Using Behavioral Economics to Curb Workplace Misbehaviors:
Evidence from a Natural Field Experiment

By Jeffrey A. Flory, Andreas Leibbrandt, and John A. List*

Abstract: Workplace misbehaviors are often governed by explicit monitoring and strict punishment. Such enforcement activities can serve to lessen worker productivity and harm worker morale. We take a different approach to curbing worker misbehavior— bonuses. Examining more than 6500 donor phone calls across more than 80 workers, we use a natural field experiment to investigate how different wage contracts influence workers' propensity to break workplace rules in harmful ways. Our findings show that even though standard relative performance pay contracts, relative to a fixed wage scheme, increase productivity, they have a dark side: they cause considerable cheating and sabotage of co-workers. Yet, even in such environments, by including an unexpected bonus, the employer can substantially curb worker misbehavior. In this manner, our findings reveal how employers can effectively leverage bonuses to eliminate undesired behaviors induced by performance pay contracts.

Key Words: Cheating, expressive law, natural field experiment, sabotage, social norms, tournament.

JEL Codes: C93, J22, J33.

*Flory: Department of Economics, University of Chicago, Chicago, IL 60637 (Visiting); and Robert Day School of Economics and Finance, Claremont McKenna College, Claremont, CA 91711 (email: jflory@cmc.edu). Leibbrandt: Department of Economics, Monash University, Clayton, Vic 3800, Australia (email: andreas.leibbrandt@monash.edu.au). List: Department of Economics, University of Chicago, Chicago, IL 60637 (email: jlist@uchicago.edu). An earlier version of this paper was previously distributed as "The Effects of Wage Contracts on Workplace Misbehaviors: Evidence from a Call Center Natural Field Experiment."

"We're taking technology that was built for counter-terrorism and using it…. If you want to be proactive, you have to get people before they act."

> Tim Estes, CEO of Digital Reasoning Systems Inc, describing his monitoring system to curb employee misbehavior. His clients include Goldman Sachs Group Inc. and Credit Suisse Group AG, among others.

"[Coworkers] sabotage other people's efforts. One of the most valuable things I learned was to give the appearance of being courteous while withholding just enough information from colleagues to ensure they didn't get ahead of me on the rankings."[1]

> A worker from Microsoft, describing the effects of the company's employee ranking system.

1. Introduction

Misbehavior on the job is an issue of central and growing concern for many organizations, who annually spend billions of dollars to attenuate such behaviors. JP Morgan, for instance, recently spent $730 million and hired 2,500 compliance workers to improve operations through surveillance of employees.[2] Likewise, the National Security Agency is developing a computerized system to continuously monitor employees across the government.[3] These industry leaders are not alone, as roughly two-thirds of U.S. employers monitor Internet connections, and almost half use video surveillance.[4] The large investments in employee monitoring despite ethical concerns, legal issues, and negative impacts on job satisfaction and productivity (Enzle and Anderson, 1993; Frey, 1993; Barkema, 1995) suggest many organizations understand the import of misbehavior in the workplace.

Laws and workplace regulations are in place to protect organizations from employee misbehavior. Employee sabotage, for example, can be punished as a crime and like other misbehaviors, such as cheating, can result in loss of employment.[5] Such laws and regulations might also shape norms if they express societal values (Sunstein, 1996;

---

[1] http://www.vanityfair.com/news/business/2012/08/microsoft-lost-mojo-steve-ballmer
[2] http://www.bloomberg.com/news/articles/2015-04-08/jpmorgan-algorithm-knows-you-re-a-rogue-employee-before-you-do
[3] http://www.pbs.org/newshour/rundown/us-intelligence-officials-monitor-federal-employees-security-clearances/
[4] According to the latest Electronic Monitoring & Surveillance Survey (2007) which interviewed 304 U.S. companies. Visit www.amanet.org/research for further information.
[5] While sabotage is already illegal in the US, 30 states have additional laws at various stages in their legislatures that would make a broader array of harmful behaviors by employees j (including sabotage) toward their coworkers against the law (see http://www.healthyworkplacebill.org/).

Cooter, 1998; Posner, 1998; Benabou and Tirole, 2011), and thus dampen misbehaviors. Beyond such laws, workplace regulations, and norms, worker misbehavior might critically rely on the nature of incentives and wage contracts. In particular, economic theory predicts that relative performance pay regimes exacerbate sabotage and cheating (Lazear and Rosen, 1981; Dye, 1984; Lazear, 1989; Konrad, 2000).

Indeed, there is some anecdotal evidence for such misbehavior among workers on the job that supports these predictions. For example, after heavy criticism of the adversarial environment it created among coworkers, Microsoft abolished its employee evaluation system based on relative performance. Amazon.com has also been heavily criticized for a work culture in which employees regularly believe that they are being secretly sabotaged by unidentifiable co-workers.[6] Furthermore, findings from the laboratory show that among strangers relative performance pay and ranking feedback can cause unproductive and unethical behavior (Carpenter et al., 2010; Harbring and Irlenbusch, 2011; Charness et al, 2014). In addition, Courty and Marschke (2004) show that explicit incentives trigger gaming responses in a large government organization, in which employees are less likely to know each other.

To provide initial empirical insights into the prevalence, drivers, and potential moderators of employee misbehavior in the workplace, we conduct a natural field experiment. Similar to many modern work environments, employees in our setting interact with other workers across an information channel that is less intimate than face-to-face environments.[7] This characteristic, which is common to many information technologies in wide and growing use, may facilitate different kinds of employee misbehaviors, as the level of depersonalization makes detection harder and reduces the fear of personal consequences – including reputation costs or fear of retaliation. Conceptually, this depersonalization of interactions has the advantage that it leads to clearer predictions and allows for a cleaner test of behavioral theories than if we were to study a more traditional workplace with repeated face-to-face interactions among a limited set of workers. From a practical

---

[6] "At Amazon, workers are encouraged to tear apart one another's ideas …," write Kantor et al. (2015). "The internal phone directory instructs colleagues on how *to send secret feedback to one another's bosses*. Employees say it is frequently used to sabotage others." http://www.nytimes.com/2015/08/16/technology/inside-amazon-wrestling-big-ideas-in-a-bruising-workplace.html

[7] Consider, for example, interactions in the rapidly-growing gig economy, among telecommuting employees, among office workers via software, email communications, etc.

perspective, it also has the advantage of providing insights on a quickly expanding type of work setting in the modern era.

Our field experiment uses an actual employer of call center employees that recruits laborers who work independently, to investigate the effect of the wage regime on their propensity to sabotage (increase the probability of outperforming a co-worker by interfering with coworker production output) and cheat (increase the probability of outperforming a co-worker through illicit actions, but without interfering with others' production output).[8] In doing so, we observe worker behavior under exogenously varying wage contracts in a natural setting (see, e.g., List and Rasul, 2011 for a survey of labor market field experiments more generally). Thus, we are able to generate a first glimpse into how social norms and beliefs, which may be shaped by laws and workplace regulations, constrain worker misbehavior, and how this is related to incentive structures.

We operationalize the study of sabotage and employee misbehaviors in this setting by observing workers who are collecting donations via telephone for a charitable organization. A key consideration is that the solicitors contact two types of individuals: those who have not recently donated (called the 'cold' list in the industry) and those who have donated to the charity in the past three years (the 'warm' list). As the shift changes, the solicitors pass along their lists of potential donors to the next shift of workers, and therefore control the quality of solicitees that they pass along to a co-worker. By marking outcomes of soliciting calls untruthfully, the workers can sabotage or cheat in a manner that is quite difficult to detect.

Yet, we are able to detect whether callers have engaged in such behaviors, as we gather dialing information from the phones after the callers have finished their shift. This information includes what phone numbers were dialed by the caller, whether the call connected, and how long the connection lasted. Thus, we are able to determine whether workers skipped cold numbers and lied about the call outcome, or marked warm numbers to remove them from their co-workers' call list without having valid reason.

We report several insights. First, performance pay induces worker cheating and workplace sabotage. Employees who can receive a bonus for relative performance are

---

[8] We define workplace sabotage as illicit behavior intended to increase one's probability of out-competing potential competitors by interfering with their production output. In contrast, we define cheating as illicit behavior intended to increase one's probability of winning, without interfering with competitors' production output.

4

more likely to skip and transfer cold list numbers to their competitors. However, the incidence of sabotage is not unduly high. Despite the clear pecuniary incentive to sabotage co-workers, employees are no more likely to sabotage than to cheat although cheating does not negatively affect the competitors' productivity. In addition, we find that almost all employees refrain from invalidly removing warm numbers from the list; i.e. they do not exhibit severe forms of sabotage. Another way that competition-based performance incentives might inadvertently create or exacerbate employee misbehavior is by causing workers to strategically misrepresent their actual work output. Despite making it easy for the employees in our experiment to do this, there is no evidence they commit this form of misbehavior.

These findings provide evidence that workers do not exhaust all opportunities to sabotage and cheat to improve their wages and suggest that their hesitance is driven by social norms and moral considerations. If this is true, then it could be beneficial for employers to use mechanisms that activate social norms. One candidate for such a mechanism is gift-exchange (Akerlof, 1982; Akerlof and Yellen, 1988; 1990). The gift-exchange hypothesis is that higher wages lead to more effort, or, more generally, that employees reciprocate good treatment from their employers. In the standard gift exchange game, for example, when given a monetary gift and an option to transfer some money back to the giver, many individuals return positive amounts (Fehr et al, 1993). In a workplace setting, while employees do not have a mechanism to reciprocate monetarily, a handful of studies have shown that employees respond by investing greater effort in work tasks, raising their productivity, after a bonus (Gneezy and List, 2006; Kube et al 2012; Cohn et al, 2014a; Cohn et al, 2014b).

Another dimension along which workers may choose to reciprocate towards the employer is curtailing deviant behaviors that violate job responsibilities and hurt the firm. In our field experiment, we find evidence for such gift-exchange. In a further treatment, we test the impact of higher base wages in performance pay contracts on employee misbehavior. A second main finding is that the incidence of cheating and sabotage in relative pay environments can be curbed with unexpected higher fixed-wages. These findings show that misbehavior may be related in important ways to the level of compensation. Thus, it takes the literature on gift-exchange in a new direction by taking

into account undesirable forms of effort and their relationships to wages. In addition, we find evidence for standard gift-exchange as employees generate significantly more donations if their performance pay is combined with an unexpected boost to their base wage.

