

Cassar, Lea; Meier, Stephan

**Working Paper**

## Intentions for Doing Good Matter for Doing Well: The (Negative) Signaling Value of Prosocial Incentives

IZA Discussion Papers, No. 11203

**Provided in Cooperation with:**

IZA – Institute of Labor Economics

*Suggested Citation:* Cassar, Lea; Meier, Stephan (2017) : Intentions for Doing Good Matter for Doing Well: The (Negative) Signaling Value of Prosocial Incentives, IZA Discussion Papers, No. 11203, Institute of Labor Economics (IZA), Bonn

This Version is available at:

<https://hdl.handle.net/10419/177007>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

DISCUSSION PAPER SERIES

IZA DP No. 11203

**Intentions for Doing Good Matter for  
Doing Well: The (Negative) Signaling  
Value of Prosocial Incentives**

Lea Cassar  
Stephan Meier

DECEMBER 2017

## DISCUSSION PAPER SERIES

IZA DP No. 11203

# Intentions for Doing Good Matter for Doing Well: The (Negative) Signaling Value of Prosocial Incentives

**Lea Cassar**

*University of Cologne*

**Stephan Meier**

*Columbia University, IZA and NBER*

DECEMBER 2017

Any opinions expressed in this paper are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but IZA takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The IZA Institute of Labor Economics is an independent economic research institute that conducts research in labor economics and offers evidence-based policy advice on labor market issues. Supported by the Deutsche Post Foundation, IZA runs the world's largest network of economists, whose research aims to provide answers to the global labor market challenges of our time. Our key objective is to build bridges between academic research, policymakers and society.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

## ABSTRACT

---

# Intentions for Doing Good Matter for Doing Well: The (Negative) Signaling Value of Prosocial Incentives\*

Prosocial incentives and Corporate Social Responsibility (CSR) initiatives are seen by many firms as an effective way to motivate workers. Recent empirical results seem to support the expectation that prosocial incentive, e.g. in the form of a charitable donations by the firm, can increase effort and motivation - sometimes even better than monetary incentives. We argue that the benefits crucially depend on the perceived intention of the firm. Workers use prosocial incentives as a signal about the firm's type and if used instrumentally in order to profit the firm, they can backfire. We show in an experiment in collaboration with an Italian firm, that monetary and prosocial incentives work very differently. While monetary incentives used instrumentally increase effort, instrumental charitable incentives backfire compared to non-instrumental incentives. This is especially true for non-prosocially-motivated workers who do not care about the prosocial cause but use prosocial incentives only as a signal about the firm. The results contribute to the understanding of the limits of prosocial incentives by focusing on their signaling value to the agent about the principal's type.

**JEL Classification:** D03, C93, M52

**Keywords:** prosocial incentives, Corporate Social Responsibility, signaling

**Corresponding author:**

Stephan Meier  
Graduate School of Business  
Columbia University  
710 Uris Hall  
3022 Broadway  
New York, NY 10027  
USA

E-mail: [sm3087@columbia.edu](mailto:sm3087@columbia.edu)

---

\* We thank Vanessa Burbano, Robert Dur, Constanca Esteves-Sorenson, Florian Engl, Sebastian Fehrler, Ray Fisman, Matthias Heinz, Alex Imas, Karine Nyborg, Dirk Sliwka, Mirco Tonin and participants at seminars and conference at Munich, Bonn, Frankfurt, Chicago, Yale, and Uber for valuable comments.

# 1. Introduction

Prosocial incentives and Corporate Social Responsibility (CSR) initiatives are becoming important and wide-spread tools in motivating and attracting employees. 67% of CEOs think that prosocial incentives are becoming crucial for their top talent (PricewaterhouseCoopers, 2016, p. 13). A number of empirical studies confirm those expected benefits of prosocial incentives: they increase, for example, effort and productivity (e.g., Tonin and Vlassopoulos, 2015, 2010; Imas, 2014; Charness et al., 2014; Cassar, 2017a; Armouti-Hansen et al., 2017), lower reservation wages (e.g., Nyborg and Zhang, 2013; Burbano, 2016), and increase retention (e.g., Bode et al., 2015; Carnahan et al., 2016). Those empirical results support theoretical arguments (e.g. Brekke and Nyborg, 2008; Bénabou and Tirole, 2010; Cassar, 2017b) that prosocial incentives can be motivating to workers that have prosocial preferences, i.e. for motivated agents<sup>1</sup>. The idea is that motivated agents care about prosocial incentives *directly*, e.g. a donation to a charitable foundation, and work harder or lower their reservation wage in order to support the cause. As a result, prosocial incentives have positive effects on effort but should never lower effort. Not surprisingly then, a large number of firms are engaged in CSR by, for example, having charitable foundations that donate part of the firms' profit to good causes. And many more are considering adding prosocial incentives instrumentally to the mix of HR tools in order to reap the benefits (see, e.g. Bhattacharya et al., 2008).

