NBER WORKING PAPER SERIES

MOTIVATING EFFORT IN CONTRIBUTING TO PUBLIC GOODS INSIDE ORGANIZATIONS: FIELD EXPERIMENTAL EVIDENCE

Andrea Blasco
Olivia S. Jung
Karim R. Lakhani
Michael Menietti

Working Paper 22189
http://www.nber.org/papers/w22189

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2016

We gratefully acknowledge the financial support of the MacArthur Foundation (Opening Governance Network), NASA Tournament Lab, 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. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Motivating Effort In Contributing to Public Goods Inside Organizations: Field Experimental Evidence

Andrea Blasco, Olivia S. Jung, Karim R. Lakhani, and Michael Menietti

NBER Working Paper No. 22189
April 2016
JEL No. D03,D23,H41

ABSTRACT

We investigate the factors driving workers' decisions to generate public goods inside an organization through a randomized solicitation of workplace improvement proposals in a medical center with 1200 employees. We find that pecuniary incentives, such as winning a prize, generate a threefold increase in participation compared to non-pecuniary incentives alone, such as prestige or recognition. Participation is also increased by a solicitation appealing to improving the workplace. However, emphasizing the patient mission of the organization led to countervailing effects on participation. Overall, these results are consistent with workers having multiple underlying motivations to contribute to public goods inside the organization consisting of a combination of pecuniary and altruistic incentives associated with the mission of the organization.

Andrea Blasco
Harvard Institute for Quantitative Social Science
1737 Cambridge St
Cambridge, MA 02138
mrblasco@gmail.com

Karim R. Lakhani
Harvard Business School
Morgan Hall 433
Soldiers Field Road
Boston, MA 02163
and NBER
klakhani@hbs.edu

Olivia S. Jung
Harvard Business School
10 Soldiers Field Road
Boston, MA
ojung@hbs.edu

Michael Menietti
1737 Cambridge Street
Cambridge MA 02138
mmenietti@fas.harvard.edu

1 Introduction

Public goods problems are ubiquitous inside of organizations. Employees are often required to make contributions that go beyond an effective and efficient execution of their own assigned tasks, such as those towards improving the overall operations and performance of the shared enterprise. For example, professors are expected to teach and publish research (with pay, promotion, and tenure tied to performance in those activities) and also to serve on various internal committees that directly benefit the university and department operations, with indirect benefits to committee members.[1] Similarly in companies, employees are expected to work on production activities and also to collaborate in teams, contribute to common resources, provide feedback on strategy and direction of the firm, and drive innovation efforts.

A central question for many organizations that rely upon these types of public goods to increase productivity is determining what motivates their members to participate. In this study we compare two perspectives on this question: the canonical economic approach that presumes agents have no social preferences and they will free-ride unless given appropriate pecuniary incentives; and a behavioral approach that treats participation as a function of internal motivations towards improving the organization, and thus presumes some form of social preferences or mission orientation towards others (e.g., organization members, employers, end users).

There has been a longstanding interest in this problem by economists. This interest has led to a large body of theoretical works directing much of their attention to the free rider problem, especially focusing on team production (e.g., Holmstrom, 1982; Itoh et al., 1991). In organizations, as well as in teams, agents are asked to perform jobs that generate benefits accruing to the whole organization, but that may not benefit them individually. Personal rewards based on performance (e.g., prizes, reputation) will help agents better internalize the effects of their own actions on others. But the incentives to free-ride typically remain, given the limited ability of organizations to contract and monitor performance.[2]

An equally important stream of research in the area of public goods has empirically shown that people voluntarily contribute to public goods, despite strong incentives to free

[1] Professors serve committees pertaining to student experience, university programs, personnel and staff, seminars, curriculum, etc. According to Neumann and Terosky (2007), professors get less recognition for their contribution to these types of faculty service than for research and teaching; administrators and professors tend to view contributing to this type of institutional good as a less meaningful and important activity.

[2] For example, it can be hard to define an optimal level of information sharing or to monitor an individual's effort towards improving the organization. But even if contracts are permitted, incentives may not fully solve the problem, given the presence of multitasking.

ride. This occurs in various real-life situations (e.g., charities, blood donation, social work).[3] It is also observed in anonymous, single-shot public good games in the laboratory (e.g., Marwell and Ames, 1981). Building on this insight, many studies have adopted a different perspective on what makes people contribute to public goods inside of organizations. Agents may decide to contribute because of internal preferences towards improving the organization, such as the personal satisfaction derived from driving value to its customers (Delfgaauw, 2005; Delfgaauw and Dur, 2008; Prendergast, 2007). These types of altruism recur in organizations for social public goods, such as hospitals, universities, schools, administrations, and the army. Sometimes this is also referred to as identity or internal preferences for the organization's mission (Akerlof and Kranton, 2005; Besley and Ghatak, 2005). It is also possible that agents derive direct utility from contributing to the wellbeing of the other members of an organization, such as their co-workers (Bandiera et al., 2005).

Lying in between these two perspectives is the work of Morgan (2000) on financing public goods through lotteries. Morgan (2000) develops a model to study the role played by a competition for fixed prizes on individual choices of participation and contribution towards a public good. The model shows that the mechanism of internal competition associated with prizes may countervail the positive externality of contributing, thereby alleviating under-provision. This result suggests a complementarity between the intensity of pecuniary incentives and internal preferences toward improving the organization.[4]

Despite a large theoretical literature on the topic, there is still limited empirical evidence about the motivations of agents to contribute to public goods inside of organizations. Laboratory experiments on public good games suggest that participation might be driven by some form of impure altruism (see Ledyard, 1997; Vesterlund, 2012), but these results may not apply inside of organizations. Empirical studies on team production show that individual productivity in teams reacts to pecuniary incentives (see Prendergast, 1999), but these do not

[3]The World Health Organization estimates 108 blood donations collected globally each year mostly from voluntary unpaid blood donors. According to List (2011), charitable gifts of money are worth 2 percent of gross domestic product for the United States. Lacetera et al. (2014) further reports that "27% of Americans volunteer with formal organizations, for a total of about 8 billion hours per year."

[4]Large retail companies, such as Apple or Levi Strauss, routinely promote company-wide competitions to gather new inputs from employees on how to improve their stores (Source: *Apple seeks 'Pie in the sky' ideas for innovation*, Computerworld, 2013; *Superusers turned programmers*, InfoWorld, 1985). AT&T has reportedly spent over $44 million funding between 2009 and 2014 funding projects generated by setting up a web portal for running periodic competitions for ideas (Source: *AT&T Develops Employee Ideas for Innovation*, Wall Street Journal, 2014). Similarly, IBM sponsors periodic internal competitions ("jams") to promote the process of innovation and collaboration (Source: *An Inside View of IBM's 'Innovation Jam'* MIT Sloan Management Review, 2008).

usually allow distinguishing between self-interest and altruistic preferences.[5] Similarly, the theory of Morgan (2000) has been investigated in the laboratory (e.g., Morgan and Sefton, 2000; Lange et al., 2007) and in the field of charitable giving (Landry et al., 2006), but not inside organizations. Lacking enough empirical evidence, many important questions are still open. What drives contributions to public goods inside organizations? Will organizations be better off by offering pecuniary incentives? Or should they rather focus on workers' social preferences and encourage voluntary contributions? What is the net outcome of incentives in terms of increased contributions compared to non-pecuniary incentives alone?

To make progress towards our research question, we conducted a randomized field experiment testing the relative effectiveness of pecuniary prizes and internal motivations to contribute effort in a public good among over 1,200 employees of the Massachusetts General Hospital's Corrigan Minehan Heart Center (hereafter "Heart Center"), a prominent medical organization in the United States. The experiment uses a call for proposals run at the Heart Center where employees at all levels (e.g., physicians, nurses, administrative staff, technicians) were solicited to identify project ideas addressing common problems in their workplace. Executives from the Heart Center also made the commitment that the best proposals, as judged by peers and hospital managers, would be implemented over the coming year.

The Heart Center executives worked closely with us in the communication strategy for the call for proposals, and we designed four treatments related to our research questions: a control introduced the call for proposals as a contest with no pecuniary prizes to be won, a second explicitly offered individual pecuniary prizes for the best proposals, a third appealed to the internal motivation of the employees by emphasizing the patient care mission of the organization, and a fourth emphasized improvement in the work processes of the organization.

Our experimental intervention changed the communication regarding the goals of the call for proposals among randomly selected staff members. By doing so, we were able to obtain causal estimates of the effect of pecuniary and non-pecuniary incentives on two main outcomes: (a) the decision to submit a proposal and (b) the quality of the proposal submitted (as measured by 12,000 ratings made by peers). Using data on profession and gender, we then characterized the heterogeneity of responses to the treatment and evaluated the implications

[5]There are several field experiments about production in teams at the firm level (e.g., Erev et al., 1993; Bandiera et al., 2013). While these studies focus on individual productivity, here we instead consider a more general setting of individuals contributing to a common resource (e.g., a common repository, a cost-saving idea) that may affect the productivity of the whole organization and potentially drive more value for customers.

of the intervention for the organization.

In all treatments, individuals could anticipate more costs and responsibilities from winning the competition, such as providing guidance on the implementation of their proposal, as the Heart Center executives made it clear that there would be special implementation teams for the proposals. They could also expect no direct benefits in terms of promotion or salary increases. Thus, the control treatment properly isolates underlying intrinsic motivations to work, such as prestige, peer recognition and altruism, in the absence of which a classical model would predict no participation. A comparison of participation rates between the control and the prize treatment will then identify differences in incentives exclusively due to pecuniary prizes to be won in the contest. The two framing conditions serve, therefore, to check the relative effectiveness of intrinsic motivations aligned with the mission of the organization compared to other incentives.

We find that the treatment introducing the call for projects as a contest for winning a pecuniary prize generates a significant increase in submissions compared to the control. Further, the mean quality of proposals was comparable to the submissions in the other treatments. Our data provide no evidence of a trade-off between the number of submissions and quality. We also find that winning a prize is not the only motivation to contribute. Participation was greater under the two interventions appealing to intrinsic motivations than under the control. In particular, emphasizing improvement in the work processes of the organization significantly increases participation (at a 10 percent significance level), while emphasizing the patient care mission of the organization has countervailing effects on participation. Controlling for profession, women's participation is significantly greater when emphasizing the patient care mission compared to the control whereas men's participation is significantly lower.

We unpack these differences with the help of an economic model of public goods, in the spirit of Morgan (2000). According to our model, the effect of social preferences on the probability of contributing with a submission is comparable to a 45 percent reduction in the cost of effort. This result is complementary to the evidence on social preferences provided by Della Vigna et al. (2016). Della Vigna et al. (2016) study freelance workers to show that social preferences towards their employers may affect the level of effort at work. Here, we show that the same may happen for workers inside of an organization. Compared to Della Vigna et al. (2016), our estimated effect appears larger, which may be attributed to the combined consequences of the organization pursuing a social goal (e.g., providing healthcare) and the individuals being members of an organization, rather than freelance workers.