Finally, we find interesting heterogeneities amongst workers. There is evidence that deviant behavior by men, in particular, may be more responsive to relative-performance based incentives, as compared to women. This is intriguing, as it suggests a potential dark side to the findings in the literature on gender differences in effort-response to competition incentives: while output of men may increase under competition, they may also be more likely to break the rules and misbehave in order to win the competition (Schwieren and Weichselbaumer, 2010; Dato and Nieken, 2014). Such an effect can neutralize the positive impacts of relative performance contracts amongst men.

We view this study as complementary to an active line of research on cheating and sabotage (Konrad, 2009; Chowdhury and Gurtler; 2014), which is based on laboratory experiments (Preston and Szymanski, 2003; Harbring et al, 2007; Carpenter et al, 2010; Schwieren and Weichselbaumer, 2010; Vandegift and Yavas, 2010; Harbring and Irlenbusch, 2011; Gilpatric, 2011), and sports competitions (Balafoutas et al, 2012; Deutscher et al, 2013). These studies provide key empirical evidence for the scope of misbehavior within well-defined game environments when individuals are under scrutiny. In contrast, we investigate employee misbehavior that is difficult to detect in a workplace where it can adversely affect other employees and the employer, and is moderated by laws and workplace regulations.

Our findings also contribute to a rich vein of studies that cuts across many different literatures, and examine different types of hidden unethical behavior by individuals in a variety of different economic settings. Zitzewits (2012) provides an excellent review of papers in this vein, grouping them loosely under the umbrella of "forensic economics". To our knowledge, none of these studies on misbehavior examine workplace cheating and coworker sabotage through a field experiment that directly observes employee responses to randomly assigned variation in pay structure. The closest field experiments we are aware of test the impacts on unethical activity from randomly increasing perceived probability of detection – for example in cheating on taxes (Slemrod et al., 2001), skimming from

government projects (Olken, 2007), and secretly skipping out on work (Duflo et al., 2008). In the context of employee misbehavior, a related study by Nagin et al. (2002) finds that lowering perceived monitoring rates at a call center increases donation pledge misreporting. Our study stands apart from these experiments in two important ways. First, rather than looking at monitoring and punishment, we examine this question from a different angle – how do differences in wage contracts affect employee misbehavior? Second, we expand the analysis of misbehavior to include actions against an often-overlooked agent type – other employees.

The remainder of our study proceeds as follows. The next section provides the experimental design. Section 3 summarizes the field experimental results. Section 4 concludes.

2. Experimental Design

The workplace setting we use to investigate worker misbehavior is a call center. We hired 81 people to participate in a phone-a-thon to raise money for a charitable organization. Workers were hired through an online job board in Chicago. The job advertisement described the nature of the work, the length of the work-shift, our location, and the pay range (see Appendix A). The ad was designed to attract a broad array of employees – for example, those with and without call center experience. The ad requested individuals interested in the position to email us their resume. Communication with applicants before their work shift was through email, and every interaction was carefully scripted.

All applicants were offered a position for a single shift. After receiving resumes, we randomly assigned each applicant to a different wage treatment, then offered him or her a date to work a shift. Workers were informed of the pay regime after arriving at the call center, just before their shift. Wage treatments were constant within a work shift. In the event that the job applicant was unable to come on the initial assigned date, he or she was offered a different date within the same treatment.

Upon arrival at the call center, each worker was greeted by a shift coordinator and led to a desk with a phone. Once situated, a worker was given a packet that described the organization for which donations were being raised, and included the script to use for each call as well as the instructions about how to move through the list of contacts. The first

page welcomed the worker, and provided wage details, which varied across three main treatment arms. The second page provided detailed background information about the organization. The third and fourth pages had instructions for calling and marking call lists. The fifth and sixth pages included scripts for each caller to use in requesting donations, and the final page had a list of reminders of job-tasks.

Treatments were administered through changes in the wage structure, announced on the first page of the caller's instructions packet, and were thus given at the very beginning of the work shift (see Appendix B)[9]. Table 1 summarizes the three main treatments. Treatment 1, the baseline wage regime, paid a flat wage of $12 per hour, i.e., $30 total for the two and a half hour shift. Treatment 2 consisted of a relative-performance pay incentive in which the base pay was $8 per hour ($20 for the entire shift), with a bonus of an additional $8 per hour paid to the worker who raised the most in donation pledges within a pair of workers. Thus, the caller that raised the most in pledges would earn $40 for the two and a half hour shift ($16 per hour), while the losing caller of the pair would earn just $20 ($8 per hour).

Treatment 3 announced an additional bonus to both callers on top of the relative performance pay incentive: the base wage was now $16 per hour ($40 for the entire shift), with a bonus of $8 per hour paid to the worker within a pair that raised the most in donation pledges.[10] Thus, callers were paid an unexpected bonus regardless of their performance; accordingly, the losing caller earned $40 for the shift ($16 per hour) and the winning caller earned $60 ($24 per hour).

Treatment	T1	T2	T3
Wage structure	$12/hour	$8 or $16/hour depending on relative performance	$16 or $24/hour depending on relative performance

Table 1: Summary of main treatments.

---

[9] Treatments 2, 3, and 4 also have two small changes in the caller instructions to make them consistent with the wage structure: one additional sentence was added to the end of paragraph 3, and one additional sentence to the end of paragraph 5, the sentence varying by treatment (Appendix C).

[10] The extra bonus across the board to both workers pushed the pay range $20 above what was advertised when they first applied for the job, and thus should be evident to workers as an extra payment on top of what was initially promised. To ensure that callers were aware this was an extra bonus, the first sentence of the payment terms also explicitly tells the callers that we increased the wage compared to what was advertised.

Beyond testing the effects of tournament incentives and an unmerited bonus on workplace behavior, we examine whether workers have a preference between relative-performance based pay and a fixed hourly wage. To this end, we implemented an auxiliary treatment which offered workers the opportunity to select their own pay regime – either an hourly rate, as in T1 ($12 per hour), or according to relative performance, as in T2 ($8 per hour if they underperformed, $16 per hour if they outperformed their coworker). We discuss the findings of the auxiliary treatment separately.

Measurement of Outcomes

In naturally-occurring workplace settings, employee misbehavior (such as lying, cheating, or co-worker sabotage) can take a variety of forms – many diffuse or hidden, and therefore difficult to detect or measure. In order to study misbehaviors among workers, we create possibilities for individuals to break rules on the job in a manner that we can observe covertly, and subtle enough that workers should expect a low probability of being caught or punished.[11] We do this in the way that we construct the call lists (list of contacts that callers use), how we instruct callers to mark the call lists, and the way in which call lists are shared by co-workers.

We construct the call lists such that there is substantial variation in probability of generating a donation pledge across different types of potential donors. The list contains "warm" contacts—individuals who have donated to the organization in the past three years—in addition to "cold" contacts—individuals pulled randomly from published telephone directories and who have no known history of having donated to any organization. We explain to the callers that one of the two functions of their job is to expand the donor base by calling cold contacts. This divides the resources for generating donations into low- and high-productivity, as the probability of a pledge differs sharply across the two types of contacts.[12] Each page of the call list has nine contacts, following the same pattern: three repetitions of two cold contacts followed by one warm contact (Appendix D). Each list has 22 pages of contacts, or 198 total contacts.

---

[11] There is no particular reason that callers should be aware that we have the means or inclination to check whether they dialed certain phone numbers, or the length that such phone numbers were connected. (Indeed, verifying that marked contacts were dialed, and checking connection durations, is a time-intensive process that can only be completed after workers have left the call center).

[12] For example, throughout the experiment, only 0.6% of calls to a cold number were reported to have generated a pledge, compared to 7.1% of calls to warm numbers.

9

Callers are informed that they have two main jobs of equal importance – generating pledges from past donors ("warm contacts") and expanding the donor base by getting pledges from new donors ("cold contacts"). Therefore, they are instructed to go straight through the list in the order that the contacts appear. They are given two different scripts to use during the call – one for each type of contact. After each call, the caller checks a box on the call list to mark the outcome of the call. One type of outcome tells a future caller to try the contact again (e.g. "no answer"), and the other indicates the contact should not be dialed again (e.g. "wrong number"). If the contact agrees to pledge an amount, callers also specify the amount pledged. Table 2 shows how the call list looked for each contact.[13]

Name	Contact Info	Alum	Gave in 08 or 09	Response
XYZ	0123456	YES {or blank}	{Amount recently given, if alum}	☐ No Answer ☐ Machine ☐ Disconnected ☐ Wrong Number ☐ No Pledge ☐ Pledged: $ ☐ Other, Explain

Table 2: One entry in a call list.

One important dimension of misbehavior in the workplace we wanted to study throughout the experiment is material interference in the job performance of one's co-workers – i.e. sabotage. To create opportunities for coworker sabotage, we have callers share resources by passing forward untapped contacts to later shifts. For any contact that is marked as not reached by a caller on the first try, it becomes the job of a different worker to phone that contact at a later date. This indicates that a resource has not yet been used and remains productive for generating pledges. In addition, a caller may mark a contact to indicate it should *not* be tried again in the future by another caller (e.g., by indicating the contact already pledged, will not pledge, etc.). This removes the contact from the pool of potential donors to be phoned by the later caller.

The structure and appearance of each list given to every caller is always the same: it is unmarked, and has the same pattern of two cold contacts followed by one warm. We intervene to replace contacts marked as no longer productive with untapped contacts of the same type (cold for cold, warm for warm). However, callers are not informed that we take

---

[13] Caller marks "*no answer*" if no one answers the phone, "*machine*" if an answering machine or voicemail answers, "*disconnected*" if the number is no longer in service, "*wrong number*" if someone answers but the contact in the list does not live there, "*no pledge*" if the contact is reached and refuses to pledge a donation amount, "*pledged*" if the contact agrees to pledge to donate, and "*other*" to capture other call outcomes.

this extra step to make these replacements before the contacts are passed forward. This allows them to believe that they are altering the composition of contacts the next caller receives, a very reasonable supposition given the large amount of time and energy required to refresh used portions of the list. Importantly, in the relative performance pay treatments, the person with whom the list is shared is the caller's competitor.

This design links co-workers through access to a shared resource, a resource that is one of the most critical inputs to their individual production of donation pledges. Every caller has the ability to deviate from the protocol of the job in order to take advantage of the resources in the call list or to deviate from truthful markings on the call list in order to try to manipulate the quality of the list before passing it forward to the next caller. However, doing so explicitly violates the job instructions, runs against the job functions, and hurts the stated objectives of their employer.

