Title

Incentives for public goods inside organizations: Field experimental evidence

Authors

Blasco, Andrea
Jung, Olivia S
Lakhani, Karim R
et al.

Publication Date

2019-04-01

DOI

10.1016/j.jebo.2019.02.029

Journal of Economic Behavior and Organization 160 (2019) 214–229

Contents lists available at ScienceDirect

Journal of Economic Behavior and Organization

journal homepage: www.elsevier.com/locate/jebo

Incentives for public goods inside organizations: Field experimental evidence

Andrea Blasco[a,*], Olivia S. Jung[b], Karim R. Lakhani[b,c], Michael Menietti[b]

[a] Harvard Institute for Quantitative Social Science, Harvard University, 1737 Cambridge Street, Cambridge, MA 02138, United States
[b] Harvard Business School, Soldiers Field, Boston, MA 02163, United States
[c] National Bureau of Economic Research, Cambridge, MA 02138, United States

ARTICLE INFO

Article history:
Received 13 July 2018
Revised 10 January 2019
Accepted 25 February 2019
Available online 16 March 2019

JEL classification:
D23
H41
M52

Keywords:
Innovation contests
Relative incentives
Organizational improvement
Free rider problem
Social incentives
Organization of work
Healthcare organization

ABSTRACT

Understanding why employees go the extra mile at work is a key problem for many organizations. We conduct a field experiment at a medical organization to study motivations for employees to submit project proposals for organizational improvement. In total, we analyze 1237 employees, 118 proposals, and quality evaluations for more than 12,000 evaluator-proposal pairs. The analysis shows that solicitations offering a personal reward for top submissions boost participation rates without affecting submission quality. We show that this is due to workers partially internalizing the positive effects of their submissions on the other individuals in the workplace. We also find that offering employees project funding to implement their own proposals potentially backfires, undermining participation. And solicitations emphasizing mission-oriented goals, like improving patient care, are sensitive to the solicited person's gender, with women responding more than men. These results shed light on the factors that drive employees' engagement in organizational tasks, beyond regular duties, and provide insights on how to design incentives to foster contributions to public goods inside organizations.

1. Introduction

One of the primary ways for employees to boost organizational performance is to provide collective goods, such as contributing to common knowledge repositories, taking extra assignments, and making suggestions for organizational improvement. Multiple factors, however, prevent the management from deploying compensation schemes to encourage a higher provision of collective goods from their staff. The basic idea is that linking an agent's pay to an imperfect measure of individual contributions may backfire producing multitasking problems (Holmström and Milgrom, 1991), free-riding incentives (Che and Yoo, 2001), less experimentation (Ederer and Manso, 2013) and, sometimes, it is simply infeasible given the non-standard nature of the desired contributions (Hellmann and Thiele, 2011).

In this study, we conduct a natural field experiment to examine what drives employees' participation and effort in tasks that are above and beyond their contractual duties. Specifically, we focus on how employees respond to a call soliciting

* Corresponding author.
E-mail addresses: ablasco@fas.harvard.edu (A. Blasco), ojung@hbs.edu (O.S. Jung), k@hbs.edu (K.R. Lakhani), mmenietti@hbs.edu (M. Menietti).

https://doi.org/10.1016/j.jebo.2019.02.029
0167-2681/

innovative project proposals for organizational improvement. Within this context, we are interested in how employees respond to *prize incentives* (Lazear and Rosen, 1981; Mary et al., 1984), such as the prospect of winning a personal reward, compared to *mission-oriented* motivations (Akerlof and Kranton, 2005; Besley and Ghatak, 2005; Delfgaauw and Dur, 2008; Prendergast, 2008), such as the inner satisfaction from helping their fellow workers and the organization as a whole.

The experiment involves more than 1200 health workers (doctors, nurses, technicians, administrative workers) in a medical organization in the United States. Workers are encouraged to participate in an internal innovation contest hosted by the organization. The main goal of the innovation contest is to invite all staff to contribute innovation project proposals for improving the operations of the organization. Workers' participation and involvement in the innovation contest are discretionary and go beyond their regular duties.[1]

To assess the effect of prize incentives versus mission-oriented motivations on participation, we use a randomized encouragement design where we select staff members to receive a personalized email from the management announcing the innovation contest. The email encourages agents to participate by submitting project proposals. The text of the encouragement is at random one of four types: "prize," "funding," "workplace," and "patient care."

The "prize" encouragement reveals the opportunity of winning a personal reward (Apple iPad) for the best project proposals. The "funding" encouragement reveals the opportunity of winning project funding (an individual grant) to lead the execution of the proposed innovation project. The "workplace" encouragement emphasizes the need for workplace improvements. And the "patient care" encouragement emphasizes the opportunity of improving health care delivery. Importantly, the information of a personal reward (either an Apple iPad or an individual grant) is withheld in the last two encouragements (workplace and patient care) in order to make altruistic motives more salient.

Each encouragement email has in its body a generic commitment by the management to implement the best projects. Such a commitment ensures the legitimacy and credibility of the contest in all the treatments. Only the solicitation in the funding treatment, however, reveals to the workers that they will be in charge of their own project's implementation through an individual grant.

Workers may view the grant opportunity as an additional incentive to participate, such as a higher satisfaction from being immediately responsible for the organizational change (a kind of warm-glow, e.g., Andreoni, 1990) or a formal recognition (something to put on a CV). But they could also view the grant opportunity as a demotivating factor due to the additional responsibilities and costs from being in charge of their own project's execution.

We evaluate the causal impact of our four experimental conditions on two main response variables: (a) the employees' participation in the contest, which is measured by the decision to submit a proposal and engage in an organizational improvement task and (b) the quality of the submissions as measured by over 12,000 peer ratings and the assessment made by the management.

Testing for both participation and quality effects is important as the presence of systematic quality differences associated with different motivations would substantially complicate the problem of incentives for the organization. Higher employee participation is a less desirable goal if it is driven by low-quality proposals.

We also investigate the presence of sorting effects based on the gender, profession, and position inside the organization of the solicited employee, and estimate the extent of treatment interaction effects associated with these characteristics. Following the economic literature on gender-based differences in preferences (Croson and Gneezy, 2009), one may expect sorting to arise due to gender differences in altruistic preferences (Andreoni and Vesterlund, 2001) or differences in competitive inclinations (Niederle and Vesterlund, 2007), complicating the analysis of the incentives.

Several features make the above empirical setting ideal for our research. First, healthcare professionals are commonly seen as willing to step beyond the boundaries of their contractual duties to offer better care (Delfgaauw, 2005), which makes the comparison of different incentives towards the provision of the public good (e.g., organizational improvement) especially relevant and interesting. Second, understanding what drives internal innovation seems crucial in the context of healthcare delivery, as the need for organizational improvement and internal innovation is vastly noted (Cutler et al., 2012). Third, it has become increasingly common for organizations to run internal innovation contests with the goal of raising employee participation in various organizational activities,[2] which makes our results immediately applicable to many other empirical settings.

Our analysis establishes that prize incentives boost participation without affecting the quality of the proposals. The size of the estimated effect on participation is nontrivial. Had all the staff members been offered prize incentives, the organization would have received more than *three times* the submissions had they all been offered the funding incentives, with the same quality on average. Give its size, the effect appears to go beyond the value of the awarded prize suggesting that workers

[1] Within this context, participation is costly to the individual worker because of the time to identify a problem, form a proposal, write up a concise description, and the potential for further work and involvement during proposal implementation. But the benefits of improvement efforts are shared with the fellow workers as well as the patients. Therefore, each staff member has an incentive to free ride on the work of others.

[2] Among the many examples of internal contests that have appeared in the news are the Apple's 2016 contest among its store employees seeking ideas on how to improve iPhone sales ("Apple seeks 'pie in the sky' ideas for innovation," Computerworld, 2013); Xerox's internal contest seeking employees ideas on how to make a more environmentally friendly workplace environment ("Xerox employees green ideas save company $10.2 million," The Guardian, 2010); and AT&T's ideation contests seeking employee ideas about new products ("AT&T develops employee ideas for innovation," The Wall Street Journal, 2014).

partially internalize the spillovers of their contributions for the other members of the organization, including their fellow workers and patients.

We also find that women respond more than men to mission-oriented incentives, such as the goal of improving patient care, and offering workers project money to run their own projects potentially backfire undermining participation.

We discuss how these results highlight a relatively underscored role of prize incentives in organizations, which is to raise internal contributions to public goods, and may help to point out a few limitations.

2. Related literature

The present study contributes to the literature on social incentives in the workplace (Bandiera et al., 2005, 2008, 2010, 2011; Della Vigna et al., 2016). In particular, Bandiera et al. (2005) finds evidence that workers internalize the negative externality of their actions on the other workers by showing that individual productivity drops when the firm under study switches from a relative incentive scheme (where workers' pay depends on their productivity relative to others) to piece rates (where workers' pay depends on their individual productivity). In this study, we look at similar social incentives but we show that this kind of negative externality can be positive for the firm when it serves to mitigate free riding incentives of the employees.

