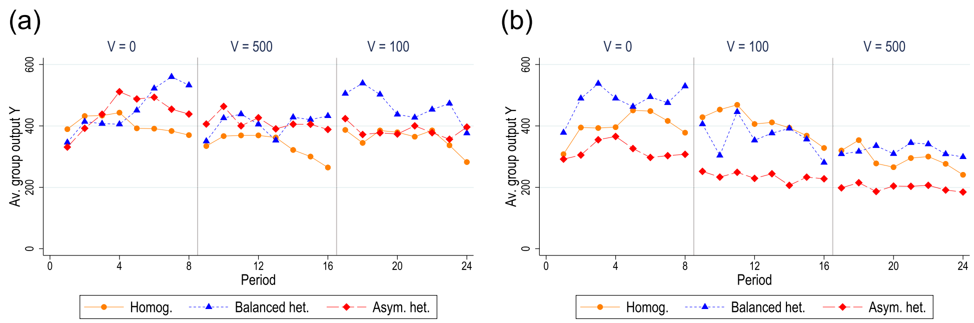

FIGURE A1 Average group output for Homogeneous, Balanced heterogeneous, or Asymmetric heterogeneous team composition over time. (a) Order 1 and (b) order 0.

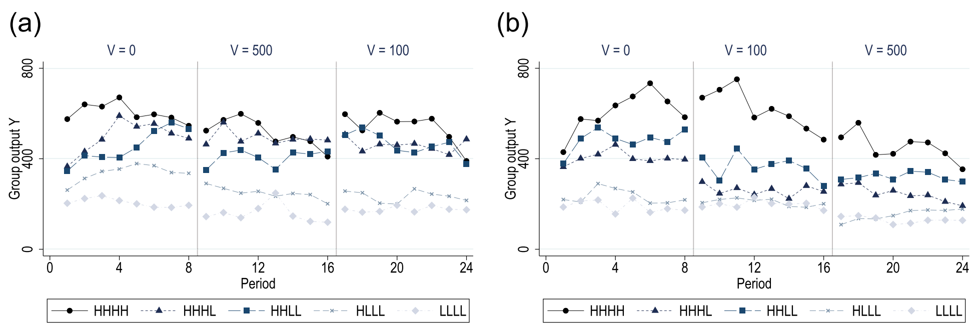

FIGURE A2 Average group output over time. (a) Order 1 and (b) order 0.

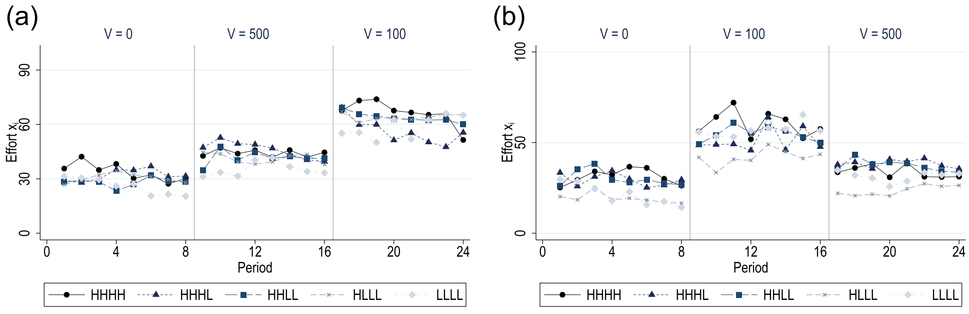

FIGURE A3 Average effort over time. (a) Order 1 and (b) order 0.

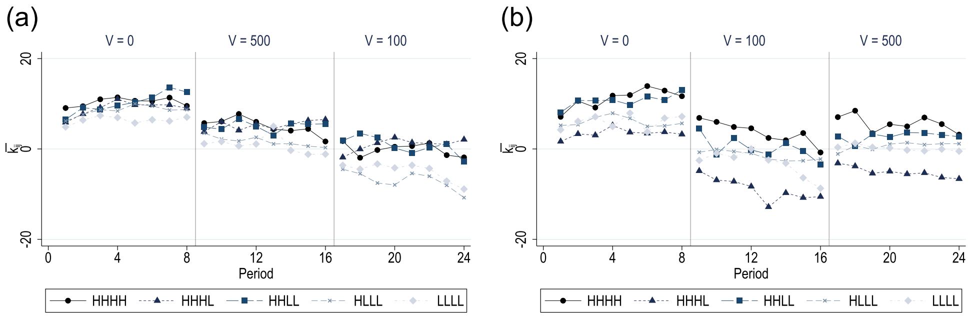

FIGURE A4 Average help over time. (a) Order 1 and (b) order 0.