In this paper, we present a simple model and evidence from an experiment that illustrate the limits and potential negative effects of prosocial incentives. The main ingredient creating the behavioral predictions is that all employees use prosocial incentives as a signal about the firm's type. All agents care about the principal's type because it will determine how agents (motivated or not) are treated.<sup>2</sup> Offering charitable incentives can signal the firm's type. In

---

<sup>1</sup>Throughout the paper, we classify an agent as 'motivated' if she derives direct utility from the impact of the prosocial incentives, i.e. the agent values the charitable contribution positively independent of any signaling properties, similar to the definition in Besley and Ghatak (2005).

<sup>2</sup>Even for simple jobs, e.g. micro jobs on M-Turk, workers depend on firms treating them nicely by paying baseline payment as promised (i.e. approving their work at all and in a timely fashion), being honest about required time and effort to complete job, and rewarding them potentially with bonuses.

fact, individuals or firms that engage in prosocial acts are seen as nicer and more trustworthy (e.g. Elfenbein et al., 2012; Fehrler and Przepiorka, 2013, 2016; Kajackaite and Sliwka, 2017). Burbano (2016) shows in her study on M-Turk that workers expect to be treated better by employer that offer CSR incentives.

But the functioning of prosocial incentives depends critically on the perceived intention of the firm for using those incentives and not just the direct effect of the incentive for the good cause. If prosocial incentives are seen as being used instrumentally, i.e. the firm’s perceived intention is to benefit from the prosocial incentive, employees see the firm as less prosocial and react negatively to the incentive. Our model shows that there exists indeed a separating equilibrium in which nicer firms are offering non-instrumental charitable incentives while not prosocial firms are only offering instrumental charitable incentives. So the agent’s belief that only not-so-nice firms offer instrumental prosocial incentives is correct in equilibrium. Supporting evidence show that prosocial acts with an instrumental benefit are indeed seen as less nice (e.g. Newman and Cain, 2014; Berman et al., 2015). Our model also predicts that the effect is not the same for motivated and non-motivated agents: While the negative effect of strategic use of prosocial incentives might be offset by a motivating effect for motivated agents, the effect is predicted to be particularly negative for non-motivated workers. Using prosocial incentives strategically can therefore backfire – especially for non-motivated agents.

The paper tests this argument in an experiment in collaboration with an Italian firm. the firm hires around 3,000 workers on Amazon Mechanical Turk (M-Turk) to create taglines for products of the firm that they want to use for their English website. Workers are asked to come up with three slogans as the baseline. The treatments vary the nature of the incentives for an additional three slogans on two dimensions:

First, we vary whether incentives are private, monetary bonuses or prosocial in the form of a donation to charity. Second, we vary whether incentives (monetary or prosocial) are instrumental for the firm or not. The difference between an instrumental and not-instrumental incentive is whether its usage is clearly profit maximizing or not, i.e. whether the agents interprets the incentive as a way for the firm to instrumentally increase effort and profitability.

Exactly as a social mission is “a concrete goal or objective for the firm that reaches beyond profit maximization” (Henderson and Van den Steen, 2015, p. 327), an incentive is less instrumental if the incentive risks being not profitable. In our experiment, we vary whether the incentives (monetary or prosocial) are performance-based or not. That is, we vary whether the extra incentive is conditional or unconditional on workers doing three more slogans.<sup>3</sup> The size of the extra incentive is half the baseline pay (paid in addition to the baseline pay) and makes it clear that for the firm conditional incentives are profit maximizing as it is cheaper to get three more slogans with the extra incentive than to hire a new worker, pay the baseline pay and get three slogans this way. Performance-based incentives give therefore a stronger signal that the incentive is instrumental given that it is paid out if and only if it increases the firm’s profit. In some treatments we even made this logic explicit in order to analyze whether an explanation will make the effects stronger.