Using the model, we also find evidence of sizable non-pecuniary incentives (e.g., repu-

tation, peer recognition) in contributing to public goods in organizations, other than just social preferences or pecuniary prizes. This evidence is consistent with the empirical work of Kosfeld and Neckermann (2011) and Blanes i Vidal and Nossol (2011) on congratulatory rewards and social status. Here, the main difference is that reputation or social status may be acquired through a public good contribution to improving the organization rather than on the basis of differences in individual productivity or skills.

Overall, this study shows that winning a prize can foster individual participation and contributions to a public good even among motivated agents. While the precise estimates will not apply to every setting, these results are remarkable because they support the hypothesis of a complementarity between the use of pecuniary incentives and the degree of social preferences.

The rest of the article is organized as follows. The next section summarizes the relevant literature followed by hypotheses that are tested in our field experiment. Section 4 introduces the context and the experimental design. Results are presented in Section 6. Section 7 concludes the paper by giving an interpretation of the findings and a discussion of the main assumptions.

2 Literature

Our work is closely related to the literature that analyzes the problem of moral hazard in team production (Holmstrom, 1982; Itoh et al., 1991). In this literature, multiple agents are engaged in several production activities. Some tasks are independent of others, whereas others are closely intertwined and serve to help everyone be more productive. Given an organization's imperfect ability to monitor individual effort, pay-for-performance schemes may not provide an effective incentive for workers to execute the collective tasks. Possible solutions include relative compensation schemes, such as internal contests (e.g., Erev et al., 1993), incentives based on group performance (Holmström and Milgrom, 1990), design of specialized task structures to encourage teamwork (Itoh et al., 1991), and reputation and peer monitoring (Che and Yoo, 2001).

Our study contributes to the existing literature by empirically assessing the relative effectiveness of incentives provided by an internal contest to raise individual participation in improving the organization, instead of a focus on performance in a specific task. Using existing literature on individual motivations at work (Gibbons, 1998; Prendergast, 1999; List and Rasul, 2011), we argue that the incentives to participate in such a competition can be grouped into two main dimensions: the traditional distinction between pecuniary

and non-pecuniary psychological motives (Deci and Ryan, 1985; Ryan and Deci, 2000) and another distinction between the private and public nature of economic incentives. By private incentives, we simply mean rewards enjoyed exclusively by the winner of the competition that can be either pecuniary (e.g., a sum of money, a wage increase) or non-pecuniary (e.g., an increase in reputation or prestige). Public rewards, by contrast, are not exclusive and can be enjoyed by more than one participant at the same time (e.g., the solution of a vexing problem may grant fame and recognition to the problem solver and, at the same, enable other individuals to enjoy benefits from the discovered solution).

These types of incentives have long been investigated in different fields. A vast literature in public economics has considered the issues generated by the non-exclusive nature of certain goods (see Vesterlund, 2012, for an up-to-date review of the literature), stressing the risks of under-supply of private provisions. The public and private incentives associated with prize competitions have been examined in the contest theory literature (Lazear and Rosen, 1981; Green and Stokey, 1983; Mary et al., 1984; Moldovanu and Sela, 2001). However, most contest theory models focus on rewards that are privately enjoyed by the winners without necessarily distinguishing between pecuniary or non-pecuniary private rewards.[6]

The model by Morgan (2000) on the use of fixed-prize competition for financing public goods is one case in which the role of private prizes and public incentives is discussed in the same model. In this particular model, mixing private and public incentives in a lottery can be an effective way to coordinate individual actions and, thus, prevent problems of free-riding. Based on this insight, several studies have found evidence consistent with a positive effect of contests on public good contributions in the laboratory (Morgan and Sefton, 2000; Dale, 2004; Lange et al., 2007) and on individual charitable giving in the field (Landry et al., 2006).

The first test of the effectiveness of fixed-prize competition on free-riding in a real setting was conducted by Erev et al. (1993), involving small teams of orange pickers. They find evidence supporting a positive effect of a fixed-prize competition on participation and effort. However, the result may not extend to more complex organizations or different types of tasks. In particular, in most organizations, agents respond to monetary incentives. However, in organizations that are centered around a mission, individuals may also be motivated to work and exert additional effort because they care about achieving the goal of the organization. This is the case, for example, of organizations engaged in a public service, such as hospitals, schools, public administration, and the military. Several studies have investigated the role

[6]Important exceptions are those models that presume individuals have preferences for social status (e.g., Moldovanu et al., 2007), other than just the value of the prize awarded to the winners.

of peer recognition and status to motivated workers (Auriol and Renault, 2008; Besley and Ghatak, 2008; Delfgaauw, 2005; Delfgaauw and Dur, 2008), emphasizing how organizational mission may impact motivation.

Our study also builds on a large literature focused on the role of framing manipulations on economic behavior (Rabin, 1998), including the effects of positive framing on the private provision of public goods (Andreoni, 1995) and on individual productivity inside firms (Hossain and List, 2012; Hong et al., 2015). We contribute to the existing literature on framing inside organizations by providing evidence that presenting tasks appealing to the internal motivation of contributing to the goal of the organization can be an effective incentive to raise levels of effort and participation in tasks that have the potential to benefit more than the performer.

Our work is also related to the literature on the provision of incentives for innovation and creativity at work. In particular, our field experiment is similar to that of Gibbs et al. (2014) focusing on the effect of pecuniary incentives on the production of employee ideas. Gibbs et al. (2014) conduct a field experiment in which they compare teams' performance inside a consulting firm across two treatments. In one treatment, teams obtain personal rewards whenever one of their own ideas gets adopted by clients. In the other, teams are not incentivized by pecuniary prizes. Our work is different in several respects. First, in Gibbs et al. (2014), agents are grouped into teams and the team is the main unit of analysis while our study examines the effect at the individual level. Second, contribution to the public good is not a relevant factor in Gibbs et al. (2014), whereas our study involves a contest that seeks public good contributions.

Finally, our study is also related to the large literature on differences in preferences between men and women. We find that women respond to a solicitation appealing to the mission of the organization by increasing participation, whereas men respond by reducing it. This difference can be explained in many different ways. Several studies in psychology and behavioral economics have shown that women behave as if they are less competitively inclined than men (Gneezy and Rustichini, 2004; Niederle and Vesterlund, 2007; Gneezy et al., 2009), and more risk averse compared to men (Croson and Gneezy, 2009). Thus, women may be expected to bear higher costs from taking part in an ideation challenge for free compared to men, and this is consistent with our observation of a generally higher propensity of women to participate in the treatment with pecuniary prizes, given expected payoffs can be larger than fixed costs to participate.

3 Predictions

We conceptualize an internal call for proposals to improve the organization as a voluntary contribution mechanism for an organizational public good. Successful proposals would improve the workplace for employees or improve quality and efficiency of patient care. Submitting proposals during the competition requires costly effort by submitters, such as the time necessary to identify a problem, form a proposal, write up a concise description, and potentially lead to the likelihood of implementation of proposals.

For illustrative purposes, consider a linear model of the utility of a typical employee, who contributes x and benefits from total contributions of Y:

$$u(R,\ Y) = \gamma Y + \delta x + \frac{x}{Y}R - cx. \tag{1}$$

The benefits of contributing derive from three sources. First, there is an altruistic benefit from the improved workplace, γY. The altruistic benefits are the crux of public goods. Only the existence of an improved workplace is desired and the source of contributions is irrelevant. Thus, everyone would prefer to free-ride on others' efforts. Second, participants may also directly benefit from personally contributing to improving the workplace, δx. The benefit from personally contributing is often referred to as warm-glow (e.g., Andreoni, 1995), as in the participants experience a warm-glow of feelings when contributing. Finally, participants have some chance of winning the contest and can expect to derive benefits from the prizes, $\frac{x}{Y}R$, where, for simplicity, all efforts have an equal chance of being selected as the winner, as in Morgan (2000). The personal reward R can be thought of as a pecuniary prize, but it could also be an increase in prestige or recognition, or any combination of the above. Contributors incur some effort cost from developing and submitting a proposal cx.

If there are n employees the public goods dilemma arises when $\gamma + \delta < c < n\gamma + \delta$. Then no individual would contribute without a reward as costs exceed individual benefits, but everyone would be better off if everyone contributes.

Suppose contributing a proposal is a discrete choice by employees. An employee can either contribute a single proposal $x = 1$ and receive utility of

$$u_1 = \gamma \hat{Y} + \delta + \sum_{k=1}^{n} \Pr(Y = k)\frac{R}{k} - c, \tag{2}$$

where $\hat{Y}$ denotes the expected level of contributions and $\Pr(Y = k)$ is the probability of having k total contributions. Or they can contribute nothing $x = 0$ and receive utility of

$$u_0 = \gamma(\hat{Y} - 1). \tag{3}$$

If there are n employees, then a symmetric mixed-strategy equilibrium is for each employee to contribute a proposal with probability $p > 0$. After using the binomial probability for $\Pr(Y = k)$, the payoff-equating condition to find a mixed-strategy equilibrium is:

$$\frac{1-(1-p)^n}{np} = (c-\gamma-\delta)/R. \tag{4}$$

This equation admits one single solution p^* which cannot be expressed explicitly. Using a first order Taylor expansion around p, the equilibrium probability can be approximated as follows:

$$p^* \approx \frac{2(R-c+\gamma+\delta)}{(n-1)R}. \tag{5}$$

The analysis of the above model is used to derive three predictions as follows.

1. The probability of contributing a proposal to improving the organization is zero when the prize for winning is sufficiently small relative to the individual cost of effort minus the preference for the public good (i.e., $R < c - \gamma + \delta$).

2. The probability of contributing a proposal to improve the organization increases with the value of the prize for winning.

3. The probability of contributing a proposal to improve the organization increases with the extent of individual preference for the public good ($\gamma + \delta$).

This framework can be extended to the case of individuals with heterogeneous costs. In the appendix, we explicitly consider the case of two types of individuals with different marginal costs of effort that form two groups of equal size. The symmetric mixed-strategy equilibrium is then characterized by the vector of probabilities of contributing with a proposal (p_1^*, p_2^*). Here, the analysis of the payoff-equating conditions for the mixed-strategy equilibrium shows that the higher the marginal cost of effort minus preference for contributing, the lower the equilibrium probability of individuals (i.e., $p_1^* > p_2^*$ when $c_1 < c_2$, and vice versa). This leads to our last prediction.

4. If individuals have heterogeneous costs, then the probability of contributing a proposal to improve the organization is higher for agents with lower costs (positive sorting).