One way that a worker can earn the wage bonus in the relative performance pay treatments is to invest greater effort in the job in order to generate more donation pledges. This is generally the objective of relative performance based rewards – to motivate workers to be more productive. In addition, however, there are other actions a worker might take in order to earn the wage bonus, but which violate the job duties and hurt the firm's goals. In particular, a caller trying to earn the bonus in T2 or T3 might skip over cold numbers entirely (never call them) and misreport the call outcome (e.g. removing from list), in order to spend as much time as possible trying to generate pledges through the warm list numbers.

An important design feature of our call lists is that there are two main types of misreported calls to skipped cold contacts, each with different implications: one leads to cheating, the other leads not only to cheating but to sabotage. On the one hand, the caller may mark the skipped contact's call outcome as "*wrong number*", "*no pledge*", or "*disconnected*". Marking the contact in this way removes it from the pool of potential donors to be contacted in the future. This breaks the rules for performing the job to confer an unfair advantage to the worker. Furthermore, it is harmful to the firm's interests, since by refusing to call cold contacts and then marking so as to remove them from the list, the caller hinders the employer's stated objective of expanding the donor base. We refer to this type of behavior as *non-sabotage cheating*.

On the other hand, a caller that entirely skips cold numbers and lies about the outcome may mark the call list so as to indicate the contact was simply not home. This would force the next caller using the list to call this cold number. That is, this is a violation of workplace rules that not only raises one's own expected performance but also *lowers* the expected performance of one's co-worker. By forcing her co-worker to use a given low-valued resource rather than use it herself, a worker would manipulate the quality of her co-worker's production resources. This is therefore not only cheating; it is sabotage: one worker is breaking the rules to materially harm the production process of another and thereby increase the chance of earning the performance reward. Note that this action directly hurts one's co-worker, but does not directly hurt the firm beyond violating its established work protocols: neither of the firm's objectives (expanding the donor base and generating pledges) should be directly affected by this action. We refer to this as *sabotage*.

Another important feature of the design is that it enables sabotage of varying degrees. In particular, a caller may try to interfere with a co-worker's productivity not only by forcing him to spend time on low-quality resources, but also by destroying his colleague's productive resources. If a caller is unable to speak with a given warm contact before the end of his shift, marking that contact to prevent it from being called in the future (e.g. "No Pledge") would increase his chance of earning the bonus. This is a more severe manipulation of a co-worker's production resources. It also clearly hurts the employer since it destroys a high-value resource likely to generate pledges. We refer to this as *severe sabotage*.

Since all of the above types of worker behavior constitute violations of workplace protocol and job responsibilities in order to gain an unfair advantage in trying to win the bonus, they fall under the general rubric of *cheating*. Acts of *sabotage* and *severe sabotage* are therefore subsets of *cheating*. We use the term *non-sabotage cheating* for actions that are cheating but do not constitute either type of sabotage (i.e. do not directly interfere with coworker productivity).

In order to detect whether callers have engaged in any of these misbehaviors, we gather dialing information from the phones after the callers have finished their shift. This information includes what phone numbers were dialed by the caller, whether the call

connected, and how long the connection lasted. Thus, we are able to determine whether workers skipped cold numbers and lied about the call outcome, or marked warm numbers to remove them from their coworkers' call list without having valid reason. This represents the basis of our key outcome measures.

Predictions

Table 3 presents the predictions from a neoclassical model. We would expect the least misbehavior in the baseline hourly wage treatment T1 as compared to T2 and T3, since there is no monetary reward in T1 that can be gained from misbehaving. Callers might nevertheless misbehave in T1 if they derive disutility from their job task. For example, it is conceivable that callers avoid calling cold numbers because they may find it less pleasant to talk to a person who has never heard of the charity than to talk to a previous donor.

In contrast, a caller in the relative performance treatments T2 and T3 should sabotage, since this will raise his expected wage from the shift. Taking this to the extreme, a caller should go straight through his call list, skipping all cold contacts and marking them such that his co-worker must call them, trying all warm contacts, and permanently removing from the call list any warm contact, even those not reached. That is, in order to maximize the expected wage, a worker should take full advantage of every opportunity to sabotage.

Treatment	T1	T2	T3
Type of Misbehavior	Possibly some misbehavior	Sabotage (including severe forms)	Sabotage (including severe forms)

Table 3: Standard treatment predictions for misbehavior.

The treatment predictions are the result of three features of our setting: a one-time work experience in an environment of minimal risk of detection and minimal effort cost of sabotage. First, there is no reason for callers to believe that we can check which numbers were actually dialed, how long they were connected, and whether this matched what they marked on the call list. Second, the shift is for a single night, and there is no way to detect misbehavior until after a caller has left the call center. The effort costs of sabotage and cheating are also virtually non-existent and both forms of misbehavior detract little to

nothing from a caller's own work output. (When performing the job honestly, a box must be checked anyway; checking the box untruthfully requires little to no additional effort.) The features of minimal detection risk and effort cost account for a broad class of possible acts of cheating and sabotage in actual workplace settings (e.g. delaying an action that is supposed to be part of your job, in order to hurt the output of a coworker; withholding information to handicap a coworker; spreading gossip to discredit a coworker; etc.).

These features also facilitate clear treatment predictions. In particular, assuming individuals are driven by wage maximization, the performance pay treatments T2 and T3 should induce sabotage, but they should not induce non-sabotage cheating. The reason is that opportunities for non-sabotage cheating are also opportunities to sabotage, a caller may choose only one or the other, and sabotage maximizes a caller's relative advantage in relation to the competing coworker – by not only enhancing his own production output but also negatively affecting his competitor's output.

Further, the comparison between T2 and T3 allows us to test the gift-exchange hypothesis and to examine how specific people's reciprocation reflexes are in the workplace. Do workers who receive the unmerited bonus respond with generosity directed specifically towards the giver, or do they reciprocate in a diffuse manner by treating everyone in their workplace better because someone gave them a gift? More specifically, does the bonus from the employer reduce cheating only in ways that directly hurt the employer's stated objectives, or does it also reduce cheating that only directly hurts co-worker competitors?

3. Experimental Results

We have a total of 81 workers who participated in the experiment. They come from a variety of backgrounds. Average age was 31 years (min 18 years, max 53 years). About two fifths (41.8%) had work experience in office settings where making and receiving phone calls was a major part of the job. Our workers also come from a variety of educational backgrounds: 8.5% have a graduate degree, 35.2% have a bachelor's, 15.5% have an Associate's degree (2-year community/junior college), 33.8% have some post-secondary schooling but without a degree, and 7% are high school graduates with no higher education. While 59% of the callers were female, 41% were male.

In total, callers indicated having attempted to contact 6,573 potential donors – 4,308 cold contacts and 2,265 warm contacts. The call connected for at least 10 seconds for 2,122 contacts (1,336 cold contacts; 786 warm contacts), for at least 30 seconds for 960 contacts (542 cold, 418 warm), and for 60 or more seconds for 417 contacts (172 cold, 245 warm). Callers indicated obtaining donation pledges from 187 contacts (27 cold, 160 warm). A total of 68 of these contacts actually donated (1 cold, 67 warm), with an average donation amount of $323 (median $250, min $50, max $1,000). Total amount raised in donations by the 81 callers was roughly $22,000.

3.1. Baseline Misbehavior: Fixed Wages

We first note that we observe worker misbehavior even in the baseline hourly wage treatment. Even when there is no clear monetary incentive to do so, many workers are misreporting by entirely skipping cold numbers but marking the call outcome as if the contact had actually been called. This not only explicitly violates the rules of the job, but also runs directly counter to the employer's stated goal of expanding the donor base. In the baseline hourly wage treatment (T1; N = 16 workers), 50% of the callers skipped a cold number at least once and lied about the call outcome to make it seem as if the number was indeed dialed.

Methodologically, the T1 findings provide a general lesson on the importance of using data to establish the baseline. In this case, the deleterious effects of relative performance pay contracts, discussed next, would be difficult to interpret if relying on standard theory alone without an empirical reference for the baseline. While perhaps surprising at first, the T1 findings provide field evidence consistent with some lab evidence showing that even in settings without pecuniary incentives, individuals may engage in unethical behavior to make themselves appear better than others (e.g. Charness et al., 2014).[14] It is also possible that calling cold contacts to convince them to donate to an organization they do not know incurs higher psychological costs than calling warm contacts who have already donated in the past, which may cause individuals to shirk job responsibilities and then lie in order to hide it.

3.2. Competition-Induced Misbehavior: Sabotage and Non-Sabotage Cheating

---

[14] It is possible the incidence in T1 could also reflect the human rate of error for this task. In either case, it is clearly preferable to use data to establish the baseline.

In the context of an explicit competition, as is the case for relative-performance based pay, skipping cold list numbers and misreporting the call outcome by marking them as if they had actually been called is more clearly interpreted as *cheating*: it raises the probability of winning an extra reward in an unfair manner. Figure 1a shows the percentage of our solicitors who take this action in the fixed hourly wage treatment compared to the relative performance wage treatment.

In the relative performance treatment (T2; N = 16 workers), almost every solicitor (94%) took this action at least once. This approximate doubling of the incidence of cheating is highly significant (χ^2-test, p = 0.006). A rank-sum test indicates the frequency distribution under T2 is also significantly higher (p=0.027).

Table 4 reports results from a Probit regression of the decision to cheat in this way. Column 1 reports coefficient estimates in a model without controls. Column 2 includes controls and observables we gathered during the job process: minor variations in call list quality (average past gift amount by warm contacts), whether the worker had previous job experience in similar settings, education level, and work activity during the shift (quantity of phone numbers dialed).[15] [16]

Both models in Table 4 confirm that T2 has a large and statistically significant effect on the probability that a caller will cheat in this way, raising the likelihood by an estimated 33.5 percentage points (column 1, p=0.0085) to 36.1 percentage points (column 2, p=0.0044). These findings lead to our first main experimental result:

Result 1: *Relative performance pay causes worker cheating that directly hurts the firm's stated interests.*

---

[15] Work experience is missing for 2 individuals. Education is missing (not included in their resumes) for 10 callers, and was imputed as being at a level of high school diploma or lower under the assumption that individuals at higher education levels are more likely to consider their educational attainment an asset, and also know that resumes are considered incomplete without education information. Omitting education from the regression does little to change the results: T2 raises the probability of cheating by an estimated 35.7 percentage points (p=0.005).