The gains of prize-based mechanisms to foster the provision of public goods have been studied before. Morgan (2000) posits that engaging a set of potential contributors in a competition for prizes will increase the equilibrium provision of a public good. The intuition behind this result is that two opposing externalities (one negative externality associated with the prize competition and the positive externality of contributing to public goods) will mitigate the free-riding problem, thus increasing the equilibrium provision. This result has received much attention in public economics (Vesterlund, 2012), as it provides an economic rationale for the use of prizes in a variety of contexts, particularly charity donations. However, the organizational context is more challenging to analyze because, among many other things, social connections among co-workers may influence employees contributions and the free-riding incentives (Bandiera et al., 2005). We address the problem by providing field experimental evidence.

The present study contributes more broadly to the literature on the incentive effects of tournaments on productivity at work (Ashraf et al., 2014; Barankay, 2012; Boudreau et al., 2011, 2016; Delfgaauw et al., 2013; Ehrenberg and Bognanno, 1990; Gallus, 2016; Gibbs et al., 2015; Knoeber and Thurman, 1994; Vidal and Nossol, 2011). Unlike previous studies, however, competitors are aware that their performance potentially yields positive spillovers to everyone in the workplace, including their opponents. This makes the analysis particularly interesting because the incentive effect of contests with spillovers is controversial (Drago and Turnbull, 1988) and we lack empirical evidence on the issue. In addition, while past literature has been mainly focused on contests with standardized tasks (e.g., fruit picking, growing chicken) or sports (e.g., golf), we examine performance in a contest with a non-standard task that involves some degree of creativity (i.e., submitting project proposals). This distinctive feature of the task to perform allows us to study the causal effect of our treatment on the participation, quality, and content of the contributions, showing the lack of a quantity versus quality trade-off.

This study is also related to the literature on mission-oriented incentives at work (Akerlof and Kranton, 2005; Besley and Ghatak, 2005; Carpenter and Gong, 2016). This literature departs from the traditional assumption of agency problems inside firms to consider a situation of almost aligned incentives between workers and their enterprise. This scenario is often considered plausible in the context of firms pursuing social public goods, such as hospitals, schools, the government (Delfgaauw and Dur, 2008; Prendergast, 2007). Within the hospital under study, we report gender-based differences in the response of the staff members to a solicitation emphasizing different intrinsic motivations associated with the mission of the firm (e.g., patients versus workplace). We start unpacking this effect by looking at the literature on gender-based differences in preferences (Croson and Gneezy, 2009).

Finally, this paper is also broadly related to the empirical literature on the free rider problem in team production (Boning et al., 2007; Erev et al., 1993; Gibbs et al., 2015; Hamilton et al., 2003; Mas and Moretti, 2009). However, it differs from much of the existing literature because it focuses on a situation where conventional *peer effects* (peer pressure, monitoring, reciprocity among team members, and other kinds of social interactions that are usually seen as opposing to free riding) are muted — teams are not allowed, participation is confidential, and staff members receive no feedback about the participation of the other workers during the contest.

3. Context, experimental design, and data

3.1. The Heart Center's internal innovation contest

The medical organization under study is the Massachusetts General Hospital's (MGH) Corrigan Minehan Heart Center (here referred to as the "Heart Center"). Founded more than a hundred years ago, the Heart Center is a prominent medical organization in the United States and a teaching hospital of the Harvard Medical School. It serves thousands of patients every year, occupies more than 35,000 square feet of office space, and employs more than 1200 people (nurses, physicians, researchers, technicians, and administrative staff) scattered across several buildings on the MGH's main campus in downtown Boston and a few other satellite locations.

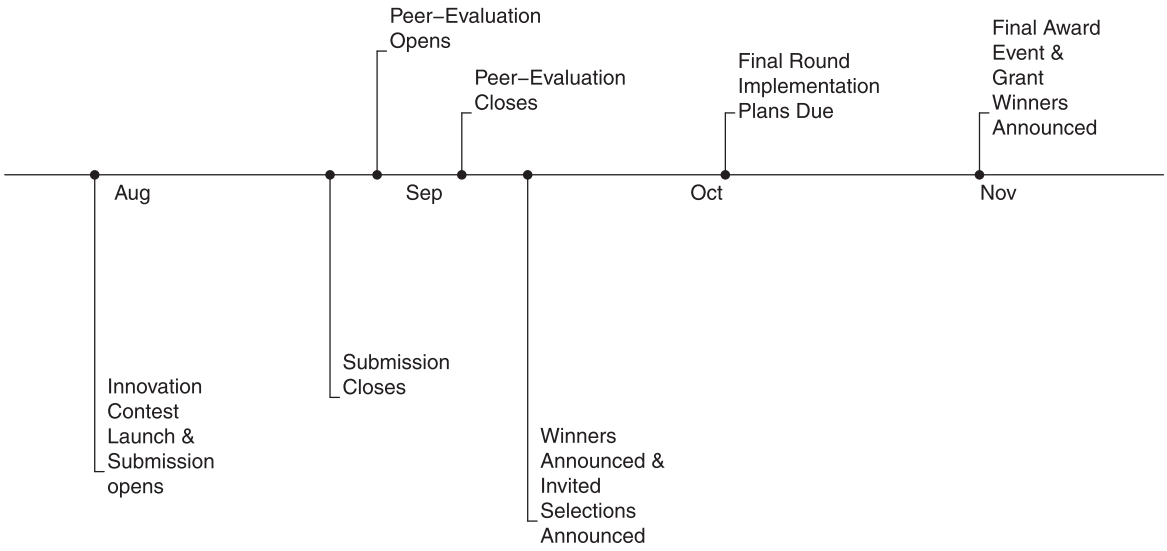

Fig. 1. Timeline showing the key phases of the internal innovation contest.

The opportunity for this study arises upon the Heart Center's launch of the Healthcare Transformation Lab (HTL), an initiative aimed at developing innovative healthcare process improvements to enhance the healthcare safety and delivery of the hospital.[3]

The launch of the HTL is accompanied by the announcement of an internal innovation contest, called the Ether Dome Challenge,[4] seeking to engage all staff members to participate. The communication around the innovation contest highlights the opportunity for staff to help in the selection process of the ideas and a commitment by the Heart Center Management that the leading ideas would be provided appropriate resources so that they could be implemented. The announcement on the contest website reads:

> "If you've noticed something about patient experience, employee satisfaction, workplace efficiency, or anything that could be improved; if you've had an inspiration about a new way to safeguard health; or if you simply have a cost-saving idea, then now is the time to share your idea."

We collaborate with the HTL administrators to design the innovation contest, which consists of three main phases (Fig. 1): submission, evaluation, and implementation.

In the submission phase, all staff members are encouraged to identify one or more organizational problems and submit proposals addressing them. Employee participation is voluntary and all project submissions can be done online via the website of the contest. There is no limit on the number of project proposals to submit and proposals could cover any issue within the organization. The only constraints are: a limit of approximately 300 words per proposal and team submissions are not permitted. The first constraint is to lower the costs of entry and encourage broader participation. The second constraint is to ensure that randomly assigned treatment effects, which we describe next, could be isolated, identified, and matched to individual participants.[5] In addition, the website provides no feedback during the submission period, so that individual decisions could not be influenced by the perceived popularity of the contest or by previous submissions.

In the evaluation phase, all staff members are invited to rate the merit and potential of all the proposals. The contest website presents a full list of anonymized proposals to read and rate. Proposals appear in batches of 10 and in random order to ensure even exposure. Each proposal has a title, the main description of the problem to solve, and the proposed solution. The evaluation is then presented by the following text: "Rate this idea" followed by a five-point rating scale (1-low; 2; 3; 4; 5-high). Evaluators can pick how many proposals to rate and they could win a limited edition T-shirt as compensation for their efforts. Ratings are confidential and the website provides no feedback or any other information that would influence individual judgment until the end of the evaluation.

In the implementation phase, employees who have submitted proposals highly rated by peers and judged as particularly promising by the HTL staff are invited to submit a full proposal detailing plans for implementation. Following evaluation by MGH senior leadership, the best among these full-length proposals are selected to receive support and funding for im-

[3] See HTL's website (www.healthcaretransformation.org).

[4] The name is taken from a historical place on MGH's main campus where the first public surgery using anesthesia was demonstrated in 1846.

[5] Another advantage of the individual submission constraint is that it lowers the incentives to communicate or exchange information among employees by preventing groups to form, thus lowering the risk of "interference" among individuals in the different treatments, a problem we will discuss later.

Table 1
Experimental design.

	Solicitation text	Assignment	
		N	%
FUND	Submit your ideas to win project funding up to $20,000 to turn your ideas into actions.	308	25
PCARE	Submit your ideas to improve patient care at the Heart Center.	310	25
PRIZE	Submit your ideas to win an Apple iPad mini.	312	25
WPLACE	Submit your ideas to improve the workplace at the Heart Center.	307	25

Note: Table showing the text of the solicitation email in each treatment along with the number and percentage of randomly selected staff members assigned to receive the solicitation treatment.

plementation. This final phase takes a few months to complete, essentially the time necessary to select and implement the best projects.

3.2. Experimental design

We use an encouragement design where we randomize the content of the personalized email communication that announces the innovation contest to all staff members and we study the effect of the encouragement on participation.