We predict substantial differences in how monetary vs. prosocial incentives work. For both types of incentives, workers use their instrumental nature as a signal about the firm’s type (for negative reaction to instrumental *monetary* incentives and control, see Gneezy and Rustichini, 2000; Carpenter and Dolifka, 2013; Falk and Kosfeld, 2006). For monetary incentives, the net effect is, however, predicted to be more positive as all workers (motivated or not) care about the monetary incentive and not just their employer’s type. For prosocial incentives, however, non-motivated workers do not care about the prosocial cause *directly*. Depending on the strength of the effect, instrumental prosocial incentives can backfire relative to non-instrumental prosocial incentives or no incentives at all.

The results of the experiment show a number of patterns consistent with our story.

First, conditional charitable incentives backfire compared to unconditional incentives. The proportion of workers who do more than the basic three slogans *decreases* by 5 percentage points when prosocial incentives are made performance-based compared to unconditional ( $p = 0.10$ ). This cannot be due to the fact that workers would rather have the charitable donation as a

---

<sup>3</sup>Throughout the paper, we will call the extra bonus/transfer “incentive” – whether it is conditional or unconditional on performance.

bonus as in both the conditional and unconditional treatment the incentive is charitable and not a monetary bonus. However, when we offer monetary incentives instead of charitable incentives, workers react very differently. Making monetary incentives conditional on performance, increases effort substantially. This is despite the fact that conditional monetary incentives are also more profitable for the firm – and perceived as less generous – than unconditional monetary incentives. The proportion of non-motivated workers who do more than three slogans increases by 11 percentage points if monetary incentives are made performance-based ( $p < 0.01$ ). Using conditional incentives *per se* does not generally backfire, i.e. workers are not generally lowering effort in conditional incentives relative to unconditional incentives – they do so only for charitable incentives.

Second, consistent with the mechanism proposed in our model, making incentives conditional on performance reflects negatively on workers’ perception of the firm’s extent of social responsibility, i.e. workers see firms as less socially responsible if incentives are conditional on effort compared to un-conditional.

Third, as predicted by our theory, the effect of performance-based, prosocial incentives depends on the workers’ prosociality. For motivated workers, for whom the two effects might cancel each other out, making prosocial incentives performance-based has only a marginally positive effect. However, for non-motivated workers who do not care about the charitable giving that much, the effect of making prosocial incentives performance-based is clearly negative. The proportion of workers that create more than three slogans decreases by 11 percentage points ( $p < 0.01$ ).

Fourth, prosocial incentives can even backfire compared to no additional incentives. The evidence comparing conditional and unconditional prosocial incentives is in line with the notion that workers use the nature of the incentives as a signal about the firm’s type. We are not making any sharp predictions about the difference between unconditional incentives and no additional incentives as it all depends on the perception of the firm’s intentions to offer the extra transfer. But, our experiment shows that even unconditional prosocial incentives can negatively affect the workers’ motivation and lead to lower effort compared to no incentive.



The paper makes contributions to at least two strands of literatures:

First, a growing literature investigates the effect of non-monetary and prosocial incentives on workers' productivity (for a review, see Cassar and Meier, 2016). The empirical papers (e.g. Tonin and Vlassopoulos, 2015, 2010; Imas, 2014; Charness et al., 2014; Cassar, 2017a; Fehrler and Kosfeld, 2014; Burbano, 2016; DellaVigna and Pope, 2016; Ashraf et al., 2014) mainly focus on comparing piece-rate monetary to piece-rate prosocial incentives and what its effect can tell about *motivated* agents' preferences.<sup>4</sup> For example, Imas (2014); Charness et al. (2014) show that prosocial incentive outperform monetary incentives when they are relatively low in levels but not when they are high. This result point to motivated agents having 'warm glow' preferences (Andreoni, 1989, 1990). Our paper shifts the focus to what the principal's choice of incentive signals to the agent about the principal's type. Just as agents use the choice of monetary incentives schemes as a signal about the principal or the attractiveness of the task (see, e.g. Bénabou and Tirole, 2003), they do so for prosocial incentives.