[16] The correction procedure of List et al. (2016) is not applicable with our current analysis since we use non-parametric tests and OLS with controls. Using Bonferroni and Holm corrections where applicable yield results similar to those presented.
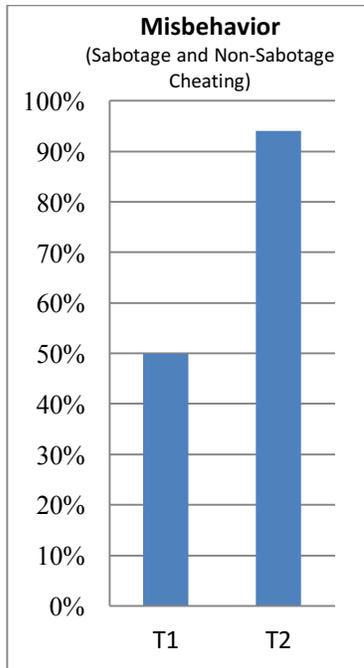
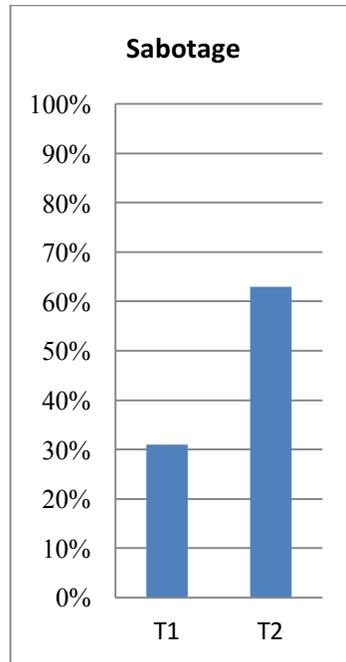**Figure 1a**: Percentage of workers that misreport cold contact outcomes.

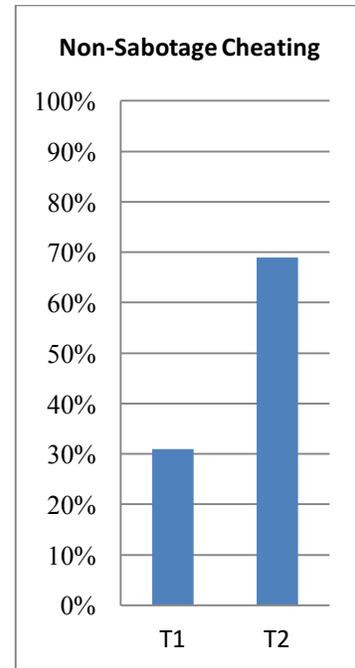**Figure 1b**: Percentage of workers that skipped cold contact and marked for coworker to call.

Figure 1c: Percentage of workers that skipped cold contact and removed from list.

3.3. Sabotage

We next turn to an analysis of whether relative performance pay induces sabotage against co-workers. Recall that workers who skip cold list numbers and lie about the outcome can either remove the cold number from the list of contacts to be called in the future or instead send their co-worker to call that number in the future, depending on how they mark the outcome. The former constitutes an act of *non-sabotage cheating*, which hurts the employer's goal of expanding the donor base. The latter manipulates the quality of the call list passed on to the co-worker, which we denote as direct *sabotage* against the co-worker.

Figures 1b and 1c illustrate the incidence of each type of action across the fixed hourly regime and relative performance pay regime. Note that a worker motivated purely by self-interest should always cheat in the relative performance pay regimes by sabotaging, since this raises her chance to earn the bonus in two ways – by raising her own expected output and also by lowering her opponent's expected output. Looking first at sabotage, we see that in the baseline treatment (T1), 31% of callers skip a cold number and mark it such that it is added to the list of contacts of one's coworker, while in the relative performance treatment (T2) 63% of callers do this (χ^2-test, p=0.077). The frequency distribution of this

action is also higher in T2 (rank-sum test p=0.097). Regression results reported in columns 1 and 3 of Table 5 indicate T2 raises the probability of engaging in sabotage-acts of this nature by an estimated 30.8 (p=0.077) to 33.2 (p=0.067) percentage points, or roughly 100%.

Turning to non-sabotage cheating, we see that in the baseline treatment (T1), 31% of callers skip a cold number and mark it such that it is permanently removed from the list while in the relative performance treatment (T2), 69% of callers do this (χ^2, p=0.034). A rank-sum test also shows that the frequency distribution of this action is significantly higher in T2 than in T1 (p=0.032). Finally, regression results in columns 2 and 4 of Table 5 confirm that T2 raises the probability of cheating in a way that does not interfere with co-worker productivity by an estimated 35.5 percentage points (p=0.035) to 40.6 percentage points (p=0.021) – an increase of 115%-131%.

These findings indicate that our solicitors are not only willing to cheat in ways that hurt the employer, but they are also willing to sabotage coworkers in order to get ahead. This leads to our second result:

> **Result 2:** *A dark side of relative performance pay is that it induces workers to not only cheat in order to unfairly enhance their own performance, but it also induces them to sabotage co-workers.*

Interestingly, however, this result is less strong than our initial prediction. Recall that in order to maximize one's wage in T2, a worker should take full advantage of opportunities to sabotage. Yet our findings suggest that, while our solicitors are indeed willing to sabotage their coworkers, they are just as likely to ignore some sabotage opportunities and instead cheat in ways that do not interfere in their coworker's productivity.[17]

How severely are workers willing to sabotage?

The purely self-interested solicitor in T2, in order to maximize her probability of winning the bonus, should go down the list at the end of her shift and mark warm numbers that she was not able to contact in such a way that they are removed from the list. This

---

[17] Cheating without sabotage of course still harms a coworker's probability of winning the competition since it increases one's own chances.

would prevent her co-worker from being able to use this resource. As this destroys a highly productive resource, we interpret this action as severe sabotage.

Despite the strong pecuniary incentive to take such actions in the relative performance pay regimes, this type of activity occurs rarely. Of the 3,301 cases in which a warm number was not dialed by a caller, only 0.6% were marked in such a way that would remove them from the next caller's list – with three occurrences in T1 and only 4 occurrences in T2. Regression coefficients reported in Table 6 show the estimated effect of each treatment on the probability of engaging in this form of severe sabotage at least once. As the estimate for T2 shows, relative performance pay has no measurable effect on whether callers sabotage in this way. This leads to our next result.

Result 3: *While workers are willing to sabotage in less severe ways under the competition-based pay regime, they appear unwilling to engage in egregious sabotage, such as destroying co-workers' productive resources, despite the ease of doing so.*

3.4. Turning off cheating and sabotage

While deviant behavior is certainly not uncommon in our experiment, there is significantly less than would be predicted by a simple neoclassical model. This suggests that there are costs to acts of cheating and sabotage beyond effort investment (and potential pecuniary costs of getting caught). In particular, this suggests that other-regarding preferences might play a role in decisions by workers of whether to cheat or sabotage on the job. If so, then is it possible to take advantage of such preferences to make the workplace more productive?

Treatment T3 (N = 32 workers) takes on this issue by providing a surprise wage bonus to the worker regardless of how she performs compared to her co-worker, in addition to the bonus based on relative performance.[18] If workers have social preferences such that they respond with positive reciprocity to gifts from the employer, then they may be less likely to violate job responsibilities in order to win the bonus. Figure 2 shows the percentage change in the proportion of workers misbehaving across the deviant behaviors

---

[18] The experiment stopped earlier than expected due to external factors that caused the charity to stop the fund-raising drive early. By chance, more shifts were assigned to T3 before this occurred, leaving more workers in the T3 wage regime.

of cheating, sabotage, and non-sabotage cheating when switching from relative performance pay (T2) to relative performance plus a surprise bonus (T3). The incidence of misbehaviors substantially drops across all three measures of workplace misbehavior.

For example, we find that the percentage of callers cheating drops from 94% to 69% – a 27% (25 pp) drop (χ^2, p=0.052). Turning to the regression results, we see in Table 4, columns 1 and 2 that, while T2 has a statistically significant effect on the probability of cheating compared to the baseline hourly wage (p<.01), T3 does not have a significant effect at conventional levels (p>0.2). When we switch from T2 to T3, the estimated probability of engaging in cheating drops by 18.3 to 19.8 percentage points, or 55% (p = .055 to p=.032, χ^2-test of difference in coefficient estimates).
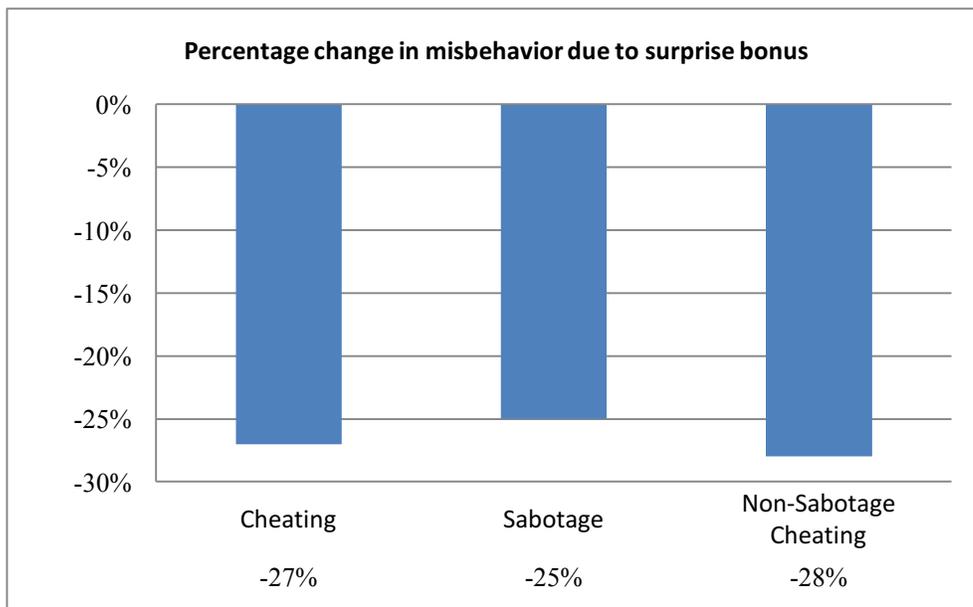

Figure 2: Percentage change in the proportion of workers that engage in each type of deviant behavior in moving from the relative performance based wage (T2) to relative performance plus an unmerited bonus (T3).

If a bonus from employers triggers positive reciprocity among employees, we may expect that it would reduce cheating that directly harms employers but not affect co-worker sabotage that harms other employees (who did nothing to merit reciprocity). This would suggest that T3, while lowering non-sabotage cheating, would not affect workers' willingness to sabotage their coworkers by skipping low-valued (cold) contacts and marking them as unavailable in order to force their coworker to call them instead.

Sabotaging co-workers in this way would directly harm one's co-worker, without directly harming the firm, and may thus be perceived as less harmful to the employer.