Below is the body of the announcement with a placeholder for our solicitation treatment. The email is reproduced in the appendix.

> Dear Heart Center team member,
>
> **Submit your ideas to [TREATMENT HERE]**
>
> The Ether Dome Challenge is your chance to submit ideas on how to improve the MGH Corrigan Minehan Heart Center, patient care and satisfaction, workplace efficiency and cost. All Heart Center Staff are eligible to submit ideas online. We encourage you to submit as many ideas as you have: no ideas are too big or too small!
>
> Submissions will be reviewed and judged in two rounds, first by the Heart Center staff via crowd-voting, and then by an expert panel. Winning ideas will be eligible for project implementation funding in the Fall of 2014!

We randomize the first paragraph of the above invitation into *four* different solicitation treatments (Table 1). In the "prize" solicitation, employees are given information on the opportunity to win a prize (iPad mini's) for the top submissions.[6] In the "funding" solicitation, employees are given information on the opportunity to win up to $20,000 budget for developing their own project proposals. In the "patient care" solicitation, employees are invited to participate in the contest for the opportunity to improve the healthcare of their patients. Finally, in the "workplace" solicitation, employees are invited to participate in the contest for the opportunity to improve their workplace.

Importantly, in the prize and funding treatments, the encouragement invites employees to participate for the opportunity of either winning a prize or leading the implementation of their own projects. Whereas, in the patient care and workplace treatments, the encouragement withholds any information about these rewards to make altruistic motives more salient.

By splitting the sample into four treatment groups, we obtain a sample size per treatment group of more than 300 staff members. This ensures a sufficiently high statistical power based upon standard power calculations on the difference of proportions. In particular, the statistical power is just below 80% for *small* differences at 5% significance level but higher than 80% for *medium* and *large* differences, where the definition of size is given by Cohen (1992).[7]

To reinforce the treatment effects, the invitations are sent three times during the contest (at the launch, eight days from the launch, and two days before the end of the submission phase of the challenge). Additionally, the website of the innovation contest has personalized graphics and layout to reinforce the effect of the announcement. That is, the headings, background images, a short video, and space just below a "submit your ideas" buttons are designed to match the first paragraph of the solicitation email that the employee received (i.e., those in Table 3.2). These elements are reproduced in the appendix.

To prevent a potential bias in the communication of the innovation contest, the MGH management and the HTL staff members are blind to group assignment and we rely only on official channels for communication to strengthen the effect of the announcement and signal legitimacy of the contest.

To generate a "safe" environment for employees submitting proposals, the submission form tells that the identity of the workers will be kept private to the management unless they self-identify. Management could not identify workers without their consent in the evaluation of proposals.

To encourage everyone to take part in the event, we distribute flyers and posters starting from the second week of the submission phase. These flyers and posters are based on a generic, undifferentiated version of the solicitation email without the additional information of the treatments.

[6] A total of three iPad minis were available to the winners.

[7] For instance, a difference of 5 percentage points of the pair (0.05, 0.10) is considered a small effect: see Cohen (1992) p. 158.

Table 2
Summary statistics.

group.var	Treatment assignment				N	p.val
	FUND	PCARE	PRIZE	WPLACE		
Gender						
Female	68.8%	69.4%	74%	75.6%	890	0.159
Male	31.2%	30.6%	26%	24.4%	347	
Office						
Office	50.3%	54.5%	55.8%	52.8%	660	0.556
Ward	49.7%	45.5%	44.2%	47.2%	577	
Profession						
Nurse	51%	52.3%	51%	55.7%	649	0.844
Physician	19.2%	18.1%	17.6%	18.2%	226	
Residual	29.9%	29.7%	31.4%	26.1%	362	
Tenure						
−10	39.8%	31.2%	36.9%	35.6%	132	0.956
10–19	27.3%	32.3%	29.8%	30.8%	111	
20–29	13.6%	17.2%	10.7%	12.5%	50	
+30	19.3%	19.4%	22.6%	21.2%	76	
Age						
18–25	5.7%	7.5%	6%	6.7%	24	0.656
26–35	30.7%	29%	28.6%	27.9%	107	
36–45	21.6%	21.5%	20.2%	24%	81	
46–65	34.1%	33.3%	44%	38.5%	138	
65+	8%	8.6%	1.2%	2.9%	19	

Note: Table showing the percentage of randomly selected employees assigned to each treatment cell by the gender, profession, and whether the employee had a fixed office location or was assigned to a ward. The table also shows the age and tenure (in years) at the Heart Center for a subset of about 30% employees who participated in a survey conducted a few months before the experiment. The last column shows the *p*-values from a Chi-square test of independence between the employee characteristic and the treatment assignment, showing a balanced sample.

Finally, since we study a workplace, we try to carry out treatments having equal chances of being successful. This prevents us from having a truly 'null' treatment with no personalized incentives messaging as a control group. Nevertheless, if we are to think of one treatment as the benchmark against which to compare the others, the funding treatment would be our best candidate because that is the default option for announcing internal grant programs and was part of the HTL's initial design before our collaboration.

3.3. Data

Our subject pool comprises the entire population working at the Heart Center as of the end of 2014, a total of 1237 individuals. For each individual, we have administrative data on the gender, profession, and whether they had a fixed office location within the hospital. Additional, complementary demographic information (i.e., age and tenure at the center) is available for a limited group of 378 employees (31%) who participated in an online survey conducted about two months prior to the launch of the innovation contest. Summary statistics are reported below (Table 2).

Note that the large majority (72%) of employees in our sample are women and only half of the employees (53%) have fixed office locations. The gender asymmetry is due to the high fraction of workers being nurses (52%) and the presence of a gender separation by profession with nurses being predominantly women (92%). The variation in space allocation is because medical workers may be on duty in multiple wards. Nurses, for instance, are likely to be assigned to a ward, due to the nature of their job. But it is also associated with experience and status inside the organization. A staff member with a fixed office location has a median of 2 more years of tenure compared to an employee assigned to wards (based on our survey data).

To check whether groups are statistically balanced in their distributions of pretreatment variables, we use a series of Chi-square tests of independence. Results (Table 2) fail to reject the null hypothesis of independence between the solicitation treatments and each pretreatment variable, which means that our initial randomization is successful and the observed group differences are non-systematic.

Differences in income and education are additional factors that affect employees choices inside firms. Though we lack data on individual incomes, medical professions exhibit considerable gaps in wages that make the employee's profession a good proxy for such information. According to the United States Bureau of Labor Statistics, the median annual wage of a physician in 2015 ($187,200) was about 60% higher than that of a registered nurse for the same period ($67,490), which, in turn, was about 70% higher than that of a laboratory technician ($38,970).

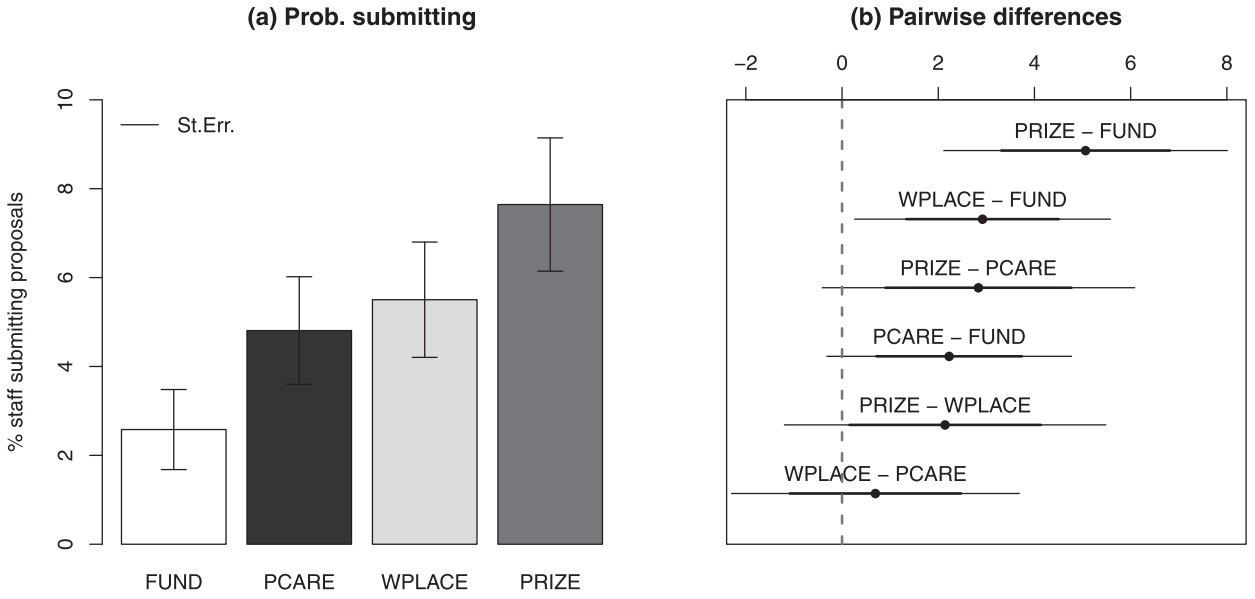

Fig. 2. Participation differences by treatment assignment. The left panel shows estimates and standard errors of the probability of staff members submitting proposals in each solicitation treatment. The right panel shows the corresponding pairwise differences with 50 and 90% confidence intervals.