Our paper makes it clear that the firm's perceived motives matter for the effect of prosocial incentives. As a result, prosocial incentives are much more delicate than previously thought and firms cannot just exploit its effect instrumentally. This might explain why a number of studies find very small or no positive effect of prosocial incentives (Tonin and Vlassopoulos, 2010; Fehrler and Kosfeld, 2014; DellaVigna and Pope, 2016). While there is a substantial literature on the detrimental effects of monetary incentives (for a review, see Gneezy et al., 2011), our results show that non-monetary incentives might be as problematic (if not more delicate) if used the wrong way.<sup>5</sup>

Second, our paper contributes to the growing literature on the importance and functioning of CSR in motivating and selecting workers (for a review about the management literature on CSR as a HR tool, see, e.g., Flammer and Luo, 2017).<sup>6</sup> For example, in a recent field

---

<sup>4</sup>Tonin and Vlassopoulos (2015, 2010) have treatments with a lump-sum donation of £10 and pay-for-performance charitable incentive of 5-10p. While this is different from what we do in this paper, we discuss their results in the result part.

<sup>5</sup>This paper focuses on prosocial incentives, but firm's perceived intentions are expected to matter also for other non-monetary incentives, such as awards (Kosfeld and Neckermann, 2011; Neckermann et al., 2014; Bradler et al., 2016; Gallus, 2016; Frey and Gallus, 2017).

<sup>6</sup>There is a related literature about the effect of CSR on consumers (e.g. Elfenbein et al., 2012; Bartling et

experiment, Hedblom et al. (2016) show that jobs with a CSR component are able to attract more productive workers and generate higher quality work output (for other papers on the effect of having a social mission, see, e.g. Fehrler and Kosfeld, 2014; Kosfeld et al., 2014; Koppel and Regner, 2014; Carpenter and Gong, 2016; Gartenberg et al., 2016). Our paper contributes to this literature by arguing that the intention behind any CSR activity is very important – and not just the outcome as in how much good in the world they achieve. Our result point to a substantial limitation of CSR and what types of firms can effectively enjoy the benefits of it: firms cannot use CSR as another traditional HR tool, but have to think about it as a signal about their type. If CSR is not perceived as genuine and sincere, the benefits will disappear.

Our results indicate that the argument that firms should align their CSR strategy with their business interest (e.g. Porter and Kramer, 2007) or use CSR instrumentally to motivate workers (e.g. Bhattacharya et al., 2008), misses the negative impact this will have on their image. The benefits of CSR might disappear when used strategically. Think about the note in every hotel room saying that the hotel cares about the environment and therefore does not wash the towels if the hotel guest does not explicitly asks for it. Because we all know that this practice helps save costs for the hotel, it does not signal any prosocial attitudes. The results from a vignette study that we discuss at the end of the paper shows that if firms just do market research to find out whether CSR has a positive effect on their operation, the positive effect of CSR gets diminished as workers doubt the good intentions of the firm. A small literature points out – consistent with our results – that certain firms are reluctant to measure the impact of CSR or even publicize it initiatives (e.g. Carlos and Lewis, 2017) in order to avoid being seen as strategic. For firms that are not genuinely interested in socially responsibly behavior, instrumental CSR can be worse than ineffective – it can backfire.

The paper proceeds in the following: Section 2 presents our illustrative model and behavioral predictions. Section 3 then provides details about the experimental setting and design. Section 4 presents the main results. Section 5 discusses results from a vignette study that support our

---

al., 2015; Singh et al., 2016). There is also a literature that looks at the negative effect of using CSR to hide other problems, like “greenwashing” (see, e.g., Delmas and Burbano, 2011).

theory. And Section 6 concludes.

## 2. Illustrative model and behavioral predictions

We construct a simple model of equilibrium behavior to study the effect of charitable incentives on effort. While the model focuses on the effect of charitable incentives, it will also allow us to illustrate the differences between *charitable* incentives and *monetary* incentives. The simple model should show 1) that a separating equilibrium exists in which unconditional charitable incentives serve as a credible signal of an employer’s prosociality and 2) how in equilibrium performance-based charitable donations can decrease effort and thus backfire compared to unconditional charitable donations by signaling an employer’s low level of prosociality.