4 The context and the experimental design

4.1 Background

The study was conducted at the Heart Center from July 2014 to September 2014. The Heart Center is a leading academic medical center specializing in clinical cardiac care and

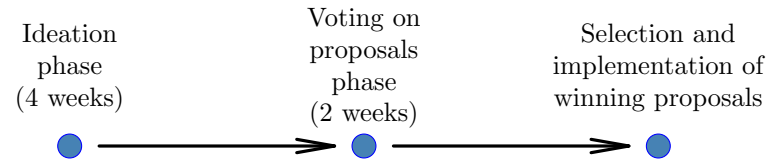


Figure 1: *The Ether Dome Challenge.*

research in the United States. Founded more than a hundred years ago, the Heart Center serves thousands of patients every year, and employs more than 1,200 people (nurses, physicians, researchers, technicians, and administrative staff) scattered across several buildings on the Massachusetts General Hospital's main campus in downtown Boston. The study was initiated in cooperation with the Heart Center's launch of the Healthcare Transformation Lab (HTL).[7] The launch of the HTL was accompanied by the announcement of "The Ether Dome Challenge,"[8] which sought to engage all staff to participate. The announcement read: "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." The communication around the challenge highlighted 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 challenge involved three main phases that are shown in Figure 1. First, in the four–week long ideation phase, employees at all levels of the organization were encouraged to identify one or more organizational problems and submit proposals addressing them. To lower the costs of entry and encourage participation, each proposal was limited to approximately 300 words, although participants could submit more than one proposal. Next, in the two–week long peer evaluation phase, all employees were asked to rate the merit and potential of submitted proposals, through a five-point rating scale. After this phase, submitters of proposals that were highly rated by peers and were judged as particularly promising by the HTL staff were invited to submit a full proposal detailing plans for implementation. Fol-

[7] An initiative aimed at developing innovative health care process improvements designed to enhance the health care safety and delivery of the hospital (`http://www.healthcaretransformation.org`).

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

lowing evaluation by MGH senior leadership, top proposals were selected to receive support and funding for implementation (up to $20,000).

Data for our study were obtained through the design of the first two phases. Participation in the challenge was voluntary and proposals could only be submitted online via a website. All Heart Center employees could submit proposals, covering any issue within the organization (as described above). The challenge was announced to all staff members in a series of emails, the content of which constituted our main experimental intervention. We designed two main constraints in the submission phase to ensure that treatment effects could be isolated, identified, and matched to participants. First, during the submission phase, employees were asked to work individually; team submissions were not allowed. Limiting submissions to individual participation allowed us to match each submitter's characteristics to the randomly assigned treatment. Second, the identity of the submitter and the content of all submissions were not disclosed during the ideation phase in order to ensure that the employee's decision to contribute would not be affected by decisions made by others.

The peer evaluation phase invited Heart Center staff to vote on the proposals submitted. The start of the (voluntary) voting phase was announced by email to every member of the Heart Center. Once logged in, the evaluator was shown a list of randomly selected proposals (the identities of the submitters were not revealed) to read and rate the overall quality of each proposal on a five point scale between 1 (low quality) and 5 (high quality).[9] Ratings were confidential and the website did not provide any feedback, or any other kind of additional information that might influence individual judgement, until the voting phase was over (e.g., there was no leaderboard system to see how a submission was performing in the voting). Evaluators were free to decide how many (and which) proposals to vote on, and because they were presented in a random order, every proposal had on average the same exposure to voting. Each evaluator received a limited edition T-shirt as a compensation for the effort in voting.

4.2 Treatments

We worked closely with Heart Center executives in the communication strategy for the challenge and designed four solicitation treatments. The control (control) treatment echoed the generic marketing of the challenge, encouraging submissions and commitment of hospital management in the form of project funding.

[9]Proposals were presented in lists of 10 each. Each proposal is described by three main elements: a title, a main description of the problem to solve, and the main proposal. Voting is then introduced by the following text: "Rate this idea" followed by the rating scale: low; 2; 3; 4; high.

The pecuniary prize (prize) treatment was the explicit offering of individual pecuniary prizes for the best proposals. The patient care (pcare) and the workplace (wplace) treatments appealed to internal motivation of the employees by emphasizing the patient care mission of the organization and emphasizing improvement in the work processes of the organization, respectively.

In the Control treatment, the call for proposals was presented via email as follows:

> "Dear Heart Center team member,
>
> Submit your ideas to win project funding up to $20,000 to turn your ideas into actions.
>
> [...] 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!"

This solicitation email conveyed the goals of the challenge and the evaluation and implementation processes following the ideation phase. No personal rewards were mentioned.

In the prize treatment, an explicit pecuniary incentive was introduced by replacing the first paragraph of the above solicitation with the following:

> "Submit your ideas to win an Apple iPad mini."

The rest of the solicitation email was exactly the same as in the control treatment.

The aim of the remaining two treatments was to appeal to internal motivations to participate. To this end, the solicitation message was "framed" in the following way. In the patient care treatment, the first paragraph was replaced by:

> "Submit your ideas to improve patient care at the Heart Center."

And in the workplace treatment, the first paragraph was replaced by:

> "Submit your ideas to improve the workplace at the Heart Center."

The remaining text was also exactly the same as in the control. Unlike the Prize treatment, these interventions mentioned no personal reward to be won. The only difference with respect

to the control treatment was in the emphasis on the goals of the challenge and the mission of the organization: improving healthcare of patients and improving the working place of the Heart Center. Each intervention was emphasized by being placed at the very beginning of the email, where readers are believed to decide whether to keep reading the full message or stop.

To encourage more participation, each employee received the same exact solicitation email three times (at the launch, eight days from the launch and two days before the end of the ideation phase of the challenge). To raise awareness and spur more participation, information booths and the distribution of fliers and posters were organized starting from the second week of the challenge, which encouraged everyone to check their emails to take part in the event.

All solicitation emails included a direct link to the website where participants could register for the challenge and submit proposals. The website required employees to enter their work email to login, which allowed us to automatically match some important features of the website with the assigned treatment.[10] The headings of the website and the space just below a "submit your ideas" button had the same exact text as the first paragraph in the solicitation email (i.e., the treatment). We also changed the main graphics of the website in accordance with the treatment (see Figure A.5 in the online appendix). This was to ensure that participants would read the paragraph of the assigned treatment at least once before a submission.

Overall, we anticipated solicitation emails to be an effective way to manipulate individual beliefs about the challenge (or any contest of similar nature) occurring for the first time in the organization. Moreover, the scattered units within the Heart Center oblige email communication to be the main source of information. To strengthen the effect of announcement and signal legitimacy of the contest, we relied only on official channels for communication.

5 Data

5.1 The Heart Center

The sample for this experiment consists of the entire population working at the Heart Center as of the end of 2014. Overall there are $1,237$ individuals. For each individual we collected

[10] The system was not restricted to Heart Center staff members only, but allowed anyone with an institutional account to login and to submit proposals. This included staff members affiliated with other centers within MGH. If for example someone not in our sample decided to login, then the website would have displayed texts and graphics at random from one of the four treatments.

data on gender, type of profession, and office location when available. The large majority (72 percent) of workers are women, however there is gender separation by profession with nurses almost exclusively women (92 pecent); physicians are, by contrast, mostly men (76 percent). We code the type of profession into three main categories: physicians (18 percent), nurses (52 percent) and a residual category called others (30 percent) including researchers, technicians, and administrative staff each almost in an equal number. Only half of the employees (53 percent) have a fixed office location within the campus.[11] For a limited group of 378 employees (30 percent) we were able to complement these data with additional self-reported information (e.g., age, tenure) that was obtained from an online survey polling employees about the quality of the workplace environment. The survey was run about two months before the start of the challenge. Based on these non-representative data we postulate that the median employee is between 36 and 45 years old and has been working at the Heart Center for 10 years.

Table 1 presents summary statistics showing that individual characteristics were statistically balanced across the four treatment groups.

5.2 Challenge outcomes

At the end of the four-week submission period, we collected 118 proposals submitted by 60 employees.[12] Proposals were later voted on by 181 employees in the evaluation phase, yielding 12,219 voter-proposal pairs. Most people judged many proposals, as the median was 68 rated proposals (ranging between 1 and 113). Participation in the challenge was representative of all different professions: 4.4 and 15.5 percent of physicians submitted and voted on a proposal, respectively; 4.9 and 14.2 of nurses; and 5 and 14.1 of others.

The spectrum of submitted proposals ranges from technological innovations, such as innovative mobile applications or wearable tools, to procedural changes to improve upon the existing staff workflows. Content of proposals appears to conform to the stated goals of the challenge, aligning with improving the work processes within the organization or providing high-quality patient care. According to a classification made by the HTL staff, the primary areas of focus were: patient support (19 percent of all proposals), information and access (18), staff workflows (17), care coordination (15), workplace efficiency and costs (12), quality and safety (9), surgical tools and support to research (2), and a residual category (7).

[11] Much of the clinical staff might be mobile and may not have fixed office locations as they may be on duty in multiple wards.

[12] We exclude an additional 20 proposals submitted by 11 people not in our sample because made by MGH employees that were not employed by the Heart Center at the time of the design of the experiment.

Table 1: Summary Statistics, By Treatment Groups

	Control n=308	PCARE n=310	WPLACE n=307	PRIZE n=312
Nursing	51.0	52.3	55.7	51.0
Physician	19.2	18.1	18.2	17.6
Other job	29.9	29.7	26.1	31.4
Sum	100	100	100	100
	$\chi^2 = 2.7$, df = 6, p-value = 0.8436			
Female	68.8	69.4	75.6	74.0
Male	31.2	30.6	24.4	26.0
Sum	100	100	100	100
	$\chi^2 = 5.2$, df = 3, p-value = 0.1586			
No office	49.7	45.5	47.2	44.2
Office	50.3	54.5	52.8	55.8
Sum	100	100	100	100
	$\chi^2 = 2.1$, df = 3, p-value = 0.5565			

Note: This table reports the number and percentage of employees in our sample cross tabulated by the assigned treatment across gender, profession and office. The variable "office" stands for an employee with no fixed office location within the Heart Center. The table also reports Pearson's χ^2 statistics, degrees of freedom and p-values as from independence tests between treatments and each category.

We observe some variation in the quality of proposals. The word count of submissions ranged from 2 (i.e., "More beepers!") to 1,098 (about 4 pages of double-spaced text with 9 different proposals).[13] With regards to voting, the average rank of a proposal ranged between 2.2 and 3.9 (on a scale from 1 to 5) and the voters per proposal ranged between 92 and 125.

Overall, these efforts resulted in 2 proposals being selected by leaders for full funding, plus 4 proposals implemented at no significant cost for the organization.

6 Results

This section presents a comparative analysis of participation and the quality of proposals across treatments.