To form our primary response variables, we use the outcomes of the internal innovation contest. These comprise the submissions made by 60 staff members (5% of the total) with a total of 118 project proposals (an average of 2 project proposals per submission) and the evaluations made by 178 staff members (14% of the total) and by the management.[8]

A 5% participation rate over a four-week submission period is unsurprising given the busy environment, the employees' conflicting priorities, and the foreseeable additional costs of leading implementation of a winning proposal. When we combine data from the submission and evaluation phases, however, the overall participation rate is 16% (i.e., the union of employees who made submissions and evaluated proposals). This means that about 1 out of 5 staff members was engaged in one way or another in the contest, indicating good participation overall.

4. Analysis

4.1. Participation in the contest

We begin to analyze the causal effect of our experimental intervention on staff participation in the innovation contest by looking at the proportion of employees who submitted project proposals within the four-week submission period (Fig. 2, left panel).

The proportion of employees submitting proposals ranges from a minimum of 2% in the funding treatment to a maximum of 7% in the prize treatment, showing a substantial variation across the treatment groups. A Fisher's exact test indicates a statistically significant ($p = 0.026$) association between the submission rates and the treatment assignment. Thus, the observed variation in participation rates is, at least in part, caused by our experimental intervention.

Inspection of pairwise differences among treatments (Fig. 2, right panel) further reveals that: (*i*) the proportion of employees submitting in the PRIZE solicitation treatment is greater by 2 percentage points or more than those in the other treatments; (*ii*) the proportion of employees submitting in the PCARE is basically the same as in the WPLACE solicitation treatment; and (*iii*) the proportion of employees submitting in the FUND solicitation treatment is lower by 2 percentage points or more than those in the other treatments, which corresponds to about half, or less, of the participation in the WPLACE and PCARE solicitation treatments.

To test to see whether these pairwise differences are statistically significant, we use a series of pairwise comparisons of proportions (Table 3). We find: a significant positive difference in participation rates between the PRIZE and FUND solicitation treatments ($p = 0.003$); an insignificant positive difference between the PRIZE and PCARE solicitation treatments ($p = 0.132$); an insignificant positive difference between the PRIZE and WPLACE solicitation treatments ($p = 0.269$); a significant negative difference between the FUND and WPLACE solicitation treatments ($p = 0.055$); and an insignificant negative

[8] Here, and throughout the paper, we exclude an additional 20 proposals from 11 employees who were not part of the Heart Center when the experiment was designed.

Table 3
Participation: pairwise comparisons.

	FUND	PCARE	PRIZE
PCARE	0.124		
PRIZE	0.003	0.132	
WPLACE	0.055	0.688	0.269

Note: Table showing *p*-values obtained from a series of pairwise comparisons of the proportion of employees submitting project proposals within each treatment assignment. These values are based on a two-sample test for equality of proportions.

Table 4
Probability of submitting proposals.

	Dependent variable:				
	Employee making submissions				
	(1)	(2)	(3)	(4)	(5)
FUND	−2.57***	−2.55***	−2.57***	−2.49***	−2.38***
	(0.86)	(0.85)	(0.86)	(0.86)	(0.85)
PCARE	−0.33	−0.31	−0.33	−0.36	−0.36
	(1.04)	(1.04)	(1.04)	(1.04)	(1.04)
PRIZE	2.53**	2.52**	2.53**	2.46**	2.45**
	(1.21)	(1.21)	(1.21)	(1.21)	(1.21)
genderMale		−0.54			−0.42
		(1.33)			(1.64)
jobPhysician			−0.45		−2.99
			(1.61)		(2.28)
jobResidual			0.03		−2.57
			(1.43)		(1.83)
officeWard				−2.79**	−4.56***
				(1.20)	(1.60)
Constant	4.84***	5.00***	4.92***	6.14***	8.38***
	(0.61)	(0.73)	(0.85)	(0.94)	(1.68)
Log likelihood	−5545	−5545	−5545	−5542	−5540
Observations	1237	1237	1237	1237	1237

Note: Table showing the estimated coefficients (with heteroskedasticity-robust standard errors in parenthesis) for a linear probability model of employee submissions as a function of the solicitation treatment, profession, gender, and whether the employee was assigned to a ward or to a fixed office location. Coefficients are multiplied by 100 to indicate a percentage point change in the probability of submitting. Solicitation treatment dummies are coded to indicate deviations from the overall probability of submitting. The asterisks ***, **, * indicate significance at 1, 5 and 10 percent level.

difference between the FUND and the PCARE solicitation treatments ($p = 0.124$). Overall, these findings are consistent with higher employee participation under a solicitation with personal awards incentives; and lower participation under a solicitation with funding incentives.

A multiple linear regression model that explicitly controls for observable differences across staff members complements the above univariate analysis. Let $Y_i = 1$ denote employee i making a submission, and $Y_i = 0$ otherwise. We assume the conditional probability of an employee making a submission is given by:

$$\Pr(Y_i = 1) = \alpha_0 + \sum_j \alpha_j \text{SOLICIT}_{ij} + \text{JOB}_i + \text{MALE}_i + \text{OFFICE}_i, \quad (1)$$

where α_0 is a constant, α_j is the causal effect of the solicitation treatment j assigned to an employee i (SOLICIT_{ij}), controlling for the employee's profession (JOB_i), the gender (MALE_i), and a dummy for office location (OFFICE_i) indicating whether the employee had a permanent office instead of being assigned to a ward.[9]

We report the OLS estimated coefficients for the above model (Table 4) expressed as solicitation treatment differences relative to the overall mean participation. Results are consistent with what discussed before. At the 95 level of statistical significance, employees in the PRIZE solicitation treatment are about 2 percentage points *more* likely to submit compared to the overall mean, whereas employees in the FUND solicitation treatment are 2 percentage points *less* likely to do so. Although these effects might appear small in absolute terms, they are fairly large compared to an overall participation of 5%.

[9] In this context, having a fixed office location is highly correlated with the type of profession Section 3.3. Within each profession, however, having a fixed office location is also correlated with the employee's tenure inside the organization Section 3.3. Hence, more than just the effect of having a fixed office location per se, this variable is potentially controlling for income and hierarchical differences occurring within each profession as well.

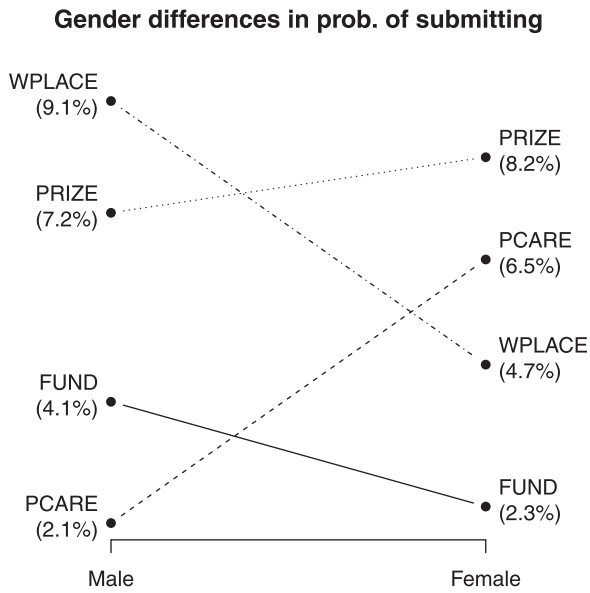

Fig. 3. Gender differences in the probability of submitting by treatment assignment. Plot showing estimates of the probability of submitting (in parenthesis) for female (right) and male (left) staff members conditional on the assigned solicitation treatment.

4.2. Sorting

Next, we examine the extent to which participation in the innovation contest is associated with relevant individual characteristics, such as the gender, profession, and office status of the solicited worker. Do female staff members participate more or less than their male counterpart? And does the profession or status of the worker matter for sorting?

Despite dissimilarity in income and education, we find no evidence of participation differences associated with the employees' profession. While physicians tend to submit less and nurses tend to submit more than the other staff members, these results are largely insignificant (Table 4). In fact, we cannot reject the null hypothesis that the corresponding coefficients are jointly equal to zero (F-test, $p = 0.989$).

We find instead strong evidence ($p = 0.003$) that having an office location, as opposed to being assigned to a ward, is associated with higher submission rates. This effect could just be a lack of opportunity: employees assigned to a ward may not have easy access to the means (workstation, privacy) to submit project proposals. But it could also be indicative of a higher utility, showing a category of employees who would benefit more from contributing to organizational goods (e.g., employees with higher tenure in the organization).

For what concerns gender effects, one may expect female staff members to shy away from the competition, as shown consistently in laboratory experiments (Niederle and Vesterlund, 2007). Contrary to this expectation, we find no evidence of sorting based on gender ($p = 0.689$); if anything, women appear more likely to submit than men (Table 4, last column), controlling for profession and office location.