Consider the following environment, which closely follows the design of our experiment. A worker (he) is hired by an employer (she) to perform a task for an exogenously fixed wage  $w$ . At the core of our predictions is the assumption that workers’ willingness to put effort into the task depends on the employer’s type, namely on whether the employer is kind or unkind towards her workers. There are a number of microfoundations for this assumption. Reciprocity a la Levine (1998); Ellingsen and Johannesson (2008); Dur (2009) makes workers care about the employer’s utility depending on how kind the employer is towards them. Kajackaite and Sliwka (2017) provide evidence that indicate that reciprocity is a driver in agent’s positive reaction to charitable incentives in the lab. More broadly, consider a setting where the employer’s type is expected to affect the worker’s current and future utility from his employment contract. Many employment contracts are incomplete and the worker has to trust that the employer fulfills her promises in terms of wage, bonus, or other benefits. Additionally, if the employer is kind, the worker expects a higher wage in the future, hence, his value for the contract increases and he is less willing to shirk for fear of losing such a great job. On the contrary, if the employer is unkind, the worker may expect to be fired in the future and, therefore, he has no incentive of working hard now. In case of changing the job, a nice employer is also writing a more positive

recommendation letter. In the specific case of our M-Turk experiment, the employer’s type can affect the worker’s expectations of the employer treating them nice, by approving the task and in a timely fashion (i.e. actually paying the worker), paying an additional bonus and giving positive reviews. Nicer employers are also expected to be more honest about their promises in the job advertising about required time and effort to finish the task.

At the beginning of a new employment relationship, the employer’s type is non-observable by the worker. However, we assume that there is a positive correlation between an employer’s utility from charitable giving and how kind she is towards her workers. Similar correlations have been assumed and empirically validated in previous work. For instance, Burbano (2016) shows in her study on M-Turk that workers expect to be treated better by employers that offer CSR incentives. Elfenbein et al. (2012) assumes that there is a positive correlation between a seller’s utility from charitable giving and his disutility from behaving opportunistically towards consumers and this correlation is consistent with their empirical findings from E-bay data. Fehrler and Przepiorka (2013, 2016) shows that observable charitable giving by an individual is correlated with trustworthiness and increases significantly the trust of others into that individual.

In our model, we make the extreme assumption that there are two types of employers: prosocial employers, who care about charities and are kind to their workers, and selfish employers, who do not care about charities and are unkind to their workers. Therefore, the worker can form a belief  $\hat{q}$  about the probability  $q$  of an employer being prosocial—and thus being kind towards him—based on the type of charitable initiatives undertaken by the employer. In our experiment, these charitable initiatives take the form of a donation  $d$  made to a charity and we vary whether donation is a lump-sum transfer, namely unconditional on workers’ performance (i.e.,  $d = d_U = \bar{d}$ ), or whether the donation is performance-based, namely conditional on the workers exerting at least effort  $\bar{e}$  (i.e.,  $d = d_C = \bar{d}$  if  $e \geq \bar{e}$  and  $d = d_C = 0$  if  $e < \bar{e}$ ). Importantly, it is obvious that the goal of conditional donations is partly to increase the profit of the firm (i.e. “instrumental”) whereas for unconditional donations one cannot exclude that main goal is to benefit the charity (i.e. “non-instrumental”).

All workers care about the employer's type, because all workers care about being treated well by their employers. However, not every worker cares about the charitable donation. We assume that there are two types of workers, charity motivated (henceforth, "motivated") and non-motivated. Motivated workers care about the donation made to the charity, while non-motivated workers do not. Consistent with our experiment where workers could not self-select into different donation treatments, we abstract from selection and screening issues here. Hence, we assume that the employer chooses whether or not to condition the donation on the worker's performance after being matched with the worker and without knowing the worker's type. However, it is common knowledge that the proportion of motivated workers in the population is equal to  $p$ .