Intriguingly, however, Figure 2 shows that *both* types of misbehavior decrease in moving from T2 to T3. Non-sabotage cheating drops by 28% (19 pp) when moving from T2 to T3 (χ^2-test, p=0.217). Sabotage drops by almost the same amount, falling 25% (16 pp) when moving from T2 to T3 (χ^2-test, p=0.307). While not statistically significant when splitting the effect across the two different forms of misbehavior, the negative effect of T3 on cheating in general is clearly driven by drops in both types of misbehavior.

Turning to the regression results, Table 5 shows the estimated marginal effects of the wage treatments T2 and T3 as compared to T1 on the propensity to commit sabotage (columns 1 and 3) and non-sabotage cheating (columns 2 and 4), with and without controls. Examining all four columns, we see the same pattern of results: whether looking at sabotage, or non-sabotage cheating, T2 raises the likelihood of this behavior by a large and statistically significant amount. By contrast, T3 has no significant effect on either behavior compared to the baseline hourly wage regime.

The bonus thus appears to prevent both forms of worker misbehavior that are induced by the relative performance pay regime. In T3, workers behave more positively in general – not only towards the firm, but also towards co-worker competitors. This suggests that the gift may trigger a more diffuse positive reciprocity towards everyone, rather than an articulate response targeted specifically at the employer. This leads to our next result:

Result 4: *A monetary gift (bonus money) reduces employee cheating and sabotage caused by relative performance pay. Giving workers a bonus at the beginning of their shift reduces misbehavior under competition to levels similar to those observed in the fixed-wage regime.*

3.5. Reported Pledges and Actual Donations

Another way that relative-performance based remuneration may cause unwanted behavior among employees is by inducing them to intentionally misrepresent their productivity, or output, whenever they have the chance to do so. In our experiment, we create the opportunity for workers to cheat in this manner by having them write down the

dollar amount that pledgers say they will donate, giving them the option to intentionally misreport donation pledges. In the tournament treatments T2 and T3, in which payment depends on relative performance measured by reported donation pledges, it is in the pecuniary interest of callers to report (i) more frequent pledges and (ii) higher pledge amounts than actually indicated by potential donors. Both types of behavior are harmful to the employer as they result in pledge cards being sent to individuals who have not agreed to donate, or pledge cards indicating donation amounts that are higher than the amount the individual actually pledged over the phone. Both actions risk harming the organization's relationship with the donor, hurting current and future donations.

To shed light on the incidence of cheating through misreporting pledged donations, we collected two different measures of donations: donation pledges reported by callers and donation receipts from contacted individuals. Table 7 summarizes a breakdown of indicators for both donation measures across the three main treatments – overall and by caller.

In total, across the three main treatments, 2.8% of the call outcomes were marked as donation pledges (147 out of 5,131 marked call outcomes). However, only 32.4% (47) of self-reported donation pledges became actual donations. The average caller-reported pledge amount was $234.90, ranging from $20 to $750, while the average actual amount given by pledgers was $314.30, ranging from $50 to $1,000.[19]

To focus on cheating, we first examine whether any of the pay regimes affect the frequency with which a contact is identified by a caller as pledging and then does not actually donate. Each caller had on average 1.53 non-performing pledges (cases in which a contact was marked by the caller as pledging but then did not actually donate). The breakdown across treatments is 1.125 in T1, 1.625 in T2, and 1.688 in T3. While the means are slightly higher in T2 and T3, the per-caller frequency distributions across the three treatments are not significantly different from each other (Mann Whitney U tests T1 vs. T2, p=0.36; T1 vs. T3, p=0.43; T2 vs. T3, p=0.96). Overall, 68.75% of callers ever indicated a pledge that was not followed by a donation: 68.75% in T1, 62.5% in T2, and 72% in T3. The differences across treatments are not significant (all pairwise χ^2 tests, p >

---

[19] Several callers left pledge amounts blank – i.e. they specified pledge amounts for some of their pledges, but not for others. The total reported pledge amount for these callers does not include amounts for the pledges of unspecified value. For such callers, total reported pledge amount is thus typically less than the total amount actually donated by their pledgers.

0.51). Regression results reported in Table 8 confirm that neither T2 nor T3 had a significant effect on a caller's likelihood of having a non-performing pledge or on her number of non-performing pledges. We conclude neither T2 nor T3 changed the likelihood or frequency of callers indicating pledges that do not ever materialize.

To test for evidence of indicating a higher pledged donation amount than actually stated, we compare caller-reported pledge amounts to actual donation amounts by donors. Among the 47 reported pledges that were followed by a donation, pledge amounts were specified for 33.[20] Only four donations received were lower than the reported pledge, with no significant differences across treatments (all pairwise Mann-Whitney and χ^2 tests, p > 0.14).

Overall, we therefore see little evidence of treatment effects on cheating through misreported donation pledges. This may indicate another limit to the willingness of workers to cheat. Despite the pecuniary incentive to over report donations in the tournament treatments, there is no evidence that T2 or T3 cause this behavior. This suggests workers in our experiment may have an aversion to misbehavior that risks damaging the firm's relationship with current donors, or an aversion to bothering individuals with false pledge cards for amounts they did not pledge.

Pay regime effects on performance

Interestingly, the most striking differences in reported and actual donations across treatments are found in performance indicators rather than indicators of cheating. As shown in Table 7, the number of self-reported donations per caller more than doubles when moving from T1 to T3 (Mann-Whitney test, p=0.029), and the number of actual donations per caller goes up by a factor of almost 10 (Mann-Whitney test, p=0.005). The total value of actual donations raised per caller in moving from T1 to T3 rises even more, increasing by a factor of almost 20 (Mann-Whitney test, p=0.004). The probability that a caller generates an actual donation is also over four times higher in T3 than in T1 – 53.1% of callers in T3 generate at least one donation, compared to 12.5% of callers in T1 (χ^2, p<0.01).

---

[20] While callers were encouraged to obtain a specific dollar-value pledge, donors had the option to pledge an unspecified amount.

Table 9 reports results from six different regression models to test for treatment effects on indicators of worker performance. Looking across columns 1 through 4 at the coefficient estimates for T3, we see that it increased the number of reported pledges by an estimated 1.52, the number of actual donations generated by 0.84, total amount raised by $296, and the probability of generating at least one donation by an estimated 40.1 percentage points—all statistically significant at the .05 level.

Performance measures are also higher in T2 than in T1. However, only one of the differences is statistically significant at conventional levels. As shown in Table 9, column 1, T2 has a marginally significant effect on the number of reported pledges, raising it by 1.33 (p=0.094). The increases in performance indicators in moving from T2 to T3 are also substantial. For example, the distribution of total donation amounts generated per caller is higher in T3 than in T2 (Mann-Whitney test, p=0.063). In addition, column 6 of Table 9 shows another dimension along which the T3 wage scheme raises worker performance over T2. A feature of our experimental design was the inclusion of a means for workers to invest extra effort beyond the explicit instructions for the job task, effort that would have no clear benefit to themselves. We did this by adding a space to the calling list for callers to write comments they thought might be helpful for later workers.[21] While there are no significant differences between T1 and T2 or T1 and T3 in the use of the comment space to provide extra help to later workers, there is a large and significant difference between T2 and T3. While 62.5% of callers in T2 add comments, 87.5% of callers in T3 do so (χ^2 test, p=0.044). In addition, regression results indicate that T3 significantly increases the likelihood that a worker will invest this extra effort by an estimated 24.5 percentage points over T2 (Table 9, column 6, χ^2 test p=0.033). Combining all of these insights leads to our next result:

> **Result 5:** *Relative performance pay contracts, relative to a fixed wage scheme, appear capable of raising productivity along several important dimensions. However, an unexpected monetary gift on top of such contracts sharply enhances the effect, significantly boosting worker performance.*

---

[21] Callers wrote comments on the calling sheet a total of 148 times. Example comments include: "Wants to donate but next week", "Call next week on cell *{cell phone number}*", "Wife answered, call back", "Not there, call back in a few days", "Call back another time in Spanish", "Out of town (call next week)".

3.6. Patterns by gender

Several studies have found that the work output of men can be substantially more responsive to competition-based incentives than the work output of women (Gneezy et al., 2003; Gneezy and Rustichini, 2004; Shurchkov, 2012). It has been hypothesized that this results from a greater desire to win competitions among men than among women. If men take extra measures to win in competitions compared to women, then this may not be limited to ethical or sanctioned activities. A greater competitive drive among males may make them more prone than women to also break the rules in order to win.[22]

We therefore examine the data for any evidence of gender differences in the response of misbehavior to our different wage treatments. Data patterns are intriguing. We first note that the proportion of men who skip a cold contact and mark it as if contacted increases by 67 percentage points in moving from in T1 to T2 (Fisher's exact 2-sided, p=.07; 1-sided, p=.049). In addition, the average number of times men engage in this activity rises from 0.8 times per worker in T1 to 6 times per worker in T2, and the frequency distribution is significantly higher among men when moving from T1 to T2 (Wilcoxon rank-sum, p=.026).

For women on the other hand, the proportion that ever skip a cold contact and mark it as if contacted increases by only 21 percentage points when moving from T1 to T2 (Fisher's exact 2-sided, p=.52; 1-sided, p=.30). And, the average number of times women engage in this activity actually falls slightly from 4 times per caller in T1 to 3.3 times per caller in T2, while the frequency distribution is no higher under T2 than T1 (Wilcoxon rank-sum, p=.70).

This suggests that there may be a flip side to the findings in the literature on gender differences in performance response to competition incentives: while males' output can increase under competition, their propensity to misbehave in the workplace in order to win may also increase. These findings should be interpreted as suggestive rather than conclusive. While significant at conventional levels, and consistent with

---

[22] This is also consistent with an evolution-based explanation for the origins of gender differences in competitiveness, as advanced for example in Gneezy and Rustichini (2004). If males are more competitive in part because of the inherent advantage of competitiveness for men's evolutionary success, the drive to outcompete opponents is less likely to be limited to actions that fall within the rules of a specific job in modern-day society.

similar findings on gender differences in the competition-response of sanctioned activities to enhance own performance, the sample size when examining by gender becomes small. For this reason, we do not make this a formal result. Replication and further examination of this pattern and the extent to which it interacts with the task-environment represents a fruitful avenue for future research.