TABLE A1 Overview of observed averages and clustered standard errors.

	LLLL	*HLLL*	*HHLL*	*HHHL*	*HHHH*
$V = 0$					
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_4	197.46	285.30	468.48	465.79	605.04
	(13.70)	(24.25)	(52.24)	(29.54)	(29.29)

TABLE A1 (Continued)

	LLLL	*HLLL*	*HHLL*	*HHHL*	*HHHH*
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_4	144.00	200.58	363.53	409.50	483.21
	(14.00)	(21.89)	(43.70)	(43.83)	(29.91)
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_4	187.09	221.08	414.29	393.58	578.33
	(16.74)	(23.44)	(42.66)	(41.13)	(61.45)

(Continues)

TABLE A1 (Continued)

	LLLL	*HLLL*	*HHLL*	*HHHL*	*HHHH*
Number of individuals	48	48	48	72	48
Number of teams	12	12	12	18	12
Number of rounds	8	8	8	8	8

Note: Standard errors are clustered on the group level and displayed in parenthesis.

TABLE A2 Deviations of observed averages from equilibrium predictions.

	LLLL	*LLLH*	*LLHH*	*LHHH*	*HHHH*
V = 0					
x_L	13.87	14.70	11.24	20.93	
x_H		6.34	17.70	11.83	13.01
k_{LL}	−1.50	−0.84	−0.49		
k_{LH}		−4.76	−2.33	−12.39	
k_{HL}		−3.77	0.17	−2.42	
k_{HH}			0.68	−7.70	−5.18
y_L	10.30	9.25	11.18	14.12	
y_H		−15.88	27.75	−18.71	−4.99
Y_4	41.21	11.86	77.85	−42.02	−19.96
V = 100					
x_L	4.76	9.25	12.54	22.83	
x_H		3.96	16.14	14.55	10.14
k_{LL}	−2.86	−5.18	−4.11		
k_{LH}		−11.16	−8.92	−17.67	
k_{HL}		−4.81	−3.45	−5.49	
k_{HH}			−6.53	−11.25	−9.31
y_L	−3.06	−4.62	2.72	9.52	
y_H		−59.00	−16.27	−35.94	−35.45
Y_4	−12.25	−72.85	−27.09	−98.31	−141.79
V = 500					
x_L	−45.71	−20.55	−5.88	0.20	
x_H		−20.15	−9.96	−15.34	−3.01
k_{LL}	16.02	−3.37	−3.96		
k_{LH}		−2.63	−4.94	−20.79	
k_{HL}		19.57	8.29	0.07	
k_{HH}			18.05	2.54	−1.56
y_L	7.71	0.77	12.32	8.58	
y_H		−54.67	−0.49	−40.94	−11.67
Y_4	30.84	−52.35	23.67	−114.23	−46.67

APPENDIX 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 will not 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 Euros. 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 min. It consists of two parts; both parts are completely independent of 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 into three blocks of eight periods; Blocks 1–3. 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 Euros 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 eight periods. To ensure anonymity, you and the three others in your group will be labeled by the computer program as members A–D. Each group member's label will be the same for all eight 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 eight 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 in the range of 0–150 (a whole number). Each point you choose will be costly to you. You can find a table of costs on a 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 Tokens, and 100 points 128 Tokens. 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 the 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 $162 * 0.25 = 40.5$ Tokens. 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 Tokens. This would result in total earnings of $40.5 - 6.72 = 33.78$ Tokens. 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 Tokens, 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 four 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 $1.28 + 1.28 + 1.28 + 1.28 = 5.12$ Tokens: 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 your 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 $5.12 + 0 + 1.28 + 1.28 = 7.68$ Tokens, 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 Tokens (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 to you by means of an example of how your payoff is calculated. Let us assume that the total group output from your group was 162 points. The gain you would receive from this total group output is $162 * 0.25 = 40.5$ Tokens. Let us 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 $5.12 + 4.15 + 1.28 + 6.77 = 17.32$ Tokens. 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 three other people to make a group of 4. You will keep this type and this group for the entire eight 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 three 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 eight 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] Tokens. Only one group member can win the bonus per period. Your probability of winning is determined as

$$\text{Chance of winning} = \frac{}{\text{TIO of A} + \text{TIO of B} + \text{TIO of C} + \text{TIO of D}}.$$

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 of 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 eight balls.