The employer's preferences can be represented by the following basic utility function:

$$U_P = \pi e + (\theta_P - 1)d - w \quad (1)$$

where  $\pi e$  is principal's revenue from agent's effort,  $e$ .  $\theta_P$  represents the employer's type. More specifically,  $\theta_P > 0$  if the employer is prosocial, i.e., if she derives some intrinsic pleasure from making the donation  $d$  to the charity, while  $\theta_P = 0$  if the employer is selfish. The financial cost of making the donation is given by  $-d$ . The worker's preferences are represented by the following utility function:

$$U_A = w + \lambda \hat{q}e - \frac{1}{2}e^2 + \theta_A d \quad (2)$$

where  $\lambda$  is the worker's marginal return of effort when he believes to be working for a prosocial employer, i.e., when  $\hat{q} = 1$ . If the worker believes that the employer is selfish – i.e.,  $\hat{q} = 0$  – he does not derive any intrinsic benefit from exerting effort, but only costs, which take the standard quadratic form  $\frac{1}{2}e^2$ .<sup>7</sup> The worker forms a belief  $\hat{q}$  about the probability  $q$  that  $\theta_P > 0$  based on the type of charitable incentive  $d \in \{d_U, d_C\}$  offered by the employer. That is, the worker has

---

<sup>7</sup>As mentioned earlier in this section there can be several interpretations and microfoundations for  $\lambda$ . Given the pure illustrative purpose of this model, we think that using this reduced form is enough to show the main theoretical insights.

beliefs of the form  $\hat{q}_{d_C} = E_A(q \mid d = d_C)$  and  $\hat{q}_{d_U} = E_A(q \mid d = d_U)$ . Finally,  $\theta_A$  captures the worker's motivation to generate the charitable donation. We assume that a motivated worker is always willing to exert effort  $\bar{e}$  in order to generate the donation under the performance-based charitable incentive, i.e.,  $\theta_A > \frac{\bar{e}}{2d}$ . Conversely, if the worker is non-motivated,  $\theta_A = 0$ .

**Proposition 1** *If  $\theta_P > \frac{\pi(p\bar{e}-\lambda)}{(1-p)\bar{d}} + 1 \equiv \bar{\theta}_P > 0$ , there exists a separating equilibrium where unconditional donations serve as a credible signal of the employer's prosociality, i.e., only if  $\theta_P > \bar{\theta}_P > 0$  employers make donations that are unconditional of workers' effort. The worker's equilibrium effort levels are then equal to  $e_{d_U}^m = \lambda$  and  $e_{d_C}^m = \bar{e}$  if the worker is motivated, and to  $e_{d_U}^{nm} = \lambda$  and  $e_{d_C}^{nm} = 0$  if the worker is not motivated. The worker's equilibrium belief on the employer's type is  $\hat{q}_{d_C} = E_A(q \mid d = d_C) = 0$  and  $\hat{q}_{d_U} = E_A(q \mid d = d_U) = 1$ .*

**Proof.** See Appendix A.1.

Proposition 1 provides some important insights. When choosing whether to condition the donation on effort, every employer faces the trade-off between two opposite effects: the negative effect of signaling to be a selfish type (in which case, workers' effort is reduced by  $-\lambda$ ), and the positive effect of motivating a motivated worker to exert extra effort in order to generate the donation ( $\bar{e}$ ) times the proportion of motivated agents in the population,  $p$ . However, compared to a selfish employer, a prosocial employer also faces the additional negative effect of renouncing the intrinsic value from making the donation should she be matched with a non-motivated worker. It is indeed this potential non-realized intrinsic value of the performance-based donation, which is captured by the term  $(1-p)\bar{d}$  and which is generated by the presence of non-motivated workers, that makes the unconditional donation relatively cheaper for prosocial employers and thus that makes it possible to obtain a separating equilibrium. Indeed, if all workers were motivated (i.e, if  $p = 1$ ), the trade-off between offering a conditional or unconditional donation would only boil down to compare the signaling versus the motivating effect—as the donation would happen with certainty in both cases—but this comparison is independent of  $\theta_P$  and, therefore, no separating contract would be sustainable in equilibrium. Hence, the presence of non-motivated workers is crucial in providing charitable incentives with a signaling

value.