Further inspection of the participation rates conditional to the employee's gender suggests relevant treatment interactions based on gender (Fig. 3).

Gender interactions may occur as a result of three factors: differences in risk taking, willingness to contribute to public goods, and competitive inclinations. If women prefer to work on activities that are less risky, more pro-social (e.g., aiming at improving people's health) and where competition is less intense, then we should observe significant treatment interactions. Similarly, one may expect treatment interactions associated with the employee's profession (e.g., the prize opportunity could be a less effective incentive for employees with a higher income, such as doctors).

To isolate gender and profession effects, we employ a version of the model (1) with gender- and profession-treatment interactions (Table 5). Compared to women, men are about 4 percentage points less likely to submit proposals in the PCARE solicitation treatment, controlling for the profession and office location. This effect could be due to gender differences in preferences, as suggested by the literature, and we will return on this topic in the discussion of the results.

4.3. Proposed projects per submission

Thus far, our comparisons have focused on the probability of staff making submissions without consideration of the number of proposed projects per submission. As submitting multiple projects require more effort from the proponent, it is plausible to expect treatment effects on the number of submitted projects as well. Despite this possibility, the median

Table 5
Gender differences.

	Dependent variable:				
	Employee making submissions				
	(1)	(2)	(3)	(4)	(5)
FUND	−3.12***	−3.12***	−2.95***	−2.92***	−3.46***
	(0.99)	(0.99)	(0.99)	(0.99)	(1.07)
PCARE	1.04	1.04	1.00	0.95	1.50
	(1.37)	(1.36)	(1.37)	(1.36)	(1.62)
PRIZE	2.78*	2.77*	2.66*	2.71*	2.11
	(1.45)	(1.46)	(1.45)	(1.45)	(1.72)
genderMale	−0.42	−0.35	−1.40	−0.25	−0.22
	(1.37)	(1.70)	(1.49)	(1.69)	(1.74)
FUND × genderMale	1.66	1.65	1.52	1.56	−0.60
	(1.99)	(2.00)	(1.99)	(2.02)	(2.12)
PCARE × genderMale	−4.57**	−4.59**	−4.50**	−4.47**	−3.75*
	(1.94)	(1.96)	(1.95)	(1.96)	(2.14)
PRIZE × genderMale	−1.20	−1.18	−1.09	−1.19	−2.53
	(2.67)	(2.71)	(2.67)	(2.70)	(3.60)
Constant	5.01***	4.98***	6.74***	8.37***	8.39***
	(0.73)	(0.85)	(1.18)	(1.68)	(1.69)
Office	No	No	Yes	Yes	Yes
Job	No	Yes	No	Yes	Yes
Job × treatment	No	No	No	No	Yes
Log likelihood	−5542	−5542	−5539	−5538	−5535
Observations	1237	1237	1237	1237	1237

Note: Table showing the estimated coefficients (with heteroskedasticity-robust standard errors in parenthesis) for a linear probability model of employee submissions as a function of the solicitation treatment, profession, gender, and whether the employee was assigned to a ward or to a fixed office location, with interactions between the treatment and the gender. Coefficients are multiplied by 100 to indicate a percentage point change in the probability of submitting. Solicitation treatment dummies are coded to indicate deviations from the overall probability of submitting. The asterisks ***, **, * indicate significance at 1, 5 and 10 percent level, respectively.

number of proposals per treatment is one in all treatments. We detect no significant differences across treatments in the number of projects per submission.

4.4. Participation over time

We turn to examine the evolution of the participation rates in each treatment, as they develop over the four weeks of the contest. It is interesting to see whether participation rates tend to converge or diverge over time, as these patterns may reveal something about the underlying communication between employees on the differences in their treatments.

There are multiple dynamics that may arise as employees share information. One is a convergence of participation rates across treatments. For example, people who are not motivated to contribute at the beginning may become motivated and contribute later as they hear about a prize being offered. Another possibility is a divergence of participation rates. People who are intrinsically motivated to contribute at the beginning may become demotivated and decide *not* to contribute as they learn of personal rewards being offered (a kind of "crowding out" in the terminology of Frey, 1997). Although both scenarios are plausible, our data support neither. The treatment differences in participation look pretty "stable" – neither converging nor diverging – over time, with a small tendency to converge only in the last week (Fig. 4). We interpret this result as evidence against strong communication between employees about the differences in their treatments, although we cannot entirely rule out the possibility.

4.5. Participation in peer evaluation

We now turn to examine the outcomes of the peer evaluation that follows the submission period. This phase involves 113 project proposals[10] that are rated by a total of 178 employees (14.4% of our sample) who volunteered for the task. Their effort yields a total of 12,219 evaluator-proposal pairs, providing a very sensitive test for differences in project quality across our solicitation treatments, as we discuss later.

Since evaluating proposals is another way for staff to contribute to the organization, one may expect treatment effects on participation in this phase as well. The proportion of employees participating in the peer evaluation phase is greater in the PCARE and WPLACE solicitation treatments (16% both) than that in the PRIZE (14%) and FUND (12%) solicitation treatments. However, we find no evidence of treatment effects on employees' participation in evaluation: a Fisher's exact test shows no association between the participation rates and the treatment assignment ($p = 0.339$); and a multiple linear regression

[10] Due to a technical problem in uploading the proposals on the website for voting, five proposals ended up with no evaluations. This problem was independent of the treatment: a Fisher's exact test rejects any association between the number of missed proposals and the treatment ($p = 0.7$).

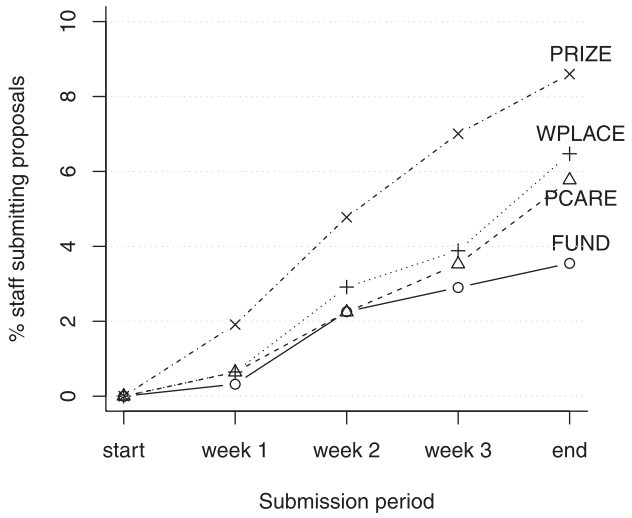

Fig. 4. Participation over time. Plot showing the proportion of staff submitting proposals by treatment assignment as it evolved over the four-week submission period of the internal innovation contest.

model of the probability of participating in peer evaluation that explicitly controls for gender, profession, and office location of the worker also fails to reject the null of the treatment coefficients being jointly zero (F-test, $p = 0.346$).

We further check for sorting to see whether the evaluators are wholly representative of the organization. We test for the statistical significance of coefficients for the profession, gender, and office location in a linear regression on the probability of evaluating proposals (results not shown). Results show no significant differences for the employee's gender and profession albeit significantly higher participation from staff members with an office location. The result mimics the previous findings of participation in the submission phase, reinforcing the idea of a positive sorting of employees with either easier access or more experience and status inside the organization.

4.6. Quality of the project proposals

The treatment interventions may not have only impacted the propensity to make a submission, but the quality of the submission as well. Of particular interest is any indication of a quantity versus quality trade-off. For example, if the treatment which generated the fewest submissions (FUND) also produced the highest quality submissions, a quality-versus-quantity trade-off would increase the complexity of choosing optimal incentives for employees. We examine the issue with the assessments of quality made by peers in the evaluation phase of the contest and, subsequently, by the management.

4.6.1. Quality assessed by peers

To check whether differences in the quality of the submissions can be explained by the solicitation treatments of the submitter, we first look at differences in the distribution of ratings obtained from peers. Overall, a project proposal is given the "neutral" point (i.e., a rating of 3) on a five-point scale about 30% of the times with employees being more likely to give a positive rating (20% of the times the rating is 4 or 5) rather than a negative rating (15% of the times is 1 or 2). This pattern does not change much when we condition the data to the solicitation treatment of the proponent (Fig. 5, right panel); suggesting an equal distribution of good and bad quality projects across the solicitation treatments.

To formally test this hypothesis, we aggregate mean ratings for each proposal (Fig. 5, left panel) and regress these aggregate measures on solicitation treatment dummies. The regression results (not reported) show only an insignificant relationship between ratings and solicitation treatments. The treatment coefficients are all insignificantly different from zero, with the linear model not significantly different from a constant model (F-test, $p = 0.652$).

We also examine the distribution of ratings as generated by treatments with no aggregation (Fig. 5, right panel).[11] We have over 12,000 ratings, providing a very sensitive test for differences across treatments. Using a Chi-square test we find that the hypothesis of dependence between the distribution of ratings and the treatments is insignificant, although close to the 10% level (p-value of 0.103). Driving the p-value is a less than 2% difference between the proportion of 5 ratings in the WPLACE treatment versus the other distributions, which is probably due to outliers (the winning proposal was in the WPLACE treatment). Taken together with the fact that our sample is large, we have strong evidence suggesting that there are

[11] The analysis on the aggregated data crucially relies on the assumption that an increment in a proposal's quality as measured by an increase in ratings from v to $v+1$ is the same for any value v.