Let us go through another example. If your group members' total individuals 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 individuals 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 eight periods where you have the same four choices: your own effort and the modifications to your group members' efforts. The costs and the 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] Tokens. 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, as shown below.

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

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 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—that is, 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—that is, if your chosen number is smaller than the drawn number—you will earn an amount equivalent to the number you have chosen (Table B1).

TABLE B1 Cost table.

Your choice of effort	Costs	Your choice of effort	Costs	Your choice	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

(Continues)

TABLE B1 (Continued)

Your choice of effort	Costs	Your choice of effort	Costs	Your choice	Costs
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				

APPENDIX C: SCREENSHOTS

See Figure C1.

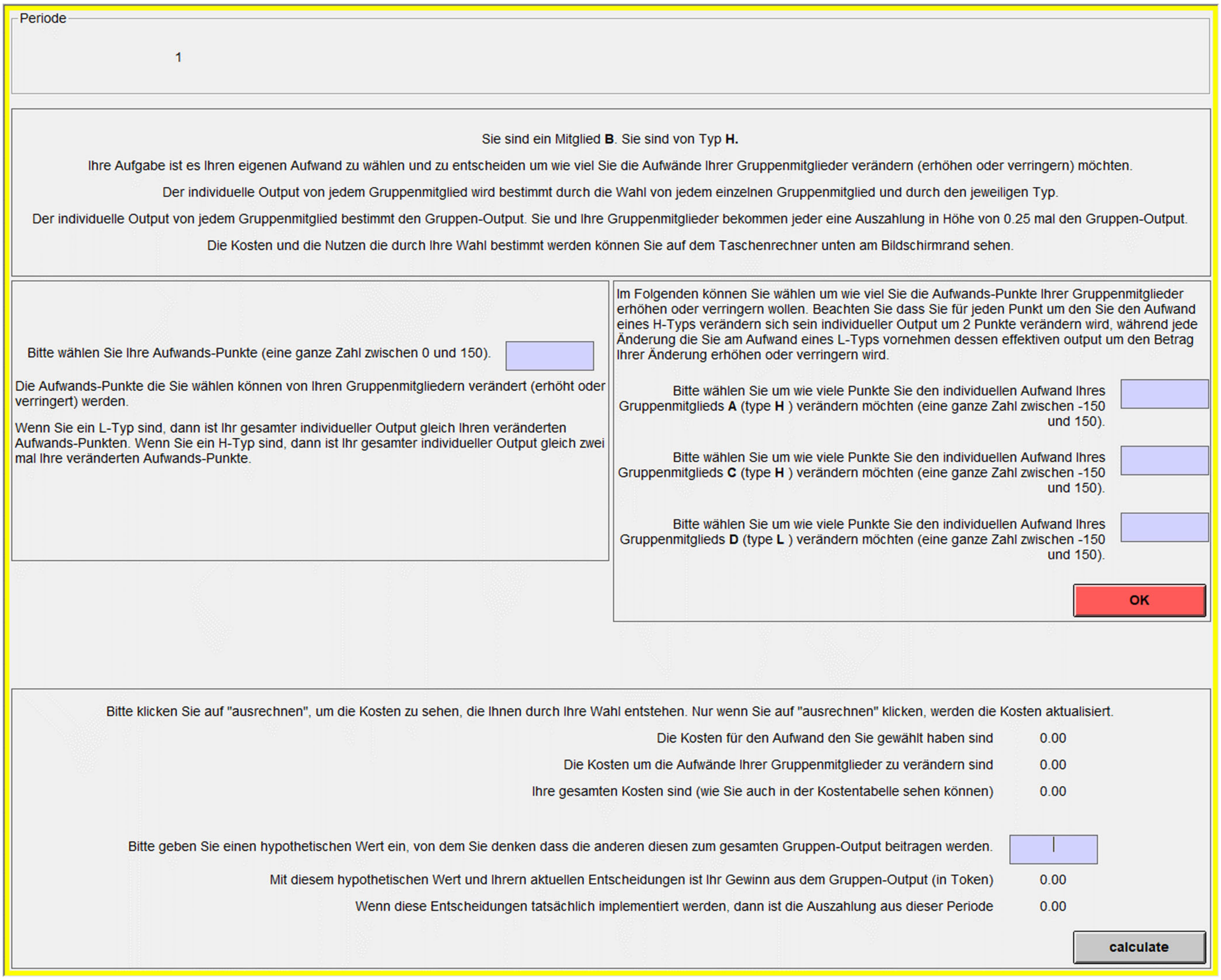

FIGURE C1 Full screen 1.