6.1 Submissions

We start by looking at individual participation in terms of submissions. Figure 2 shows the trend of employees responding with a submission over time. In the control treatment, about half of the responses took place in the first two weeks of the ideation phase and almost nothing afterward. This pattern differs strikingly from the other treatment groups where participation increased throughout the ideation phase. This trend resulted in a noticeable difference in the final number of submissions, which suggests that both the solicitation for a pecuniary prize and the framing manipulation encouraged extra participation compared to the control.

Table 2 presents the number of employees solicited and the final number of employees responding for each of the four treatment groups: 'control', 'prize', 'pcare' and 'wplace'. 4.9 percent of employees responded, which is large relative to the short duration of the initiative and the novelty of the approach (especially in the health care setting).[14] The participation rate π_j for each single treatment j ranged between $\pi_{control} = 2.2$ and $\pi_{prize} = 7.4$ percent. A Pearson's chi-squared test for the independence between participation rates and treatments gives a p-value of 0.03, which means that this is not just random variation in the data.

To interpret and test the significance of the observed differences in participation rates we

[13] Note, each submission had no limit on the number of proposals. Each proposal, however, had a limit of about 300 words.

[14] List and Lucking-Reiley (2002) find very similar participation rates (between 3 and 8 perecent) in a field experiment on charitable giving for a University capital campaign.

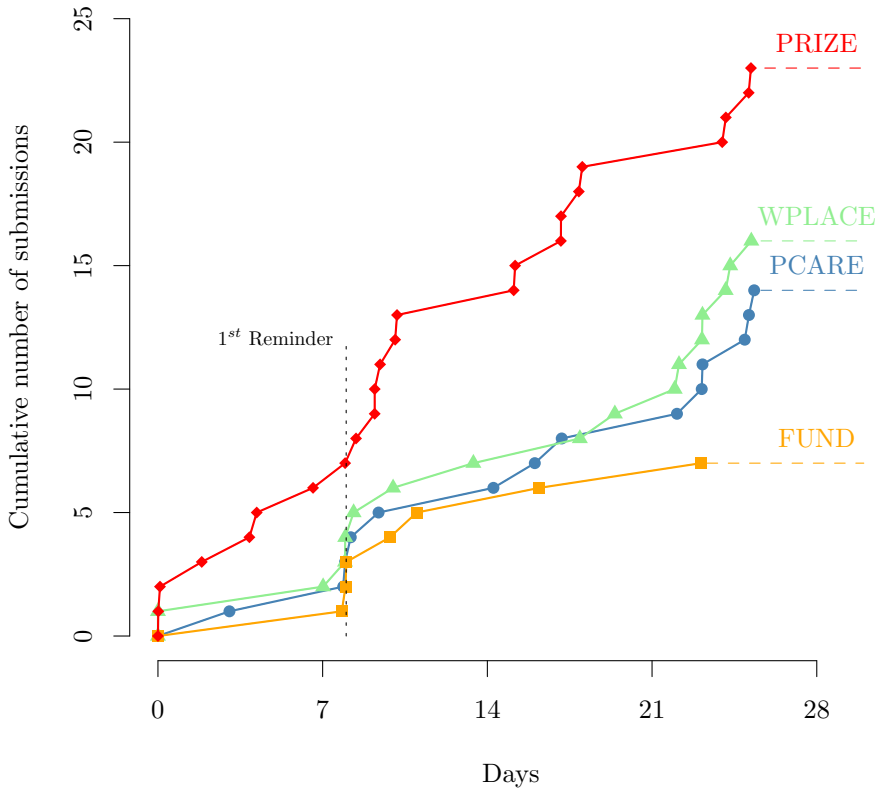


Figure 2: Cumulative submissions over time by treatment. A solicitation email was sent to everyone on the first date of the ideation phase (day 0). The same email was sent again as a reminder on day 8 and on day 26 of the ideation phase. On day 8 posters and leaflets were distributed to advertise the event encouraging everyone to check for the invitation to the challenge in their emails.

Table 2: Participation, by Treatment Groups

	Employees solicited	Employees responding	Percentage responding
Control	308	7	2.3
PCare	310	14	4.5
WPlace	307	16	5.2
Prize	312	23	7.4
Total	1,237	60	4.9

use a linear probability model:[15]

$$SUBMIT_{ij} = \alpha + \tau_j + PROFESSION_i + GENDER_i + OFFICE_i \tag{6}$$

where the dependent variable $SUBMIT_{ij}$ indicates that, at the end of the submission period, an employee i assigned to a treatment j has either made a submission ($SUBMIT_{ij} = 1$) or not ($SUBMIT_{ij} = 0$). The parameter τ_j denotes a shift in the probability of a submission that is associated with the treatment. Similarly, the other coefficients represent discrete shifts in the probability associated with individual characteristics such as profession, gender, and a dummy for office location (as a proxy for status within the organization).

Main effects. Ordinary least squares estimates of the model (6) and heteroskedasticity robust standard errors are shown in Table 3. Given the random assignment, the estimated $\hat{\tau}_j$ can be interpreted as the change in the probability that an individual will contribute with a submission caused by the treatment j.

The treatment with pecuniary prize is found to have a highly significant ($p < 0.01$) and positive causal effect on participation. In fact, as we increase the pecuniary prize from zero to a positive value (i.e., the value of an iPad), the probability that an individual will contribute with a submission increases by 5.1 percent (the 95 percent asymptotic confidence interval lies between 1.6 and 8.9). Given the baseline probability of 2.2 percent (i.e., the constant in column 1 of Table 3), the pecuniary prize treatment has essentially tripled the probability of a submission.

The two treatments with framing are found to have a positive effect on participation as well. But these estimates have higher uncertainty (i.e., type-I error). The causal effect of

[15] Nearly identical results are obtained with a logistic regression model.

Table 3: Probability of an employee making a submission

	$SUBMIT_{ij}$							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
PCARE	2.24	2.24	2.24	2.24	2.13	2.01	2.10	2.01
	(1.46)	(1.46)	(1.46)	(1.46)	(1.46)	(1.45)	(1.46)	(1.45)
WPLACE	2.94*	2.94*	2.90*	2.91*	2.87*	2.70*	2.76*	2.67*
	(1.53)	(1.53)	(1.54)	(1.55)	(1.53)	(1.52)	(1.54)	(1.55)
PRIZE	5.10***	5.09***	5.07***	5.07***	4.95***	4.84***	4.85***	4.82***
	(1.71)	(1.71)	(1.71)	(1.71)	(1.71)	(1.70)	(1.70)	(1.70)
Physician		−0.47		−0.25		−0.60		−0.42
		(1.79)		(1.84)		(1.79)		(1.84)
Nursing		−0.03		−0.16		2.69		2.57
		(1.43)		(1.55)		(1.73)		(1.83)
Male			−0.54	−0.52			−1.54	−0.42
			(1.33)	(1.65)			(1.45)	(1.64)
Office					2.79**	4.56***	3.18**	4.56***
					(1.20)	(1.60)	(1.31)	(1.60)
Constant	2.27***	2.38*	2.44***	2.57*	0.87	−1.28	1.15	−1.12
	(0.85)	(1.27)	(0.91)	(1.48)	(0.91)	(1.68)	(0.92)	(1.83)
N	1,237	1,237	1,237	1,237	1,237	1,237	1,237	1,237
Log Likelihood	150.63	150.67	150.71	150.72	153.25	155.12	153.85	155.15
Akaike Inf. Crit.	−293.26	−289.35	−291.42	−287.44	−296.49	−296.23	−295.69	−294.29

Note: This table reports coefficients from a linear probability model estimating the probability of an employee making a submission. Heteroskedasticity robust standard errors are in parenthesis. Coefficients and standard errors are multiplied by 100 so that readers can interpret them as percentage point change in the probability.

*** *Significant at the 1 percent level.*

** *Significant at the 5 percent level.*

* *Significant at the 10 percent level.*

the treatment for improving the workplace is significant at the 10 percent level ($p = 0.06$), whereas that of the treatment variable for improving patient care ($p = 0.13$) is slightly above the conventional level of 10 percent.[16] Both treatments generated large effects on participation relative to the baseline: an increase of 2.9 and 2.2 percent respectively.

With regards to profession, the estimated effects for nurses and physicians are both negative relative to the residual category of other workers but not significant. Similarly for gender, the difference between men and women is small and insignificant. This result is surprising because differences in skills and in income between hospital employees can be sharp.[17] One possible interpretation is that individual costs of effort for contributing with a submission may not change much between groups in the hospital. This interpretation makes sense because everyone may have an idea on how to improve the organization, regardless of their profession. Moreover, proposals were not required to be long or very technical in order to keep individual costs of participation small for everyone.

The office location variable is found to have a significant and positive effect on participation ($p < 0.01$). Within a profession, one may interpret this participation effect as due to differences associated with an individual's position or status within the organization.[18] Under this view, this evidence suggests that there is positive sorting based upon underlying motivations associated with the hierarchical position inside of the organization (e.g., reputation, prestige), rather than differences in the cost of submitting.

Overall, this evidence confirms a significant and positive effect of our intervention on the average effort. The treatment with pecuniary prize significantly increased participation. There is also evidence, although less strong, suggesting a positive effect of the treatment for improving the workplace. The sign of the effect of the treatment for improving patient care is positive but not statistically significant. Somewhat surprisingly, we don't find evidence of sorting based on individual costs of effort, but there is some evidence of a positive sorting based upon experience or the status inside the organization.

[16] However, a one-tailed test indicates that the probability in the PCARE treatment is *greater than* the probability in the control group at 10 percent significant level ($p = 0.09$).

[17] According to the United States Bureau of Labor Statistics, the median wage of a physician is about 40 percent higher than the that of a registered nurse; and that of a registered nurse is about 20 percent higher than that of an administrative worker.

[18] Having an office location or not is highly correlated with the type of profession. For instance, given the nature of their job, nurses are less likely to have a fixed office location than physicians or administrative workers. However, within each profession, having an office location is usually correlated with the hierarchical position inside of the organization.

Interactions. Some further insights on the effect of our intervention are provided by allowing treatment effects to change systematically across groups. There is a large literature in economics suggesting that observable differences in behavior, such as consumption or investment decisions, can be explained by difference in preferences between men and women. Economic experiments in the laboratory have been able to show differences in risk aversion, competitive inclination and altruism (see Croson and Gneezy, 2009, for a review). This led us to hypothesize systematic differences in the way men and women responded to our intervention.

We examined a more general version of the model (6) where the treatment effect $\tau_{j,g}$ can vary by gender g. Then we estimated this model in two ways: in one case we run a series of OLS regressions on the subsample of men and women separately, and in the other we use OLS to estimate a full model with interactions between treatment and gender dummies on the entire sample. In all cases, given the random assignment, the estimated coefficient $\hat{\tau}_{j,g}$ is the causal effect of the treatment j on the average effort conditional on gender.