3.7. Choice of pay scheme

As aforementioned, we included an auxiliary treatment that allowed workers to select whether to be paid according to the T2 pay regime ($8 per hour base wage, an extra $8 per hour if outperform coworker) or the T1 regime ($12 per hour). Of the 17 workers in this treatment, only one chose the relative performance pay regime (T2). This is consistent with recent evidence from other field studies that workers may have a general aversion to relative performance pay when implemented in workplaces (Flory et al., 2015a; Samek, 2015).[23] However, our finding in this auxiliary treatment should be interpreted with some caution, as it could also be driven by the fact that the bonus is paid a few days later.

Turning to worker misbehavior in the auxiliary treatment, by and large there is not strong evidence that it was significantly different from behavior in the fixed hourly wage regime. Mann-Whitney U-tests show the frequency distributions in T1 and the auxiliary group are not significantly different, whether looking at cheating (.175), sabotage (p=.325) or non-sabotage cheating (p=.107). Chi-squared tests also show no significant difference between the two groups in the percentage of callers that engage in cheating (p=.114, sabotage (p=.208) or non-sabotage cheating (p=.112). Finally, we see in Table 4, column 1 that the estimated effect of being in the wage choice group on the probability of cheating is not significant at conventional levels.

Interestingly, however, there is some evidence of a significant difference in the propensity to cheat across the two groups when adding controls. As column 2 of Table 4 shows, the estimated effect of being in the wage regime choice group is a 22 percentage point increase in the probability to cheat, significant at the .10-level. This may suggest that the mere mention of workplace competition could be sufficient to at

---

[23] The proportion of participants selecting competition in lab experiments is often on the order of 50%. (For example, see Niederle and Vesterlund, 2007; Flory et al., 2015b).

least partially trigger worker behaviors induced by tournament-incentives. Turning to Table 5, the coefficient estimates are also relatively large for sabotage (columns 1 and 3) and non-sabotage cheating (columns 2 and 4), though significant only for non-sabotage cheating with controls added.

These data patterns are similar to findings in Charness et al. (2014). They find that simply ranking task performance in the lab causes subjects to lie, artificially inflate their own performance, and take actions to lower the performance of other subjects working on the same task – even when there is no monetary gain from performing better than others. If workers in the wage choice group believe their output will be placed in a relative ranking regardless of their compensation choice, this may cause them to cheat as part of inherent status-seeking behavior.

Finally, turning to worker performance, we find some intriguing data patterns. The wage choice regime appears to significantly improve several measures of worker performance. Columns 2-4 of Table 9 show that workers in the auxiliary treatment generated more donations, raised more money, and had a higher probability of generating a donation than workers in the fixed wage treatment.

The findings in our auxiliary treatment are consistent with the hypothesis that simply alluding to workplace competition may trigger status-seeking behavior by workers which affect both undesirable and desirable workplace behavior. However, the findings of the auxiliary treatment should be interpreted with caution, as two elements simultaneously change in moving from T1 to T4. First, we give participants a *choice* of the wage contract, and second, we provide a particular choice set (fixed wage versus relative performance) – each of which may have their own effects.[24] We believe that the separation of these two elements and their impact on workplace behavior is an interesting avenue for future research.

4. Discussion

---

[24] It is possible that already providing the choice of the wage contract has an impact on workplace behavior. For example, it may provide workers with a signal that the employer is less top-down and this may correlate with expectations about monitoring and trust in workers. If workers expect less monitoring and more trust in T4 than in T1, it could explain why there is slightly more misbehavior and better performance assuming self-interested workers increase their misbehavior and reciprocal workers reciprocate trust by working harder. On the other hand, it is possible that the particular choice set – fixed wage or tournament – sends a signal that the employer values competition and this may trigger status-seeking behavior.

Using a field experiment in the workplace, we document misbehavior by employees with a high degree of specificity. We have also shown how an employer can take advantage of incentives uncovered by behavioral economics (e.g. social preferences) to intentionally affect and mitigate misbehavior that is damaging to its interests. To our knowledge, this is the first field experimental evidence of this kind.

The wage regime effects we find on worker behavior and performance have important implications for the relative costs and benefits of the different pay structures. The pure tournament regime causes a large and statistically significant rise in the damaging behaviors of cheating and sabotage on the one hand, with little to no significant effects on worker productivity and performance indicators on the other. The unmerited-bonus plus tournament wage scheme creates improvements in worker behavior on two accounts.

First, it eliminates the negative workplace behaviors of cheating and sabotage that the pure tournament wage creates. Second, it does so without negatively affecting indicators of worker performance and productivity with respect to T2. On the contrary, while the boost in performance measures from T1 to T2 are not significant, the positive effects in moving from T1 to T3 are large and significant. Furthermore, several performance indicators are significantly higher in T3 than in T2. To summarize, T3 significantly raises measures of worker performance, without inducing misbehavior.

One surprising finding from the experiment is that the gift treatment (T3) caused a reduction in both of the major forms of deviant behavior – cheating that directly hurts the employer and sabotage that directly hurts one's coworker. One might have expected that only the former and not the latter would have been affected by this treatment, since it was the employer that provided the gift and not one's coworker. However, receiving the unmerited bonus appears to have caused a positive response from workers toward not only the firm but also fellow employees. We are not aware of any studies in the gift exchange literature that looks at whether reciprocal responses are restricted to the one providing the gift or might extend to other agents. Our finding that T3 inhibits behavior that would harm the employer as well as behavior that would harm coworkers suggests that specificity in targeting reciprocity-response behavior may be an interesting line of future research. An alternative interpretation of the effects of T3 on sabotage is that the

unmerited bonus triggered reciprocity that was expressed as stricter adherence to the employer's rules, irrespective of who the deviant behavior would harm. Further research on specificity of reciprocity responses would yield valuable insights in this regard.

In addition to the beneficial effects of the unmerited-bonus wage regime in eliminating worker cheating and sabotage induced by relative performance pay, we also find a surprisingly large positive effect on worker performance. Its reduction of misbehavior harmful to the firm combined with its sharp positive effects on performance suggest combining relative performance pay with a bonus regardless of performance could lead to large improvements in worker behavior. Further research on the effects of this type of incentive represents a promising avenue for future research.

Some closing words of caution are in order. It is possible that the extent of worker misbehavior interacts with the job setting and the nature of work being performed. The fact that the job involves making calls on behalf of a non-profit organization as part of a fund-raising drive may decrease workers' willingness to misbehave in ways that hurt their employer. If so, our results on the extent to which relative performance incentives induce unethical behavior damaging to the firm and its objectives may represent a lower bound on likely effects. For example, in a competition among salespeople in a for-profit firm, it is possible we would observe a higher incidence and more egregious forms of cheating and sabotage.

Also, the depersonalized interactions of workers and structure of the work performed constrains our findings to the class of workplace sabotage that does not permit retaliation. Workers are compared to callers in the following shift, they can only pass the list forward, and we give each worker a calling list that appears untouched by previous callers, so that callers act as first-movers. This is similar to cases where the production process is such that one worker can harm the output of a second worker, while the latter has little or no means to harm the first worker's output – e.g. working at an earlier stage in a production process. While our field setting provides an environment to clearly identify the impacts of different wage regimes in the initial base case of unilateral sabotage, it is unclear whether our findings extrapolate to more complex work environments, where there are opportunities for bilateral or multilateral sabotage.

We therefore consider these findings an important first step in integrating field experiments with naturally occurring workplace settings to learn more about the influence of different wage contracts on unethical behavior by workers that damage the firm's interests. As such, we view this study as the opening of a critical discussion vital to the design of incentives and workplace policy, and additional research in other contexts as a fruitful and exciting avenue for future research to complement the insights generated by our study. For example, it seems important to investigate whether sabotage is more or less pronounced in firms where there are depersonalized channels through which bilateral sabotage can take place.

References

Akerlof, G. A. (1982). Labor contracts as partial gift exchange. *The Quarterly Journal of Economics*, *97*(4), 543-569.

Akerlof, G. A., & Yellen, J. L. (1988). Fairness and unemployment. *The American Economic Review*, *78*(2), 44-49.

Akerlof, G. A., & Yellen, J. L. (1990). The fair wage-effort hypothesis and unemployment. *The Quarterly Journal of Economics*, *105*(2), 255-283.

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

Barkema, H. G. (1995). Do top managers work harder when they are monitored?. *KYKLOS-BERNE-*, *48*, 19-19.

Benabou, R., & Tirole, J. (2011). Laws and norms (No. w17579). *National Bureau of Economic Research*.

Carpenter, J., Matthews, P. H., & Schirm, J. (2010). Tournaments and Office Politics: Evidence from a Real Effort Experiment. *The American Economic Review*, *100*(1), 504-517.

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

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

Cohn, A., Fehr, E., Herrmann, B., & Schneider, F. (2014a). Social comparison in the workplace: evidence from a field experiment. *Journal of the European Association of Economics*.

Cohn, A., Fehr, E., & Goette, L. (2014b). Fair wages and effort provision: Combining evidence from a choice experiment and a field experiment. *Management Science*, *61*(8), 1777-1794.

Cooter, R. D. (1998). Expressive law and economics. *University of Munich, Center for Economic Studies, Working Paper*, (161).

Courty, P., & Marschke, G. (2004). An empirical investigation of gaming responses to explicit performance incentives. *Journal of Labor Economics*, *22*(1), 23-56.

Dato, S., & Nieken, P. (2014). Gender differences in competition and sabotage. *Journal of Economic Behavior & Organization*, *100*, 64-80.

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

Duflo, E., R. Hanna, and S. Ryan. (2008). "Monitoring Works: Getting Teachers to Come to School." CEPR Discussion Paper No. DP6682.

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

Enzle, M. E., & Anderson, S. C. (1993). Surveillant intentions and intrinsic motivation. *Journal of Personality and Social Psychology*, *64*(2), 257.

Fehr, E., Kirchsteiger, G., & Riedl, A. (1993). Does fairness prevent market clearing? An experimental investigation. *The Quarterly Journal of Economics*, *108*(2), 437-459.

Flory, J., Leibbrandt, A. & List, J. (2015a) Do Competitive Work Places Deter Female Workers? A Large-Scale Natural Field Experiment on Gender Differences in Job-Entry Decisions. *The Review of Economic Studies, 82*(1), 122-155.

Flory, J., Gneezy, U. & List, J. (2015b) Gender, Age, and Competition: the Disappearing Gap. Working paper.

Frey, B. S. (1993). Does monitoring increase work effort? The rivalry with trust and loyalty. *Economic Inquiry*, *31*(4), 663-670.

Gilpatric, S. M. (2011). Cheating in contests. *Economic Inquiry*, *49*(4), 1042-1053.