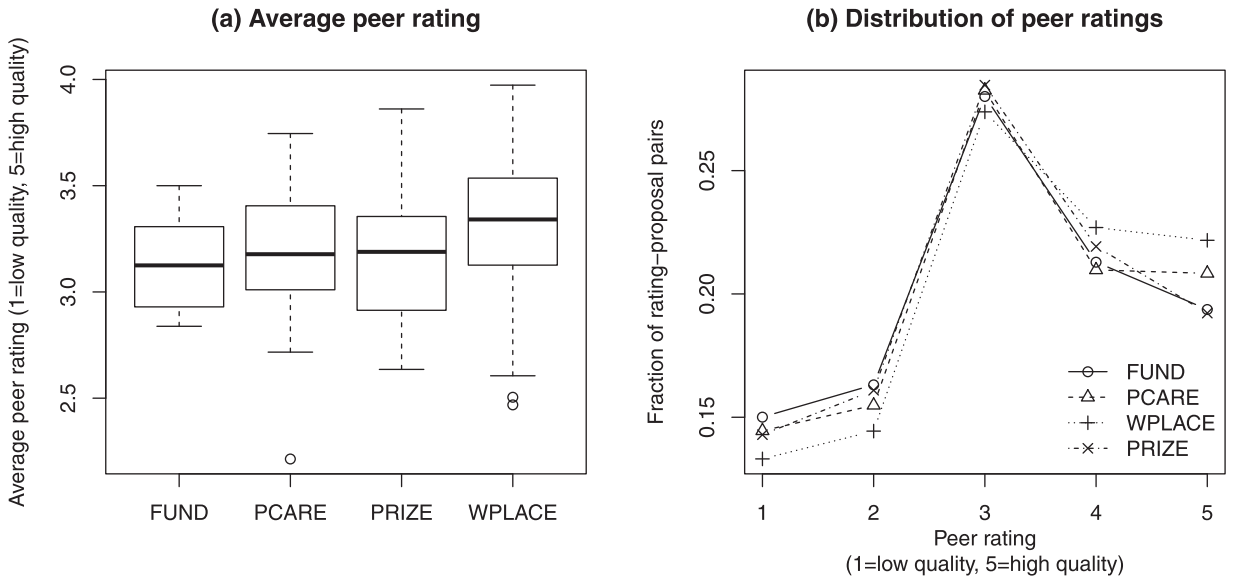

Fig. 5. Differences in the quality of the proposals by treatment assignment as assessed by peer staff members. The left panel shows boxplots of the mean rating for each treatment. The right panel shows the proportions of ratings received by a proposal in a given treatment. Both panels show differences that are statistically insignificant (see the main text).

no (economically meaningful) differences in the quality of project proposals across treatments and in particular no evidence of a quantity versus quality trade-off up to the resolution of the five-point scale.[12]

4.6.2. Quality assessed by managers

One potential limit of assessing quality from the peer ratings alone is that employees might have a different view of a proposal's quality than executives (due, for instance, to a misalignment of incentives). Indeed, to ensure alignment between managerial goals and the peer assessment, all project proposals have been further vetted by the HTL staff before being considered for implementation funding. We now focus on the outcomes of this vetting process to investigate more broadly the presence of treatment effects on the quality of project proposals.

The vetting process conducted by the HTL staff resulted in 113 proposals being scored on a scale from 1 to 100 points with 29 authors of the best proposals being invited to submit implementation plans. The remaining proposals were excluded (and received a score of zero) either because flagged as inappropriate for funding or because the proponent manifested no intention to participate in the implementation phase (we find no association between proposals excluded and treatments).

As a broad measure of agreement between peer ratings and the assessment made by the management, we use the Spearman's rank correlation coefficient between the scores given by the HTL staff and the average peer ratings. Results show a positive correlation (cor = 0.198), indicating good agreement between our two measures of quality.[13] As for the assessments made by peers, using the scores by the management shows no treatment effects on quality: a linear regression on the score of proposals against treatment dummies gives coefficients that are all insignificantly different from zero, with the linear model not significantly different from a constant model (F-test; $p = 0.378$).

Being selected and invited to submit additional implementation plans is another indication of quality assessed from the management perspective. A simple test for independence between the percentage of submitters being selected and invited by HTL staff and the randomly assigned solicitation treatment gives no significant results (a Fisher's exact test for Count Data gives a p-value of 0.314). Although insignificant, employees who made project proposals in the FUND solicitation treatment are *less* likely to be selected as finalist than the others (only 1 out of 7 in the FUND treatment were selected and invited by the HTL staff), providing additional evidence of a no quantity versus quality trade-off, as discussed before.

4.6.3. Quality assessed by higher word counts

Following prior research indicating that higher total word counts reflect higher quality (Blumenstock, 2008), we also look at differences in the word counts of a submission. We find that most submissions are below 200 words with little

[12] One may worry that such binning is a fairly coarse measure of quality. In particular, effects concentrated in the upper tail of the distribution may not be detected. For example, comparing the ratings of proposals A, B, C and D with hypothetical true qualities of 3, 4, 5, and 10 stars respectively. Under a five-point scale rating system, proposals A and B can be distinguished, but C and D cannot be distinguished. Hence, one needs to be very cautious in interpreting these results as evidence against quality effects in general.

[13] The managerial evaluation is blind to the results of the peer evaluation (all the peer ratings were stored on a secure server with exclusive access by the researchers) and, therefore, the correlation is not automatically influenced by the opinions of the staff.

differences between the treatments. Testing for a significant linear regression relationship between the length of submissions and treatment dummies returned an overall insignificant result (F-test, $p = 0.43$), which is consistent with our previous assertion of little, or no, differences in quality across treatments.

4.7. Content of the project proposals

While we have shown little differences in the overall project quality across solicitation treatments, it is easy to think of ways in which a solicitation may affect the *content* of the submitted proposals, while keeping the quality constant. For example, different solicitations may lead proponents to think about different kinds of problems or, indirectly, sort proponents with differing needs or knowledge of the problems inside the organization.

Before examining specific differences in the content of proposals, it is important to point out that the proposed projects broadly conform to the stated goals of the contest, which is to improve Heart Center operations by identifying problem areas and potential solutions. For example, one highly voted project proposal is to build a platform to electronically review and update patients' medicine list in the office prior to seeing the physician. Another is to develop a smartphone application showing a patient's itinerary for the day providing a guide from one test or appointment to another. This suggests the aligning with improving the work processes within the organization or providing high-quality patient care. Nevertheless, other contest organizers may have varying goals and be concerned about different aspects of the submissions.

To examine additional dimensions of submission content, we study the *area of focus* of the submissions. Of particular interest is understanding whether different wordings used in the general encouragement solicitation (either towards improving the workplace or targeting the wellbeing of patients) induce employees to concentrate on different categories of interventions. Members of the HTL categorize each project proposal into one of seven areas of focus: three categories ("Care coordination", "Staff workflow", "Workplace") identify improvements for the workplace, three ("Information and access", "Patient care", and "Quality and Safety") focus on improvements centered around patients, and one ("Surgical tools and support to research") categorizes projects developing tools to support scientific research.

The proportions of submitted project proposals in each area of focus (Fig. 6) exhibit very similar patterns for the WPLACE and PRIZE solicitation treatments and different and uncorrelated patterns for the other treatments; suggesting the presence of treatment effects. We test to determine the overall association between these proportions and our solicitation treatments with a Fisher's exact test for Count Data with simulated p-value (based on 2000 replicates). Results show a significant ($p = 0.093$) association at the 90% level, providing evidence that the variation in the content of the submitted proposals is, at least in part, attributable to our experimental intervention.

To test which areas of focus is affected by our treatment, we regress the probability of a project proposal being in a given category against solicitation treatment dummies. We use an F-test where the null hypothesis tested is that all the treatment effects have a zero effect on the probability of the proposal being in a given category. The results show that project proposals in the PCARE solicitation treatment are less likely to fall in the "Quality and Safety" category, and project proposals in the FUND solicitation treatment are less likely to fall in the "Information and access" category.

Although it is difficult to interpret these results without any theoretical model guiding our predictions, they indicate a possible trade-off between stimulating participation via personalized solicitation messages and the effect of those messages on the type of contributions to the public good, which complicates the analysis of incentives for public goods inside organizations beyond what the current literature anticipates.

5. Limitations

As the literature on the identification of treatment effects in randomized experiments has long recognized (Imbens and Rubin, 2015), the validity of our causal interpretation of the results rests on a few conventional assumptions. These include the "no interference between units" assumption (Imbens and Rubin, 2015, p. 10), which is a potential limitation of this study.