Estimates are shown in Table 4. By comparing the coefficients obtained on the subsample of women and men (Columns 1 to 4), the effect of the treatment with pecuniary prize is found to be larger for women than for men. The estimated change in the probability that a female employee will contribute with a submission is 6 percent ($p < 0.05$), whereas it is only 3 percent and insignificant for a male employee. Using the full model with interactions (Column 5 or 6), one can formally test the significance of this difference. The estimate is negative and sizable but not significant. Therefore, while we cannot formally reject the null hypothesis of no difference, this evidence is suggestive of women responding more than men to a solicitation with pecuniary prizes.

Significant gender differences are found with respect to the treatment variable for improving patient care. Controlling for professional status (Columns 2 and 4), the estimated change in the probability that a female employee will contribute with a submission is 6 percent ($p < 0.05$), whereas it is negative -2 percent for men and not significant. In the full model, the estimated difference between men and women is negative -2 percent and significant ($p < 0.05$).

Taken together with the main effects, there is evidence of a positive, statistically significant and sizable effect of the treatment with pecuniary prize on participation compared to a control group with no pecuniary prize. The pecuniary incentive may be more effective for women compared to men (although the difference is not significant). This indicates that men may potentially bear lower costs in being involved in a competition with no pecuniary prizes. There is also evidence supporting a causal effect of framing. The treatment for improving

Table 4: Probability of an employee making a submission. Differences between the genders.

	$SUBMIT_{ij}$					
	(1)	(2)	(3)	(4)	(5)	(6)
PCARE	4.16**	3.82**	−2.07	−2.17	4.16**	3.87**
	(1.88)	(1.88)	(2.08)	(2.09)	(1.88)	(1.88)
WPLACE	2.42	2.15	4.87	4.66	2.42	2.17
	(1.64)	(1.63)	(3.65)	(3.61)	(1.64)	(1.63)
PRIZE	5.91***	5.54***	3.05	2.80	5.91***	5.63***
	(2.00)	(2.02)	(3.25)	(3.35)	(2.00)	(2.01)
PCARExMale					−6.23**	−6.03**
					(2.81)	(2.84)
WPLACExMale					2.45	2.54
					(4.00)	(4.02)
PRIZExMale					−2.86	−2.75
					(3.82)	(3.89)
Male					1.24	1.31
					(2.03)	(2.09)
Constant	1.89**	−1.62	3.13*	0.54	1.89**	−1.61
	(0.94)	(2.09)	(1.79)	(3.04)	(0.94)	(1.81)
Controls	No	Yes	No	Yes	No	Yes
N	890	890	347	347	1,237	1,237
Log Likelihood	91.76	96.20	62.29	62.86	153.55	157.91
Akaike Inf. Crit.	−175.51	−178.39	−116.58	−111.72	−291.11	−293.82

Note: This table reports coefficients from a linear probability model estimating the probability of an employee making a submission using the subsample of women (columns 1-2) the subsample of men (columns 3-4) and the entire sample (columns 5-6). Heteroskedasticity robust standard errors are in parenthesis. Coefficients and standard errors are multiplied by 100 so that readers can interpret them as percentage point change in the probability. Even number columns (2, 4, 6) also include controls for profession and office location.

*** *Significant at the 1 percent level.*

** *Significant at the 5 percent level.*

* *Significant at the 10 percent level.*

the working place has on average a positive higher participation compared to the control group. On the contrary, the solicitation for improving the patient care has a countervailing effect for men, whereas, it has a positive effect on women. One interpretation is that there are differences in the way men and women perceive the mission of the organization. Another possibility is the presence of gender stereotypes in the workplace that may affect individual decisions to contribute with a submission, as we will discuss later.

6.2 Robustness

In this section, we discuss the results obtained for robustness checks regarding the main analysis.

As pointed out by Young (2015), testing for multiple treatment effects at once may cause the *actual* level of uncertainty (i.e., type-I error) to be higher than the *nominal* level: a problem also known as False Discovery Rate (FDR) or Multiple Comparison Problem (MCP). Our results, however, are robust to more conservative p-values, such as the ones computed using the Benjamini and Hochberg (1995) correction. Under the Benjamini and Hochberg (1995) correction, the effect $\hat{\tau}_{prize}$ is found significant at 5 percent level ($p = 0.01$) and $\hat{\tau}_{wplace}$ at 10 percent level ($p = 0.08$).

Another concern is the relatively small number of responses (i.e., submissions) compared to the sample size. It is widely acknowledged that asymptotic distributions of test statistics often give poor approximations in small samples (Horowitz, 2001, p. 3161). In binary variables models, this problem occurs as well when the overall number of responses is small compared to the sample size, as for rare events (King and Zeng, 2001). One direct consequence is that estimated asymptotic confidence intervals can be artificially smaller than they really are. There are several ways to address this problem, such as bootstrap methods (Wu et al., 1986), *exact* inference (Agresti, 1992), and correction methods (King and Zeng, 2001). While most of these solutions have been applied within the context of logistic regression, the bootstrap procedure is very general and also easy to implement within the linear probability models. So, we computed non-parametric bootstrap confidence intervals at the 10 percent significance level. Results are reported in Table 5 showing that our main results hold even under bootstrap confidence intervals.

Another way to check the robustness of our results is to examine the subsample of staff members (378 individuals) who volunteered in an unrelated online survey that took place a few weeks before our intervention. This sample may not be representative of the full population but had higher participation rates (10 percent), which allows us to control for systematic bias associated with the small number of submissions. Focusing the analysis on

Table 5: Bootstrap confidence intervals of the probability of making a submission.

	1	2
PCARE	−0.63, 5.28	0.25, 7.50
WPLACE	−0.03, 5.96	−1.10, 5.36
PRIZE	1.86, 8.40	1.75, 9.83
PCARExMale		−11.60, −0.69
WPLACExMale		−4.99, 10.98
PRIZExMale		−10.43, 4.72
Male		−2.46, 5.63
Constant	0.66, 4.17	−5.23, 1.91
Controls	No	Yes
N	1,237	1,237

Note: This table reports non-parametric bootstrap (percentile method with 2000 replica) confidence intervals at the 5 percent significance level from a linear probability model estimating the probability of an employee making a submission. Confidence intervals are multiplied by 100 so that readers can interpret them as percentage point change in the probability. Column 2 also includes controls for profession and office location.

this subsample can be intriguing in its own right. It can be presumed that survey respondents are more active staff members compared to the general population, especially in organization-related activities. It also allows us to use additional variables in the regression taken from the survey, such as age. So we re-run the same regressions that we described before (see Table A.5 in the online appendix). The *main* results are qualitatively unchanged: the treatment with pecuniary prize increased participation by 10.8 percent ($p < 0.05$) and the treatment for improving the working place by 8.5 percent ($p < 0.05$).

6.3 Quality of proposals

We now turn to differences in the underlying quality of submissions that were generated by the treatments. Here, the main measure of quality is based upon the ratings collected for each proposal in the peer evaluation phase of the challenge. Basic descriptive statistics are shown in Table 6.[19]

Table 6: Ratings of the submitted proposals

	Proposals evaluated	Proposal average rating	Proposal median rating	Number of ratings	Employees voting
Control	11	3.1	3	1,146	165
PCare	33	3.2	3	3,571	180
WPlace	30	3.3	3	3,283	178
Prize	39	3.2	3	4,219	178
Total	113	3.2	3	12,219	181

For the average rating, a liner regression with treatment dummies was performed. Testing for a significant linear regression relationship between a proposal's average rating and the treatments gives no significant results (F-test gives a p-value of 0.65). This result is consistent with a zero effect on average. Figure 3 shows further that the empirical densities overlap quite well across treatments. Thus, not surprisingly, a non-parametric Kolmogorov-Smirnov testing procedure finds no significant distributional differences; the test rejects significant differences between the full sample and the treatment with pecuniary prizes ($p = 0.27$),

[19]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 ($p = 0.7$).

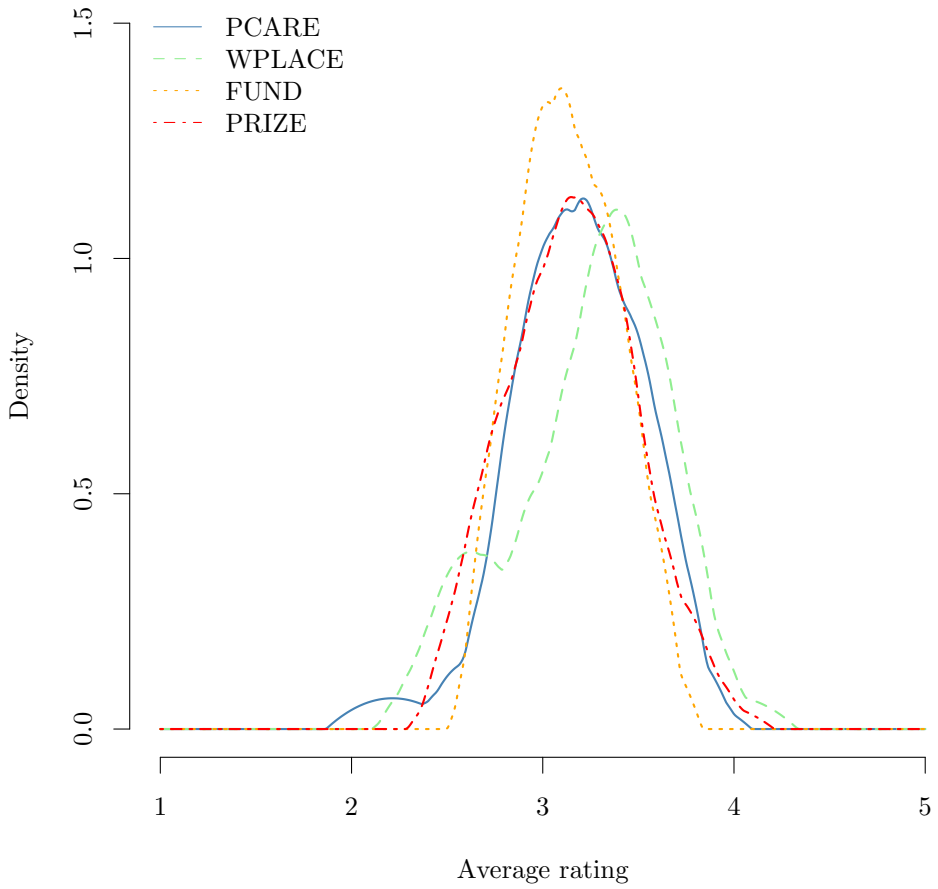


Figure 3: Empirical probability density function of a proposal's average rating conditional on the treatment of the proponent.

the improving patient care treatment ($p = 0.79$), and the improving workplace treatment ($p = 0.16$).