Gneezy, U., & List, J. A. (2006). Putting behavioral economics to work: Testing for gift exchange in labor markets using field experiments. *Econometrica*, *74*(5), 1365-1384.

Gneezy, U., Niederle, M.,& Rustichini, A. (2003). Performance in competitive environments: Gender differences, *The Quarterly Journal of Economics, 118*(3), 1049-1074.

Gneezy, U., & Rustichini, A. (2004). Gender and competition at a young age. *The American Economic Review*, *94*(2), 377-381.

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

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

Konrad, K. A. (2000). Sabotage in rent-seeking contests. *Journal of Law, Economics, and Organization*, *16*(1), 155-165.

Konrad, K.A. (2009). *Strategy and dynamics in contests.* Oxford University Press.

Kube, S., Maréchal, M. A., & Puppe, C. (2012). The currency of reciprocity: Gift exchange in the workplace. *The American Economic Review*, *102*(4), 1644-1662.

Lazear, E. P. (1989). Pay equality and industrial politics. *Journal of political economy*, 561-580.

Lazear, E. P., & Rosen, S. (1981). Rank-Order Tournaments as Optimum Labor Contracts. *The Journal of Political Economy*, 841-864.

List, J. A., & Rasul, I. (2011). Field experiments in labor economics.*Handbook of labor economics*, *4*, 103-228.

List, J. A., Shaikh, A. M., & Xu, Y. (2016). Multiple Hypothesis Testing in Experimental Economics (No. w21875). *National Bureau of Economic Research*.

Nagin, D., Rebitzer, J., Sanders, S., & Taylor, L. (2002). Monitoring, Motivation, and Management: The Determinants of Opportunistic Behavior in a Field Experiment. *The American Economic Review*, 92(4), 850-873.

Niederle, M., & Vesterlund, L. (2007). Do women shy away from competition? Do men compete too much?. *The Quarterly Journal of Economics*, *122*(3), 1067-1101.

Olken, Benjamin A. (2007). "Monitoring corruption: evidence from a field experiment in Indonesia." *Journal of political Economy* 115.2: 200-249.

Posner, R. A. (1998). Rational choice, behavioral economics, and the law. *Stanford Law Review*, 1551-1575.

Preston, I., & Szymanski, S. (2003). Cheating in contests. *Oxford review of economic policy*, *19*(4), 612-624.

Schwieren, C., & Weichselbaumer, D. (2010). Does competition enhance performance or cheating? A laboratory experiment. *Journal of Economic Psychology*, *31*(3), 241-253.

Shurchkov, O. (2012). Under pressure: gender differences in output quality and quantity under competition and time constraints. *Journal of the European Economic Association*, *10*(5), 1189-1213.

Slemrod, Joel, Marsha Blumenthal, and Charles Christian. (2001). "Taxpayer response to an increased probability of audit: evidence from a controlled experiment in Minnesota." *Journal of Public Economics* 79.3: 455-483.

Sunstein, C. R. (1996). On the expressive function of law. *University of Pennsylvania law review*, *144*(5), 2021-2053.

Vandegrift, D., & Yavas, A. (2010). An experimental test of sabotage in tournaments. *Journal of Institutional and Theoretical Economics JITE*, *166*(2), 259-285.

Zitzewitz, Eric. (2012). "Forensic economics." *Journal of Economic Literature* 50.3: 731-769.

Tables

Table 4. Probability to engage in deviant behavior: cheating

	(1) Cheat	(2) Cheat
T2	0.335***	0.361***
	(0.0770)	(0.0732)
T3	0.152	0.163
	(0.116)	(0.123)
Wage Choice Group	0.198	0.217*
	(0.105)	(0.105)
Past Gift Amounts		0.000789
		(0.000550)
Work Experience		-0.0746
		(0.106)
Shift Activity Level		0.112
		(0.118)
Education		0.0945
		(0.164)
χ^2 test p-value of $H_0: \beta_{T2} = \beta_{T3}$	.055	.032
Observations	81	79

Notes: This table reports estimated marginal effects from a Probit regression. The omitted category is T1, the treatment with a fixed hourly wage. T2 is the treatment in which compensation depends on relative performance. T3 is the treatment in which compensation depends on relative performance, but in which all callers also received a surprise bonus. Number of individuals drops by 2 when including controls due to missing information on work experience. *Wage Choice Group* is an indicator for callers that were in the group allowed to select their own compensation scheme. *Past Gift Amounts* is the average donation amount the prior year by warm contacts in the worker's calling list. *Work Experience* is an indicator for whether the worker had previous experience working in a call center or office setting with significant phoning work. *Shift Activity Level* indicates whether a worker dialed at least 100 times during the work shift. *Education* is an indicator for whether the caller had any education beyond a high school diploma. Standard errors are in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 5. Probability to engage in deviant behavior: sabotage and non-sabotage cheating

	(1)	(2)	(3)	(4)
	Sabotage	Non-Sabotage Cheating	Sabotage	Non-Sabotage Cheating
T2	0.308*	0.355**	0.332*	0.406**
	(0.158)	(0.142)	(0.159)	(0.141)
T3	0.163	0.192	0.143	0.193
	(0.154)	(0.152)	(0.162)	(0.163)
Wage Choice Group	0.220	0.270	0.167	0.329*
	(0.167)	(0.155)	(0.177)	(0.156)
Past Gift Amounts			-1.80e-05	0.000645
			(0.000601)	(0.000613)
Work Experience			0.140	-0.0881
			(0.115)	(0.119)
Shift Activity Level			0.0107	0.192
			(0.149)	(0.147)
Education			0.157	-0.0883
			(0.155)	(0.164)
Observations	81	81	79	79

Notes: This table reports estimated marginal effects from a Probit regression. The omitted category is T1, the treatment with a fixed hourly wage. T2 is the treatment in which compensation depends on relative performance. T3 is the treatment in which compensation depends on relative performance, but in which all callers also received a surprise bonus. Number of individuals drops by 2 when including controls due to missing information on work experience. *Wage Choice Group* is an indicator for callers that were in the group allowed to select their own compensation scheme. *Past Gift Amounts* is the average donation amount the prior year by warm contacts in the worker's calling list. *Work Experience* is an indicator for whether the worker had previous experience working in a call center or office setting with significant phoning work. *Shift Activity Level* indicates whether a worker dialed at least 100 times during the work shift. *Education* is an indicator for whether the caller had any education beyond a high school diploma. Standard errors are in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 6. Probability to engage in deviant behavior: severe sabotage

VARIABLES	(1) Severe Sabotage	(2) Severe Sabotage
T2	-0.0685	-0.0419
	(0.130)	(0.141)
T3	0.115	0.0684
	(0.127)	(0.131)
Wage Choice Group	-0.0774	-0.129
	(0.127)	(0.114)
Past Gift Amounts		0.000697
		(0.000463)
Work Experience		0.133
		(0.0936)
Shift Activity Level		0.0102
		(0.115)
Education		0.0550
		(0.113)
Observations	81	79

Notes: This table reports estimated marginal effects from a Probit regression. The omitted category is T1, the treatment with a fixed hourly wage. T2 is the treatment in which compensation depends on relative performance. T3 is the treatment in which compensation depends on relative performance, but in which all callers also received a surprise bonus. Number of individuals drops by 2 when including controls due to missing information on work experience. *Wage Choice Group* is an indicator for callers that were in the group allowed to select their own compensation scheme. *Past Gift Amounts* is the average donation amount the prior year by warm contacts in the worker's calling list. *Work Experience* is an indicator for whether the worker had previous experience working in a call center or office setting with significant phoning work. *Shift Activity Level* indicates whether a worker dialed at least 100 times during the work shift. *Education* is an indicator for whether the caller had any education beyond a high school diploma. Standard errors are in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 7. Summary Statistics on Pledges and Donations

	Combined	T1	T2	T3
Panel A: Overall				
Percentage of call outcomes marked as a donation pledge	2.8%	1.5%	2.7%	3.4%
Percentage of pledges that became actual donations	32.4%	10.0%	23.5%	40.7%
Mean reported pledge amount	$235	$250	$164	$260
Mean amount actually donated by donors (intensive margin)	$314	$175	$150	$357
Panel B: Caller-level				
Reported pledges per caller (mean)	2.30	1.25	2.19	2.88
Reported pledges that became donations (mean)	0.73	0.13	0.50	1.16
Reported pledges that did not become donations (mean)	1.53	1.13	1.63	1.69
Percentage of callers that report at least one pledge	76.6%	68.8%	68.8%	84.4%
Percentage of callers with at least one performing pledge	37.5%	12.5%	31.3%	53.1%
Percentage of callers with at least one non-performing pledge	68.8%	68.8%	62.5%	71.9%
Reported total value of donations generated per caller (mean)	$323	$156	$225	$455
Actual total value of donations generated per caller (mean)	$231	$22	$75	$413

Notes: Two pledges are dropped from the second row of panel A, since there are two pledges for which realized donation data is missing – one for T2 and one for T3.

Table 8. Deviant behavior: Misreporting pledge amounts

VARIABLES	Probit (1) Any non-perf. pledge	Tobit (2) No. non-perf. pledges	Probit (3) Any non-perf. pledge	Tobit (4) No. non-perf. pledges
T2	-0.0613	0.468	-0.0548	0.654
	(0.168)	(0.699)	(0.177)	(0.722)
T3	0.0316	0.650	0.0730	0.800
	(0.140)	(0.603)	(0.145)	(0.627)
Wage Choice Group	0.0184	-0.0469	0.0788	0.156
	(0.159)	(0.688)	(0.150)	(0.710)
Past Gift Amounts			-0.00107*	-0.00134
			(0.000550)	(0.00233)
Work Experience			-0.0898	-0.139
			(0.107)	(0.448)
Shift Activity Level			0.0497	-0.283
			(0.132)	(0.564)
Education			-0.146	-0.484
			(0.124)	(0.607)
Constant		0.729		1.445
		(0.497)		(0.980)
Observations	81	81	79	79

The omitted category is T1, the treatment with a fixed hourly wage. T2 is the treatment in which compensation depends on relative performance. T3 is the treatment in which compensation depends on relative performance, but in which callers also received a surprise bonus. *Wage Choice Group* is an indicator for callers that were in the group allowed to select their own compensation scheme. *Marked Contacts* is the number of all contacts marked by the caller. *Pledged* is the number of pledges the caller collected (self-reported). *Past Gift Amounts* is the average donation amount by warm contacts in the calling list during the prior year. *Work Experience* is an indicator for whether the worker had worked in a call center or office setting with significant phoning work. *Non-Dialed Warm* is the quantity of warm contacts not dialed by the caller. Standard errors are in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 9. Wage regime effects on worker performance