It is indeed possible, and we cannot exclude, that communication among staff members assigned to different treatment arms could have influenced decisions to participate, thus violating the no interference assumption. However, the magnitude of the interference effects would depend on the intensity of staff communication and the density of social interactions. Both of which should be small because (i) an individual competition may provide only weak incentives for information sharing and (ii) the staff members are scattered across multiple buildings on the hospital campus. Moreover, HTL managers did not report any concerns for spillovers and, by looking at the temporal dynamics of submissions, we find no indication of either divergence or convergence in the participation rates, which can be interpreted as an indication of communication among employees. Hence, the assumption of no interference seems appropriate.

Another potential limitation is to have to deal with large amounts of noncompliance because staff members may have left unopened the solicitation email announcing the contest, thus non-complying with the assigned solicitation treatment. As this kind of noncompliance is almost entirely unobserved,[14] the analysis follows an intention-to-treat approach – it focuses on the causal effect of the *assignment* rather than the causal effect of the *receipt of the treatment* (the solicitation

[14] The email was sent using the internal messaging system of the Heart Center, which, at the time, was not collecting individual analytics.

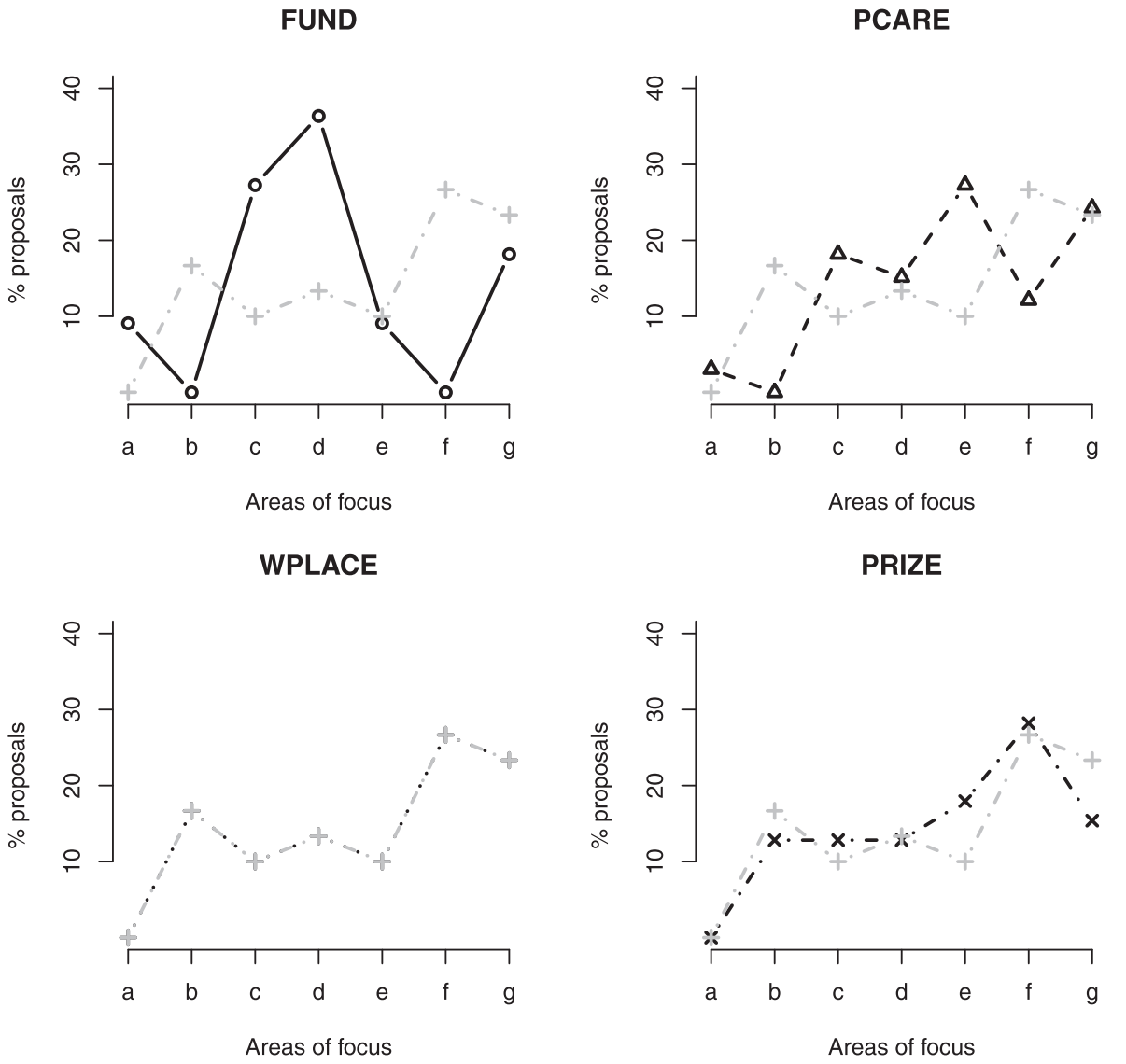

Fig. 6. Panels showing the proportion of submitted project proposals in each area of focus (a = Surgical tools and support to research, b = Quality and safety, c = Workplace,d=Staff workflow, e = Care Coordination, f = Information and access, g = Patient support) for each solicitation treatment. The proportion of proposals in the WPLACE treatment are used as a reference in all panels.

itself). Consequently, our interpretation of the results lies on the additional assumption that the effect of the assignment is indicative of the effect of the treatment.

6. Conclusion

Our analysis shows that prize incentives boost individual participation without lowering the quality of the proposals. The positive effect on participation is large and goes beyond what one might expect from the value of the offered reward alone. This evidence suggests that workers partially internalize the positive effects of their actions (submitting a project proposal) on the other individuals in the workplace. Indeed, we can exclude that the boost is due to opportunistic behavior or a negative selection, as the quality of the proposals is as high as in the other treatments. In addition, the overall effect appears *too large* to be consistent with a standard incentive effect of tournaments (Ehrenberg and Bognanno, 1990).[15]

[15] The cost of submitting is proportional to the hourly wage of the median staff member, which is $40 (the median earning per hour of a nurse according to the Bureau of Labor Statistics) and, abstracting from the non-monetary aspects, the expected value of being awarded an iPad is proportional to the retail price of the iPad ($350) times the probability of winning, which is relatively small in a contest that involves more than 1000 potential participants.

Consequently, unlike Bandiera et al. (2005) that shows a case in which prize-based compensation schemes have a negative impact on effort and firm performance, our study shows a beneficial outcome for the firm (increasing the number of project proposals without lowering the quality). From a theoretical point of view, this positive result can be explained with a very simple model of lottery-like contests to finance public good provision inspired by Morgan (2000). Intuitively, the internal competition introduces two opposing externalities (one negative externality associated with the prize competition and the positive externality of contributing to public goods) that mitigate the free-riding problem, thus increasing the equilibrium provision.

Our results also suggest that gender may be an important factor influencing the sensitivity of responses to solicitations concerning the organizational mission, with women's participation being greater when emphasizing patient care than men's. The result is robust and does not seem to be caused by the profession and status within the organization of the solicited worker. One possible explanation is that male and female staff members hold different attitudes towards the opportunity of helping their patients, which seems unlikely. Instead, we believe that gender stereotypes might have shaped the responses of the employees as if men viewed the task as falling in a female-type domain. This kind of self-selection based on stereotypes has been shown in other contexts. Coffman (2014), for example, finds that women are less likely to contribute ideas to groups when the topic falls in male-typed domains, e.g., sports, and vice versa.

Another result concerning sorting that deserves consideration is the finding that women's participation in the prize treatment is slightly higher, but not significantly so than men's. This result is robust to our controls for profession and status. So, it indicates that gender differences in competitive inclinations or risk aversion, if any, might have exerted only a little influence on our sample, which is reassuring given the risk of producing a gender imbalance in the type of projects submitted.

We also find that encouraging participation by offering individual funding to implement proposals has a negative effect on participation. Although the possibility of winning project funding implies a tangible prize, unlike appealing to altruistic motives as in improving patient care and workplace, the funding treatment induced the least participation in both submitting and evaluating proposals. To understand this interesting finding, we interviewed Heart Center leadership and staff and learned two possible explanations. First, unlike winning a one-time prize like the iPad, winning project funding implies extra work and responsibility that people are not willing or able to take on. A senior leader shared, "The whole point of a grant is that at the end, you have to write up a report and show that you did it. It's not just talk. You're obligated.... Ideation is fun. Adding more work to your busy schedule is hard and so people think, ugh, I don't really want to do it.... In the future, if we try that experiment again but said, you'll get project funding and the support of a team who will take your idea and run with it – you will serve as an advisor – I bet people would say, oh, yeah, that's okay." This offers contrary perspective to the conventional view that providing opportunities to contribute will elicit higher levels of workplace engagement.

Second, although we conducted our experiment in an academic hospital where clinicians and researchers routinely apply for grant funding, our informants shared that many employees (like medical assistants, technicians, and administrative assistants) are not familiar with the process of acquiring and managing grants and that some would even find it daunting. A nurse, who won the contest and led the implementation of her idea, explained, "I think [the prospect of winning project funding] can be scary and it would discourage people from submitting ideas. They'd think, if I win then I'll have to lead this big initiative that's worth $20,000". She also added that seeing a large dollar value in idea solicitation can be "intimidating," or difficult to "wrap your mind around what you'd do with that funding and be in charge of this big project, unless you have an idea that's really tangible and fully formed, like an app or some technology."