As a robustness check, differences in the probability of a proposal getting a particular rating v from 1 to 5 were also examined (see Table 7). This comparison has the main advantage of not making any strong assumption about how to aggregate ratings.[20] As can be seen in the table, there is an overlap in empirical densities across treatments, suggesting an underlying comparable quality. Also in this case, we do not find a significant association between ratings and treatments at conventional significance levels.

[20]By contrast, a comparison based upon differences in the average rating 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, which is not generally true for ordinal variables.

Table 7: Contingency table of proposals ratings, by treatment group

	Control n=1146	PCare n=3571	WPlace n=3283	Prize n=4219
1	15.0	14.4	13.3	14.3
2	16.3	15.5	14.4	16.1
3	28.0	28.3	27.4	28.5
4	21.3	21.0	22.7	21.9
5	19.4	20.8	22.2	19.2
Sum	100	100	100	100
$\chi^2 = 18.4$, df = 12, p-value = 0.1031				

Note: The table also reports Pearson's χ^2 statistics, degrees of freedom and p-value of an independence test between the treatments and ratings.

Objective quantitative measures of effort in preparing the proposals were further examined: the length of the submission, and the count of proposals per submission. Testing for a significant linear regression relationship between the length of submissions, as measured by the count of words, and the treatment dummies returned an overall insignificant result (F-test of significance gives a p-value of 0.43). Similarly, the relationship between the count of proposals per submission and the treatments is found not significant (F-test gives a p-value of 0.81).[21]

To sum up, based on the analysis of three different measures, we do not find evidence of any sizable difference in the quality of submissions across treatments: submissions are of comparable length; include on average an equal number of proposals; and a proposal's average rating is approximately the same. One may want to conclude that the provided incentives had no effects on the average quality. This conclusion, however, has some important caveats. As will be discussed in the next section, ratings were binned into 5 discrete categories, which may limit our ability to detect more granular differences in quality across treatments. Moreover, there were word limits for proposals that may put an upper bound on the extent of quality differentiation. Hence, we do not view these results as evidence against quality effects in general.

[21] Regression results are shown in Table A.5 in the online appendix.

7 Discussion

The empirical results presented support our theoretical predictions (Section 3). As the pecuniary value of the reward for winning goes from zero to a positive value (i.e., from the control group to the treatment with pecuniary prize), participation increases by 5 percent. Participation also increases by 2-3 percent when the solicitation sent to everyone in the Heart Center emphasizes the public good nature of the task (i.e., from the control group to the treatments with framing). This evidence supports the hypothesis that an agent's decision to contribute with a submission is proportional to the value of the expected reward (Predictions 1 and 2), and increases with the level of social preferences (Prediction 3).

Beyond these qualitative results, we now use the theoretical model developed in Section 3 to estimate the magnitude of the underlying incentives faced by individuals in each treatment. This means computing an estimate of the personal reward from winning (i.e., R) and the level of social preferences for contributing to the public good (i.e., γ) that are consistent with the experimental results in each treatment. (Here, we do not distinguish between pure or "warm-glow" preferences for the public good.)

To perform the estimates, we set the mixed-strategy equilibrium probability of contributing (i.e., defined by equation 4) equal to the empirical probability (i.e., $\hat{\gamma}_j$ for a treatment j). This gives a relationship between R and γ that must be satisfied. In addition, we calibrate the cost of submitting a proposal c using the median hourly wage of a nurse in Massachusetts (i.e., \$40)[22] and the number of competitors n using only 30 percent of the entire sample.[23] Given these are constant across treatments, such calibration is not crucial for our results.

In the control group the probability of submitting is about 2 percent. Thus, the procedure just described leads to:

$$\hat{R} = \frac{0.02 \times 0.3 \times 1237}{1 - 0.98^{0.3 \times 1237}}(40 - \hat{\gamma}) = 297 - 7.4 \times \hat{\gamma}. \tag{7}$$

Equation (7) is the relationship between an implicit reward from winning and the prefer-

[22] Here, we consider the median income of a Nurse Practitioner according to the Bureau of Labor Statistics. Nurse practitioners are the most represented jobs in our sample. Yet, one may argue that this value may underestimate the cost of contributing for other professions, such as physicians, because these workers have higher opportunity costs of time. This criticism, however, is not supported by data, as we find no empirical evidence of sorting based upon profession.

[23] This choice is our best guess of the number of active staff members at the Heart Center. This is based on the number of employees who took a survey before the experiment (378 people). Assuming greater participation would lead to artificially increasing the estimates of underlying incentives. In fact, staff members may have rational expectations about the actual number of potential participants, which may be less than the entire population.

ences for the public good that is consistent with the control group data. The implicit reward can be thought of as the value of non-pecuniary rewards from winning (e.g., reputation, peer recognition, prestige). This relationship is inversely proportional because rewards and social preferences must substitute each other to keep participation fixed.

In the treatment with pecuniary prize, the probability of submitting was 5 percent higher than in the control group. Theoretically, this gap is due entirely to the difference in rewards from winning. So, letting R_p denote the implicit reward in the treatment with pecuniary prize, we get the following relationship:

$$\hat{R}_p = \frac{0.07 \times 0.3 \times 1237}{1 - 0.93^{0.3 \times 1237}}(40 - \hat{\gamma}) = 1039.1 - 25.9 \times \hat{\gamma}. \tag{8}$$

Given the pecuniary prize was worth \$400,[24] one can use equation (7) together with equation (8) to identify R and γ separately. As shown graphically in Figure 4, this identification step can be thought of as shifting the curve for equation (7) by the value of the pecuniary prize and using the point where it intersects with the curve for equation (8) to pin down an estimate of γ (i.e., $\hat{\gamma}$).

As a result, the magnitude of the social preferences towards improving the organization is $\hat{\gamma} = \$18$. Thus, the impact of social preferences on the probability of contributing with a submission is equivalent to a 45 percent reduction in the cost of contributing. This result is complementary to the evidence on social preferences provided by Della Vigna et al. (2016). Della Vigna et al. (2016) show that freelance workers have social preferences towards their employers and these preferences may affect the level of effort at work. We show that the same may happen for workers inside of an organization. Here, a larger effect is found which may be attributed to the combined consequences of the organization pursuing a social goal (e.g., providing healthcare) and the individuals being members of an organization, rather than freelance workers.

The magnitude of the implicit reward from winning in the control group is found positive and large $\hat{R} = \$160$. This result suggests sizable non-pecuniary incentives (e.g., reputation, peer recognition) in contributing to public goods inside of organizations, other than just social preferences or pecuniary prizes. The empirical work of Kosfeld and Neckermann (2011) and Blanes i Vidal and Nossol (2011) also find positive effects associated with the role of congratulatory rewards and rankings in tournaments with no prizes. Here, the main difference

[24]Recall that the pecuniary prize was an iPad mini. The value of a regular iPad mini as a reward can be estimated by looking at the cost of purchasing the item online. The price paid by the Heart Center was \$239 at the end of 2014 (no shipping cost). Other very popular models (i.e., those with cellular data and large storage) could cost as high as \$350. Agents, however, were not aware of the specific model used for the competition and of the price paid. So, the value of \$400 is intended to be very conservative.

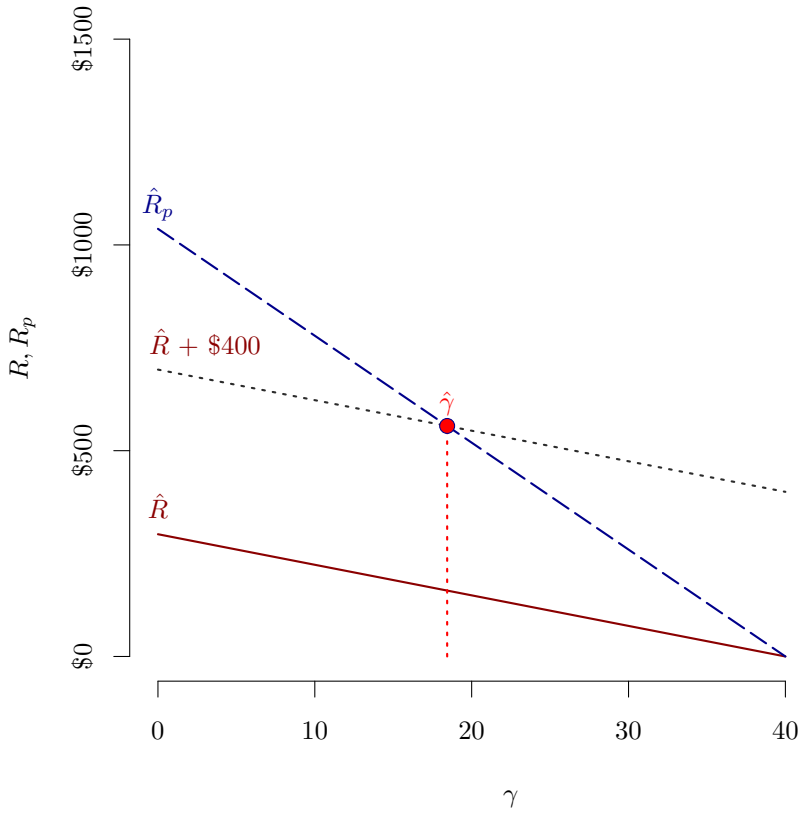


Figure 4: Solid and dashed curves show equations (7) and (8) respectively.

is that reputation or social status may be acquired through a public good contribution to improving the organization rather than on the basis of differences in individual productivity or skills.

In the treatment with framing, one may hypothesize that the increase in participation of 2-3 percent is due to an increase in social preferences, as if individuals experienced a higher level of personal satisfaction from contributing per se (i.e., γ). On the other hand, it is also possible that framing raised the level of the implicit reward from winning, as individuals expected a higher reputation increase due to contributing to a socially valuable task (i.e., R). Here, experimental data are consistent with both possibilities, and we cannot distinguish between these two effects. All we can infer without making further assumptions is that framing had a significant positive impact on one or both of these two variables. This result supports the existing literature on framing manipulations in public good games in the laboratory (e.g., Andreoni, 1995), and extends its validity to a field setting. It is also in line with the studies finding an effect of framing on effort inside of firms (e.g., Hossain and List, 2012; Hong et al., 2015).

A few remarks are in order here. To obtain confidence intervals around these estimates one may consider several sources of uncertainty. First, there is the uncertainty for the estimation of the probability of submitting in our sample. This type of uncertainty, however, is easy to quantify and evaluate (e.g., standard errors can be computed directly from the data). Another source of uncertainty is due to the calibration of the marginal cost or the number of competitors. Also in this case, it is not difficult to evaluate the variation of our estimates (e.g., one may use historical data about internal compensations for extra activities). Finally, another important source of uncertainty is with respect to the main behavioral assumptions of the model, which we discuss below.