VARIABLES	(1) OLS Reported Pledges	(2) OLS Number Donations	(3) OLS Total Donations	(4) Probit Generated Donation	(5) OLS Reported Amounts	(6) Probit Give Extra Help
T2	1.329*	0.619	148.5	0.316	189.3	-0.170
	(0.783)	(0.433)	(166.7)	(0.197)	(191.6)	(0.168)
T3	1.520**	0.838**	296.1**	0.401**	243.4	0.0749
	(0.676)	(0.374)	(144.0)	(0.164)	(165.5)	(0.116)
Wage Choice Group	1.114	1.025**	350.4**	0.411**	249.6	0.0849
	(0.761)	(0.421)	(162.1)	(0.173)	(186.3)	(0.113)
Past Gift Amounts	0.00548**	0.00560***	2.495***	0.00157**	2.165***	2.90e-05
	(0.00251)	(0.00139)	(0.535)	(0.000635)	(0.615)	(0.000450)
Work Experience	-0.222	-0.152	16.69	0.0710	-47.45	-0.0486
	(0.485)	(0.268)	(103.4)	(0.120)	(118.8)	(0.0887)
Shift Activity Level	-0.0493	0.231	117.7	0.0710	113.4	0.0529
	(0.618)	(0.342)	(131.7)	(0.161)	(151.4)	(0.102)
Education	-0.704	-0.352	-159.5	-0.0668	-157.2	
	(0.668)	(0.369)	(142.4)	(0.172)	(163.6)	
Constant	0.672	-0.792	-422.6*		-214.0	
	(1.079)	(0.597)	(229.9)		(264.2)	
Observations	79	79	79	79	79	79
R-squared	0.162	0.303	0.354		0.226	

Notes: The response variables are the number of pledges reported by each caller (column 1), the number of donations actually generated (column 2), the total amount of money raised (column 3), whether the caller generated at least one donation (column 4), the average reported pledge amount (column 5), and whether the caller ever gave extra help to her coworkers (column 6). Columns 1, 2, 3, and 5 are linear regressions; columns 4 and 6 are Probit regressions. *Education* had to be dropped from the final regression since all individuals in the low education group predicted success in the Probit perfectly. Standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Appendix A:
Job Advertisement

Posting Title: Seeking Callers for Phone-A-Thon (*{Location}*)
We are seeking individuals to help with a fund-raising phone-a-thon. No experience necessary. However, you should be comfortable speaking over the phone.

Employment is for a single 2.5-hour shift on a weekday evening. We provide brief training before the calls. The pay ranges from $20-$40 for the entire shift. We are located on *{location}*.

If you are interested, please email a short resume to:
{Name of coordinator} [Email address]

Appendix B:
Description of Payments, by Treatment Arm

PAYMENT

{Treatment 1}
The wage is $12/hour. So **you will be paid $30** for the 2.5 hours you are here this evening. We will pay you in cash at the end of the shift.

{Treatment 2}
The base wage is $8/hour. So **you'll earn at least $20** for the 2.5 hours you work tonight. You may also earn a bonus, depending on the amount you raise in pledges. The pledges you raise tonight will be compared with one caller in the next shift. If you raise more in pledges than the other caller, you earn a bonus of $20 for the entire shift. That is, whoever raises the most between you two earns $16/hour, **which means you can earn up to $40** for the entire shift.

We will pay you the base wage in cash at the end of tonight's shift. If you earn the bonus, we can either pay you in cash if you'd like to return later, or we can send you a check if you prefer. If you prefer to receive the bonus by check, please be sure to give us the address to send it to.

{Treatment 3}
We increased the wage compared to what was advertised. The base wage is $16/hour. So **you'll earn at least $40** for the 2.5 hours you work tonight. You may also earn a bonus, depending on the amount you raise in pledges. The pledges you raise tonight will be compared with one caller in the next shift. If you raise more in pledges than the other caller, you earn a bonus of $20 for the entire shift. That is, whoever raises the most between you two earns $24/hour, **which means you can earn up to $60** for the entire shift.

We will pay you the base wage in cash at the end of tonight's shift. If you earn the bonus, we can either pay you in cash if you'd like to return later, or we can send you a check if you prefer. If you prefer to receive the bonus by check, please be sure to give us the address to send it to.

{Auxiliary Treatment: Wage Choice}

For your wage, you can choose between the following two options:

1. The wage is $12/hour. So **you will be paid $30** for the 2.5 hours you are here this evening. We will pay you in cash at the end of the shift.

OR

2. The base wage is $8/hour. So **you'll earn at least $20** for the 2.5 hours you work tonight. You may also earn a bonus, depending on the amount you raise in pledges. The pledges you raise tonight will be compared with one caller in the next shift. If you raise more in pledges than the other caller, you earn a bonus of $20 for the entire shift. That is, whoever raises the most between you two earns $16/hour, **which means you can earn up to $40** for the entire shift.

    We will pay you the base wage in cash at the end of tonight's shift. If you earn the bonus, we can either pay you in cash if you'd like to return later, or we can send you a check if you prefer. If you prefer to receive the bonus by check, please be sure to give us the address to send it to.

Please indicate which wage method you want: 1. _____   2._____

{All Treatments}

All callers are hired for a single shift. Tonight is the only night you will be working for the phone-a-thon.

Appendix C:
Caller Instructions

INSTRUCTIONS FOR CALLING

There are two goals of this fund-raising drive:

1) Collecting donation pledges from past donors; and
2) Increasing the donor base by gathering donation pledges from *new* donors.

Both goals are important. Your phone list therefore has a mix of people who have donated in the past and new people who have probably never heard of *{Organization Name}* before. Since it is important to contact both types of possible donors, your job is to call each phone number in the order it appears, starting at the first number and going down the list. As you go down the list, you can clearly see who has given in the past, and how much they have given, under the column "Amount Gave Recently". A contact with a blank in this column has never given before. Under the column "Alum", you can also see which people are past graduates of the *{Organization Name}* program (who should be very familiar with the program), and which people have probably never heard of *{Organization Name}* before. Knowing which type of contact you are about to call will help you to prepare ahead of time for each call.

You will be sharing the list of contacts you have been given with a caller in the following shift. For any phone number where the contact is marked as not reached, the number will be passed on to a caller in the next shift. It is that caller's job to try that contact again at a different time. *{Treatments 2 and 3}* Whoever raises the highest amount in donations between the two of you earns a bonus of **$20** for the shift. *{Treatment 4}* If you choose the bonus-based wage, then whoever raises the highest amount in donations between the two of you earns a bonus of **$20** for the shift.

Mark the list as follows: If no one answers the phone, check the box that says *"No Answer"*. If an answering machine or voicemail answers, do not leave a message. Just hang up, and check *"Machine"*. If the person says they will not donate, mark *"No Pledge"*. If the number for the contact named in your list is wrong, mark *"Wrong Number"*. If it's disconnected, mark *"Disconnected"*. If another family member answers, and says the contact named in the list is not home, then mark *"Other"* and write *"Answered by someone else. Number correct, but contact not home."* You should also use *"Other"* to explain any other situation not covered by the other options to check-mark.

 If the contact named in the list answers and agrees to make a donation, mark the box *"Pledged $_____"* and fill in the amount pledged. Do your best to obtain from them an actual donation amount. This increases the chance that they will actually donate when they receive the pledge-card. If the donor resists specifying an amount, write "not specified" in the blank for the amount on your calling list. *{Treatments 2 and 3}* For determining the bonus, any donation amount that is unspecified will count as $10 towards your total amount of pledges for tonight's shift. *{Treatment 4}* If you choose the bonus-

based wage: For determining the bonus, any donation amount that is unspecified will count as $10 towards your total amount of pledges for tonight's shift.

If the contact is not home, thank whoever answered the call for their time, and hang up. Do not ask anyone besides the person named in the contact list for a donation pledge.

Do not call a phone number more than once. It is possible the contact may not want to answer the phone tonight, and it may hurt donations to badger the contacts. If you do not reach the contact, mark the appropriate box on your sheet, and continue moving through the list. The contact will have the chance to be called again the next shift, since it is the job of the caller in the following shift to try the number again when you mark they were not reached.

Remember that the job of the next caller using your list is to call any numbers for which you indicate you did not speak to the contact, so please mark your list carefully.

If you have any questions, please direct them to *{name of shift coordinator}*.

Please mark appropriate response clearly. Numbers marked "no answer" or "machine" will be passed to a different caller to try again the following work day. If you reach an answering machine, do not leave a message. As soon as you know it is an answering machine, please hang up and move to next number.

Name	Contact Info	Alum	Amt Gave in '08 or '09	Response		
{Contact Name}	{Contact Address}			☐ No Answer ☐ Wrong Number ☐ Other, Explain:	☐ Machine ☐ No Pledge	☐ Disconnected ☐ Pledged: $_____
{Contact Name}	{Contact Address}			☐ No Answer ☐ Wrong Number ☐ Other, Explain:	☐ Machine ☐ No Pledge	☐ Disconnected ☐ Pledged: $_____
{Contact Name}	{Contact Address}	YES	$250	☐ No Answer ☐ Wrong Number ☐ Other, Explain:	☐ Machine ☐ No Pledge	☐ Disconnected ☐ Pledged: $_____
{Contact Name}	{Contact Address}			☐ No Answer ☐ Wrong Number ☐ Other, Explain:	☐ Machine ☐ No Pledge	☐ Disconnected ☐ Pledged: $_____
{Contact Name}	{Contact Address}			☐ No Answer ☐ Wrong Number ☐ Other, Explain:	☐ Machine ☐ No Pledge	☐ Disconnected ☐ Pledged: $_____
{Contact Name}	{Contact Address}	YES	$250	☐ No Answer ☐ Wrong Number ☐ Other, Explain:	☐ Machine ☐ No Pledge	☐ Disconnected ☐ Pledged: $_____
{Contact Name}	{Contact Address}			☐ No Answer ☐ Wrong Number ☐ Other, Explain:	☐ Machine ☐ No Pledge	☐ Disconnected ☐ Pledged: $_____
{Contact Name}	{Contact Address}			☐ No Answer ☐ Wrong Number ☐ Other, Explain:	☐ Machine ☐ No Pledge	☐ Disconnected ☐ Pledged: $_____
{Contact Name}	{Contact Address}	YES	$250	☐ No Answer ☐ Wrong Number ☐ Other, Explain:	☐ Machine ☐ No Pledge	☐ Disconnected ☐ Pledged: $_____