From the perspective of the management, these results have several implications for the design of incentives for public goods inside organizations. First, prize incentives raise participation without lowering quality and the size of the effect appears to be driven by complementarity between the value of the conferred prize and the employees' preferences towards the organization. Second, participation is sometimes triggered by mission-based preferences, but these preferences may also create gender-based selection effects that are difficult to predict ex-ante. We point to the extensive literature on a gender-based difference in preferences (see Croson and Gneezy, 2009, for a review) as a potential way to explain some of these effects. Third, offering workers the opportunity of managing resources to contribute to collective goods may backfire.[16]

Finally, while the choice of focusing on health care workers may limit the generalizability of our results in some respects, we believe they extend beyond the specific organization under study. It should be noted that in the US alone health care spending accounts for 18% of the GDP (in 2016).[17] There are also other professions, such as teachers, public servants, academic researchers, who are routinely exposed to public good dilemmas and in mission-driven environments. In all these settings, our study suggests that internal contests soliciting employee contributions and awarding an individual prize to the winning contribution appear an effective strategy.

Therefore, a selfish risk-neutral person who expects to work 10 hours on the execution of the project and has no other goal than maximizing utility would need to be almost certain to win the iPad in order to make a submission.

[16] One may argue that this effect could be stronger for medical workers who have opportunities to satisfy their mission-oriented preferences (e.g., helping patients) and may be at a point of satisfaction compared to other workers.

[17] This figure is taken from the "Statistics Trends and Reports" section of the Centers for Medicare and Medicaid Services website (www.cms.gov).

Acknowledgments

We gratefully acknowledge the financial support of the MacArthur Foundation (Opening Governance Network), NASA Tournament Lab (Cooperative Agreement #NNX16AC77A), and the Harvard Business School Division of Faculty Research and Development. This project would not have been possible without the support of Eric Isselbacher, Julia Jackson, Maulik Majmudar and Perry Band from the Massachusetts General Hospital's Healthcare Transformation Lab. We also thank Margaret McCleary, Catharine Wetzel, and Randall Zusman for generous sharing of time and thoughts in an interview.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jebo.2019.02.029.

References

Akerlof, G.A., Kranton, R.E., 2005. Identity and the economics of organizations. J. Econ. Perspect. 19 (1), 9–32.
Andreoni, J., 1990. Impure altruism and donations to public goods: a theory of warm-glow giving? Econ. J.
Andreoni, J., Vesterlund, L., 2001. Which is the fair sex? Gender differences in altruism. Q. J. Econ. 116 (1), 293–312.
Ashraf, N., Bandiera, O., Lee, S.S., 2014. Awards unbundled: evidence from a natural field experiment. J. Econ. Behav. Organ. 100, 44–63.
Bandiera, O., Barankay, I., Rasul, I., 2005. Social preferences and the response to incentives: evidence from personnel data*. Q. J. Econ. 120 (3), 917–962.
Bandiera, O., Barankay, I., Rasul, I., 2008. Social capital in the workplace: evidence on its formation and consequences. Labour Econ. 15 (4), 724–748.
Bandiera, O., Barankay, I., Rasul, I., 2010. Social incentives in the workplace. Rev. Econ. Stud. 77 (2), 417–458.
Bandiera, O., Barankay, I., Rasul, I., 2011. Field experiments with firms. J. Econ. Perspect. 25 (3), 63–82.
Barankay, I., 2012. Rank Incentives: Evidence from a Randomized Workplace Experiment. Discussion papers.
Besley, T., Ghatak, M., 2005. Competition and incentives with motivated agents. Am. Econ. Rev. 95 (3), 616–636.
Blumenstock, J.E., 2008. Size matters: word count as a measure of quality on Wikipedia. In: Proceeding of the 17th International Conference. ACM Press, New York, New York, USA, p. 1095.
Boning, B., Ichniowski, C., Shaw, K., 2007. Opportunity counts: teams and the effectiveness of production incentives. J. Labor Econ.
Boudreau, K.J., Lacetera, N., Lakhani, K.R., 2011. Incentives and problem uncertainty in innovation contests: an empirical analysis. Manag. Sci.
Boudreau, K.J., Lakhani, K.R., Menietti, M., 2016. Performance responses to competition across skill levels in rank-order tournaments: field evidence and implications for tournament design. RAND J. Econ.
Carpenter, J., Gong, E., 2016. Motivating agents: how much does the mission matter? J. Labor Econ. 34 (1), 211–236.
Che, Y.-K., Yoo, S.-W., 2001. Optimal incentives for teams. Am. Econ. Rev.
Coffman, K.B., 2014. Evidence on self-Stereotyping and the contribution of ideas. Q. J. Econ. 129 (4), 1625–1660.
Cohen, J., 1992. A power primer. Psychol. Bull. 112 (1), 155.
Croson, R., Gneezy, U., 2009. Gender differences in preferences. J. Econ. Lit. 47 (2), 448–474.
Cutler, D., Wikler, E., Basch, P., 2012. Reducing administrative costs and improving the health care system 367 (20), 1875–1878.
Delfgaauw, J., 2005. Dedicated doctors: public and private provision of health care with altruistic physicians. N. Engl. J. Med. 353 (15).
Delfgaauw, J., Dur, R., 2008. Incentives and workers' motivation in the public sector*. Econ. J.
Delfgaauw, J., Dur, R., Sol, J., Verbeke, W., 2013. Tournament incentives in the field: gender differences in the workplace. J. Labor Econ.
Della Vigna, S., List, J.A., Malmendier, U., Rao, G., 2016. Estimating Social Preferences and Gift Exchange at Work. Technical Report.
Drago, R., Turnbull, G.K., 1988. The incentive effects of tournaments with positive externalities among workers*. South Econ. J.
Ederer, F., Manso, G., 2013. Is pay for performance detrimental to innovation? Manag. Sci. 59 (7), 1496–1513.
Ehrenberg, R.G., Bognanno, M.L., 1990. Do tournaments have incentive effects? J. Polit. Econ.
Erev, I., Bornstein, G., Galili, R., 1993. Constructive intergroup competition as a solution to the free rider problem: a field experiment. J. Exp. Soc. Psychol. 29 (6), 463–478.
Frey, B.S., 1997. Not Just for the Money: An Economic Theory of Personal Motivation. Edward Elgar.
Gallus, J., 2016. Fostering public good contributions with symbolic awards: A Large-Scale natural field experiment at wikipedia. Manag. Sci.
Gibbs, M., Neckermann, S., Siemroth, C., 2015. A field experiment in motivating employee ideas. Rev. Econ. Stat.
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. J. Polit. Econ.
Hellmann, T., Thiele, V., 2011. Incentives and innovation: a multitasking approach. Am. Econ. J.
Holmström, B., Milgrom, P., 1991. Multitask principal-Agent analyses: incentive contracts, asset ownership, and job design. J. Law Econ. Organ.
Imbens, G.W., Rubin, D.B., 2015. Causal Inference for Statistics, Social, and Biomedical Sciences, first ed. Cambridge University Press.
Jonny, E., 2013. "Apple seeks 'Pie in the sky' ideas for innovation". Computerworld. Accessed March 8, 2019. https://www.computerworld.com/article/2474058/apple-seeks--pie-in-the-sky--ideas-for-innovation.html.
Knoeber, C.R., Thurman, W.N., 1994. Testing the theory of tournaments: an empirical analysis of broiler production. J. Labor. Econ.
Lazear, E.P., Rosen, S., 1981. Rank-order tournaments as optimum labor contracts. J. Polit. Econ. 89 (5), 841–864.
Leon, K., 2010. Xerox employees' green ideas save company $10.2 Million. The Guardian. Accessed March 8, 2019. https://www.theguardian.com/sustainable-business/xerox-employees-green-ideas-save.
Mary, O., Viscusi, W.K., Zeckhauser, R.J., 1984. Economic contests: comparative reward schemes. J. Labor Econ.
Mas, A., Moretti, E., 2009. Peers at work. Am. Econ. Rev. 99 (1), 112–145.
Morgan, J., 2000. Financing public goods by means of lotteries. Rev. Econ. Stud.
Niederle, M., Vesterlund, L., 2007. Do women shy away from competition? Do men compete too much? Q. J. Econ. 122 (3), 1067–1101.
Prendergast, C., 2007. The motivation and bias of bureaucrats. Am. Econ. Rev. 97 (1), 180–196.
Prendergast, C., 2008. Intrinsic motivation and incentives. Am. Econ. Rev. 98 (2), 201–205.
Rachel, K., 2014. AT&T Develops Employee Ideas for Innovation. The Wall Street Journal. Accessed March 8, 2019. https://blogs.wsj.com/cio/2014/11/12/att-develops-employee-ideas-for-innovation/.
Vesterlund, L., 2012. Voluntary giving to public goods: moving beyond the linear VCM. The Handbook of Experimental Economics.
Vidal, J.B.i., Nossol, M., 2011. Tournaments without prizes: evidence from personnel records. Manag. Sci.