One possible concern is that people often make incorrect judgements about incentives. In the laboratory, several studies have indeed found over-expenditure of effort in contests (see Dechenaux et al., 2014). This finding is often associated with systematic mistakes or judgemental biases, such as overconfidence. If agents in our sample were "overconfident," our estimate of the preferences for the public good would be biased upward, as we were falsely considering "altruism" any excess of effort due to overconfidence. Yet, this type of bias alone is unlikely to explain large differences across treatments, as the ones observed in our setting.

Another possibility is that of a negative interaction between pecuniary and non-pecuniary incentives. Theoretically, pecuniary incentives for public goods or pro-social activities may crowd out non-pecuniary incentives to perform these activities (Deci and Ryan, 1985; Ben-

abou and Tirole, 2003). Empirically, non-monotonicites in price incentives have been indeed documented (Gneezy and Rustichini, 2000; Lacetera et al., 2014). Here, a crowding-out effect would bias R_p, because incentives do not simply add up. Our procedure would then underestimate the social preferences (i.e., $\hat{\gamma}$). Given the already large estimate of social preferences, this appears unlikely.

We now turn to discuss the observed difference in response between genders. We argue that this could be due to systematic differences in preferences between men and women. Several studies have provided evidence supporting the hypothesis of gender-based differences in preferences (Croson and Gneezy, 2009). Theoretically, there are three main possible explanations that one can consider. First, there is evidence on gender differences in risk aversion broadly suggesting that women are more risk averse than men (e.g., Borghans et al., 2009). Second, women seem to shy away from competitive situations, either because of feedback aversion or different attitudes toward competition (Niederle and Vesterlund, 2007). Third, there may be potential differences amongst the genders in altruism (Andreoni and Vesterlund, 2001).

In our setting, women are found to be more responsive to a competition with a pecuniary incentive than men. While one needs to be cautions in attributing field evidence of gender differences in outcomes to specific psychological traits, this finding is not consistent with the view that women are more risk averse than men, as the competition for prizes generates greater variation in outcomes. One may argue that women are more attracted by a competition with pecuniary incentives because the expected payoff is overall higher than the same competition with no pecuniary incentives. If so, the differential response can be explained as a differential in the fixed cost of competing. Or, to put it differently, women seem to bear higher costs to compete, which is consistent with the view that women are less competitively inclined than men. Differences in altruism may explain why women are more responsive than men to the treatment for improving patient care. However the existing evidence on difference in the level of altruism suggests that this may not be the case. Another possible explanation is that agents respond differently to situations affected by gender-stereotypes. For example, Coffman (2014) has shown evidence that women are less likely to contribute ideas in situations that are strongly identified with male stereotypes.

Another empirical result that merits further consideration is that the treatments had zero effects on the quality of the proposals submitted. Gibbs et al. (2014) reports results from a field experiment in an ideation contest and observes effects in both the number and quality of submissions. Here we only observed effects on the number of submissions. We conjecture the difference is due to the difference in rating systems. In our study, proposals

were rated from 1 to 5 stars; binning proposals into 5 values. Five stars is a readily reached quality level in our setting; 20 percent of ratings were 5-stars. This fact and the "hump-shaped" rating distribution suggests the bins are near the central portion of the quality distribution. Hence our quality measure is more sensitive to changes in moderate quality proposals. In Gibbs et al. (2014), the quality metric was the proportion of proposals pitched to clients of the company under study. In effect, binning every proposal as 0 or 1 and focusing on the upper-tail of quality where proposals get binned as 1. For example, compare the ratings of proposals A, B, C, and D with hypothetical true qualities of 3, 4, 5, and 10 "stars" respectively. Assume proposals with ratings above 5-stars are accepted. Under the Gibbs et al. (2014) system, proposals A and B cannot be distinguished, but C and D can be distinguished. Under our rating system, the opposite is true; A and B can be distinguished, but C and D cannot be distinguished. Hence we view our results as complementary to Gibbs et al. (2014) and not evidence against quality effects in general. Taken together, the two studies suggest that quality effects are not constant across the distribution of quality.

Overall, in recent years incentive schemes based on an employee's performance in a contest have been increasingly used by management to increase the productivity and creativity of employees. Existing studies on contest theory, both theoretical and empirical, often neglect the role of externalities exerted by agents contributing ideas to a contest, which may result in problems of under-provision. The present paper tries to fill this gap by looking at the intricate relationship between incentives and the motivation to contribute ideas in order to improve the organization. Within this context, the field experimental results presented confirm that monetary prizes have an important role in spurring participation, without affecting the quality of the ideas. But we also find significant effects of framing manipulations that we interpret as consistent with impure altruism. Overall, these results support the view that workers have multiple underlying motivations to contribute to public goods inside the organization, consisting of a combination of pecuniary and altruistic incentives associated with the mission of the organization.

References

Agresti, Alan, "A Survey-of Exact Inference for Contingency Tables," *Statistical Science*, 1992, *7* (1), 131–153.

Akerlof, George A and Rachel E Kranton, "Identity and the Economics of Organizations," *Journal of Economic Perspectives*, 2005, *19* (1), 9–32.

Andreoni, James, "Warm-Glow versus Cold-Prickle: The Effects of Positive and Negative Framing on Cooperation in Experiments," *The Quarterly Journal of Economics*, 1995, *110* (1), 1–21.

_ **and Lise Vesterlund**, "Which is the Fair Sex? Gender Differences in Altruism," *The Quarterly Journal of Economics*, 2001, *116* (1), 293–312.

Auriol, Emmanuelle and Régis Renault, "Status and incentives," *The RAND Journal of Economics*, 2008, *39* (1), 305–326.

Bandiera, Oriana, Iwan Barankay, and Imran Rasul, "Social Preferences and the Response to Incentives: Evidence from Personnel Data," *The Quarterly Journal of Economics*, 2005, *120* (3), 917–962.

_ **,** _ **, and** _ , "Team incentives: Evidence from a firm level experiment," *Journal of the European Economic Association*, 2013, *11* (5), 1079–1114.

Benabou, Roland and Jean Tirole, "Intrinsic and extrinsic motivation," *The Review of Economic Studies*, 2003, *70* (3), 489–520.

Benjamini, Yoav and Yosef Hochberg, "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing," *Journal of the Royal Statistical Society. Series B (Methodological)*, 1995, *57* (1), 289–300.

Besley, Timothy and Maitreesh Ghatak, "Competition and Incentives with Motivated Agents," *The American Economic Review*, 2005, *95* (3), 616–636.

_ **and** _ , "Status incentives," *American Economic Review*, 2008, *98* (2), 206–211.

Borghans, Lex, James J Heckman, Bart HH Golsteyn, and Huub Meijers, "Gender differences in risk aversion and ambiguity aversion," *Journal of the European Economic Association*, 2009, *7* (2-3), 649–658.

Che, Yeon-Koo and Seung-Weon Yoo, "Optimal Incentives for Teams," *American Economic Review*, 2001, *91* (3), 525–541.

Coffman, Katherine Baldiga, "Evidence on Self-Stereotyping and the Contribution of Ideas," *The Quarterly Journal of Economics*, 2014, *129* (4), 1625–1660.

Croson, Rachel and Uri Gneezy, "Gender Differences in Preferences," *Journal of Economic Literature*, 2009, *47* (2), 448–474.

Dale, Donald J, "Charitable lottery structure and fund raising: Theory and evidence," *Experimental Economics*, 2004, *7* (3), 217–234.

Dechenaux, Emmanuel, Dan Kovenock, and Roman M Sheremeta, "A survey of experimental research on contests, all-pay auctions and tournaments," *Experimental Economics*, 2014, pp. 1–61.

Deci, Edward L and Richard M Ryan, *Intrinsic motivation and self-determination in human behavior*, Springer Science & Business Media, 1985.

Delfgaauw, Josse, "Dedicated Doctors: Public and Private Provision of Health Care with Altruistic Physicians," *New England Journal of Medicine*, 2005, *353* (15).

_ **and Robert Dur**, "Incentives and Workers' Motivation in the Public Sector*," *The Economic Journal*, 2008, *118* (525), 171–191.

Erev, Ido, Gary Bornstein, and Rachely Galili, "Constructive intergroup competition as a solution to the free rider problem: A field experiment," *Journal of Experimental Social Psychology*, 1993, *29* (6), 463–478.

Gibbons, Robert, "Incentives in Organizations," *Journal of Economic Perspectives*, 1998, *12* (4), 115–132.

Gibbs, Michael, Susanne Neckermann, and Christoph Siemroth, "A field experiment in motivating employee ideas," *ZEW-Centre for European Economic Research Discussion Paper*, 2014, (14-096).

Gneezy, Uri and Aldo Rustichini, "Pay Enough or Don't Pay at All*," *The Quarterly journal of economics*, 2000, *115* (3), 791–810.

_ **and** _ , "Gender and Competition at a Young Age," *American Economic Review*, 2004, *94* (2), 377–381.

_ **, Kenneth L Leonard, and John A List**, "Gender Differences in Competition: Evidence From a Matrilineal and a Patriarchal Society," *Econometrica*, 2009, *77* (5), 1637–1664.

Green, Jerry R and Nancy L Stokey, "A Comparison of Tournaments and Contracts," *The Journal of Political Economy*, 1983, *91* (3), 349–364.

Holmstrom, Bengt, "Moral Hazard in Teams," *Bell Journal of Economics*, 1982, *13* (2), 324–340.

Holmström, Bengt and Paul Milgrom, "Regulating Trade Among Agents," *Journal of Institutional and Theoretical Economics (JITE)*, 1990, *146* (85), 85–105.

Hong, Fuhai, Tanjim Hossain, and John A List, "Framing manipulations in contests: a natural field experiment," *Journal of Economic Behavior & Organization*, 2015.

Horowitz, Joel L., "The Bootstrap," in James J. Heckman and Edward Leamer, eds., *Handbook of Econometrics*, Vol. 5 of *Handbook of Econometrics*, Elsevier, 2001, chapter 52, pp. 3159 – 3228.

Hossain, Tanjim and John A List, "The behavioralist visits the factory: Increasing productivity using simple framing manipulations," *Management Science*, 2012, *58* (12), 2151–2167.

i Vidal, Jordi Blanes and Mareike Nossol, "Tournaments without prizes: Evidence from personnel records," *Management Science*, 2011, *57* (10), 1721–1736.

Itoh, Hideshi et al., "Incentives to Help in Multi-agent Situations," *Econometrica*, 1991, *59* (3), 611–36.

King, Gary and Langche Zeng, "Logistic regression in rare events data," *Political analysis*, 2001, *9* (2), 137–163.

Kosfeld, Michael and Susanne Neckermann, "Getting More Work for Nothing? Symbolic Awards and Worker Performance," *American Economic Journal: Microeconomics*, 2011, *3* (3), 86–99.

Lacetera, Nicola, Mario Macis, and Robert Slonim, "Rewarding volunteers: a field experiment," *Management Science*, 2014, *60* (5), 1107–1129.

Landry, Craig E, Andreas Lange, John A List, Michael K Price, and Nicholas G Rupp, "Toward an Understanding of the Economics of Charity: Evidence from a Field Experiment*," *The Quarterly journal of economics*, 2006, *121* (2), 747–782.

Lange, Andreas, John A List, and Michael K Price, "Using lotteries to finance public goods: Theory and experimental evidence*," *International Economic Review*, 2007, *48* (3), 901–927.

Lazear, Edward P and Sherwin Rosen, "Rank-Order Tournaments as Optimum Labor Contracts," *The Journal of Political Economy*, 1981, *89* (5), 841–864.

Ledyard, John, "Public goods: A survey of experimental research," Technical Report, David K. Levine 1997.

List, John A, "The Market for Charitable Giving," *The Journal of Economic Perspectives*, 2011, *25* (2), 157.

_ **and David Lucking-Reiley**, "The effects of seed money and refunds on charitable giving: Experimental evidence from a university capital campaign," *Journal of Political Economy*, 2002, *110* (1), 215–233.

_ **and Imran Rasul**, "Field experiments in labor economics," *Handbook of labor economics*, 2011, *4*, 103–228.

Marwell, Gerald and Ruth E Ames, "Economists free ride, does anyone else?: Experiments on the provision of public goods, IV," *Journal of public economics*, 1981, *15* (3), 295–310.

Mary, O'Keeffe, W Kip Viscusi, and Richard J Zeckhauser, "Economic Contests: Comparative Reward Schemes," *Journal of Labor Economics*, 1984, *2* (1), 27–56.

Moldovanu, Benny and Aner Sela, "The optimal allocation of prizes in contests," *American Economic Review*, 2001, pp. 542–558.

_ **,** _ **, and Xianwen Shi**, "Contests for status," *Journal of Political Economy*, 2007, *115* (2), 338–363.

Morgan, John, "Financing public goods by means of lotteries," *The Review of Economic Studies*, 2000, *67* (4), 761–784.

_ **and Martin Sefton**, "Funding public goods with lotteries: experimental evidence," *The Review of Economic Studies*, 2000, *67* (4), 785–810.

Neumann, Anna and Aimee LaPointe Terosky, "To give and to receive: Recently tenured professors' experiences of service in major research universities," *The Journal of Higher Education*, 2007, *78* (3), 282–310.

Niederle, Muriel and Lise Vesterlund, "Do Women Shy Away From Competition? Do Men Compete Too Much?," *The Quarterly Journal of Economics*, 2007, *122* (3), 1067–1101.

Prendergast, Canice, "The Provision of Incentives in Firms," *Journal of Economic Literature*, 1999, *37* (1), 7–63.

_ , "The Motivation and Bias of Bureaucrats," *American Economic Review*, 2007, *97* (1), 180–196.

Rabin, Matthew, "Psychology and Economics," *Journal of Economic Literature*, 1998, *36* (1), 11–46.

Ryan, Richard M and Edward L Deci, "Intrinsic and extrinsic motivations: Classic definitions and new directions," *Contemporary educational psychology*, 2000, *25* (1), 54–67.

Vesterlund, Lise, "Voluntary giving to public goods: moving beyond the linear VCM," in John H Kagel and Alvin E Roth, eds., *The Handbook of Experimental Economics*, Vol. 2, in preparation, Princeton University Press, 2012.

Vigna, Stefano Della, John A List, Ulrike Malmendier, and Gautam Rao, "Estimating Social Preferences and Gift Exchange at Work," Technical Report, National Bureau of Economic Research 2016.

Wu, CFJ et al., "Jackknife, Bootstrap and Other Resampling Methods in Regression Analysis," *The Annals of Statistics*, 1986, *14* (4), 1261–1295.

Young, Alwyn, "Channeling Fisher: Randomization Tests and the Statistical Insignificance of Seemingly Significant Experimental Results," Technical Report, Working paper, October 2015.

A Online appendix

In section A.1 of this online appendix, we present a formal proof of the result of sorting based upon costs in the extended model with heterogenous costs discussed in the Predictions section of the main paper. In section A.2, the graphics used in the challenge website are shown, as well as tables with additional results are presented.

A.1 Extended model with heterogenous costs

Proof. Consider the case of two types of individuals $j = 1, 2$ forming two groups of equal size $n_1 = n_2 = n$. Individuals can decide to contribute with a single proposal or not. When an agent of type j decides to contribute, the expected utility is as follows.

$$u_1^j = \gamma \hat{Y} + \delta_j + \sum_{k_j=1}^{n} \sum_{k_l=0}^{n} \Pr(Y = k_j + k_l) \frac{R}{k_j + k_l} - c_j.$$

The utility of not contributing is as follows.

$$u_0^j = \gamma(\hat{Y} - 1).$$

Equating these two conditions for all individuals gives the following mixed-strategy equilibrium condition:

$$\sum_{k_j=1}^{n} \sum_{k_l=0}^{n} \Pr(Y = k_j + k_l) \frac{R}{k_j + k_l} = c_j - \delta_j + \gamma$$

for all $j = 1, 2$. To examine differences in equilibrium probabilities p_1^* and p_2^*, we use the ratio between the above equilibrium condition for individuals of type $j = 1$ and the same expression for agents of type $j = 2$. This gives:

$$\frac{\sum_{k_1=1}^{n} \sum_{k_2=0}^{n} \Pr(Y = k_1 + k_2) \frac{R}{k_1+k_2}}{\sum_{k_1=0}^{n} \sum_{k_2=1}^{n} \Pr(Y = k_1 + k_2) \frac{R}{k_1+k_2}} = \frac{c_1 - \delta_1 + \gamma}{c_2 - \delta_2 + \gamma}.$$

The left hand side can be rearranged as follows.

$$\frac{\Pr(k_2 = 0) \sum_{k_1=1}^{n} \Pr(Y = k_1) \frac{R}{k_1} + \sigma R}{\Pr(k_1 = 0) \sum_{k_2=1}^{n} \Pr(Y = k_2) \frac{R}{k_2} + \sigma R}$$

where $\sigma = \sum_{k_1=1}^{n} \sum_{k_2=1}^{n} \Pr(Y = k_1 + k_2) \frac{1}{k_1+k_2}$. Using $1 - p_2 = \Pr(k_2 = 0)$ and $1 - p_1 = \Pr(k_1 = 0)$ together with the density of the binomial distribution, we obtain the following simpler expression.

$$\frac{(1 - p_2) \frac{(1-(1-p_1)^n)}{np_1} R + \sigma R}{(1 - p_1) \frac{(1-(1-p_2)^n)}{np_2} R + \sigma R}.$$

If $c_1 - \delta_1 > c_2 - \delta_2$, then the above expression in equilibrium needs to be larger than one. This inequality can be expressed as follows:

$$\frac{p_2(1 - p_2)}{(1 - (1 - p_2)^n)} > \frac{p_1(1 - p_1)}{(1 - (1 - p_1)^n)}.$$

Hence, the inequality is satisfied only if p_2 is greater than p_1. *Q.E.D.*

A.2 Graphics and tables with additional results

Table A.5: Probability of an employee making a submission – Subsample with survey

	$SUBMIT_{ij}$					
	(1)	(2)	(3)	(4)	(5)	(6)
PCARE	1.95	1.90	1.93	1.19	1.94	1.16
	(3.65)	(3.68)	(3.68)	(3.67)	(3.69)	(3.70)
WPLACE	8.52**	8.54**	8.54**	7.71*	8.61**	7.85*
	(4.16)	(4.20)	(4.17)	(4.23)	(4.16)	(4.33)
PRIZE	10.78**	10.78**	10.96**	10.06**	11.08**	10.43**
	(4.70)	(4.71)	(4.71)	(4.67)	(4.72)	(4.79)
Physician		−1.30				−4.68
		(4.89)				(5.21)
Nursing		−1.11				3.35
		(3.61)				(4.83)
Male			1.39			3.55
			(3.99)			(4.91)
Office				6.97**		7.25*
				(3.06)		(4.07)
age18-25					−6.78	−7.91
					(12.11)	(12.54)
age26-35					−6.20	−6.56
					(11.50)	(11.50)
age36-45					5.02	3.55
					(12.02)	(12.01)
age46-65					1.58	0.68
					(11.72)	(11.76)
age65+					7.46	5.91
					(14.33)	(14.70)
Constant	5.49**	6.26*	5.16**	1.36	5.56	0.41
	(2.42)	(3.38)	(2.48)	(2.43)	(11.16)	(12.60)
N	378	378	378	378	378	378
Log Likelihood	−92.01	−91.95	−91.95	−89.91	−87.39	−85.37
Akaike Inf. Crit.	192.03	195.90	193.90	189.81	192.79	196.75

Note: This table reports coefficients from a linear probability model for the probability of an employee contributing with a submission using the subsample of 378 employees who have taken an online survey. Heteroskedasticity robust standard errors are in parenthesis. Coefficients and standard errors are multiplied by 100 so that readers can interpret them as percentage point change in the probability.

***Significant at the 1 percent level.*

**Significant at the 5 percent level.*

Significant at the 10 percent level.

Table A.5: Proposal ratings, submissions' word count and number of proposals

	Submission length	Proposals	Proposal ave. rating
	(1)	(2)	(3)
PCARE	0.05	−25.02	1.00
	(0.11)	(40.56)	(1.27)
WPLACE	0.11	26.13	0.37
	(0.11)	(40.36)	(1.25)
PRIZE	0.02	−8.97	0.17
	(0.11)	(38.50)	(1.19)
Constant	3.13***	127.37***	1.57
	(0.10)	(33.70)	(1.04)
N	113	60	60
R^2	0.01	0.05	0.02
Adjusted R^2	−0.01	−0.004	−0.04
Residual Std. Error	0.03 (df = 109)	77.82 (df = 56)	2.75 (df = 56)
F Statistic	0.55 (df = 3; 109)	0.93 (df = 3; 56)	0.33 (df = 3; 56)

Note: This table reports coefficients from a linear regression model for submission length as measured by the word count of the proposals made (column 1), the number of proposals per submission (column 2) and the proposal's average rating (column 3).

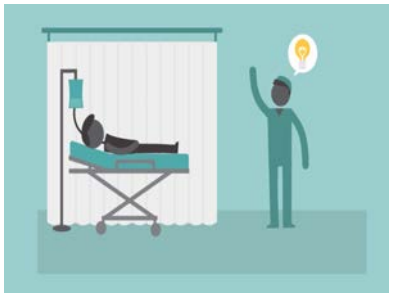
*** *Significant at the 1 percent level.*

** *Significant at the 5 percent level.*

* *Significant at the 10 percent level.*

Control

Prize



Patient care

Workplace

Figure A.5: Graphics displayed on the website of the challenge at the login matching the randomly assigned treatment.