The presence of non-motivated workers is also crucial in generating some clear effort predictions that can be tested in our experiment. As suggested by Proposition 1, it is indeed not clear a priori whether the conditional donation should increase or reduce the effort of a motivated worker compared to the unconditional donation. As can be seen by comparing  $e_{d_U}^m$  and  $e_{d_C}^m$ , it depends on the relative size of the signalling versus motivating effort (i.e, on whether  $\lambda$  is larger or smaller than  $\bar{e}$ ). On the contrary, conditioning the donation on effort unambiguously decreases the effort of a non-motivated worker. This is because for the non-motivated worker only the negative signalling of the conditional donation is at play. The main predictions of our experiment are therefore as follows:

**Prediction 1** *The effect on effort of performance-based donations relative to unconditional donations is higher for motivated workers than for non-motivated workers.*

And a stronger version of Prediction 1:

**Prediction 2** *Performance-based donations decrease the effort of non-motivated workers compared to unconditional donations.*

Next, in equilibrium, workers that are matched with employers who offer conditional donations should have different beliefs about the employer's prosociality than workers who are matched with employers who offer unconditional donations:

**Prediction 3** *Principals who offer performance-based donation are perceived as less prosocial than employers who offer unconditional donations.*

Prediction 3 should hold true for all types of workers, whether motivated or not motivated.

Finally, we want to compare the effects of these charitable incentives with a benchmark, such as standard private monetary incentives, that can be unconditional on effort, as a lump-sum bonus, or performance-based. While we can not exclude a priori that a similar signalling story

could also apply to monetary incentives and, therefore, that the choice between a conditional or unconditional bonus could work as a signal about the employer’s kindness toward her workers, we do not expect that the signalling value of the unconditional bonus can in any case outweigh the motivating effect of the conditional bonus. In fact, differently from the charitable incentives, all workers are “motivated by money”, there are no “not-motivated” workers. Furthermore workers’ financial motive is probably much larger in size than workers’ charitable motives because workers can always maximize their income and make private donations later. This leads us to the last prediction for behavior in our experiment:

**Prediction 4** *The effect on effort of performance-based incentives relative to unconditional incentives is higher when the incentives are monetary rather than charitable.*

### 3. Study Design

#### 3.1. Description of Intervention

In order to test our behavioral predictions, we designed an experiment in collaboration with an Italian company, named PharmaGIC S.R.L, which distributes pharmaceutical products in Italy and abroad. PharmaGIC is looking for suggestions on how to improve their English website. We partnered with them to recruit workers on Amazon Mechanical Turk (henceforth, M-Turk) with the task of generating marketing slogans for the products sold by PharmaGIC. These slogans are then considered for potential use in the English version of their website. All communication with the workers came from PharmaGIC.<sup>8</sup>

M-Turk is an online platform where companies and researchers can recruit cheap labor force to perform quick and easy tasks, called Human Intelligence Task (HIT). This platform is being increasingly used by economists (see, e.g., DellaVigna and Pope (2016); Ambuehl et al. (2015)). Goodman et al. (2013); Imas (2016); Imas et al. (2016) discuss that M-Turk participants behave

---

<sup>8</sup>We also got permission from IRB to waive the typical consent form.



very similarly to other samples in many (classical) decision-making tasks and differ little on many other dimensions. Interestingly for our study, M-Turkers apparently do not differ in their attitudes about money compared to a student sample, but they care more about money than a community sample (Goodman et al., 2013).

Workers were recruited on M-Turk and directed to a webpage that informed them about the company and the main task (see Appendix A.4 for all the instructions). We allowed only M-Turk workers who live in the US, have approval ratings of 98 or higher and have more than 100 HITs approved.

*Task and base payment:* The main task of workers was to create slogans for different products. We explained what a slogan or tagline is and that they have to create at least three slogans (but that we would appreciate more). It was clearly explained that the task is considered completed if they create at least three slogans and that they would receive a fixed wage of \$1.50.

*Treatments:* After reading the basic information about the firm and the task, workers were randomly assigned to different treatments which varied the type of incentives to motivate them to create more than three slogans. More specifically, we had a  $2 \times 2$  between-subjects design which varied:

1. *Nature of incentive:* We varied whether workers were offered a) a private monetary incentive that was a financial bonus of 75 cents, or b) a charitable incentives that was a donation of 75 cents made to the charity “Doctors without Borders”.
2. *Performance-based:* Additionally, we varied whether the incentive was i) conditional on performance, namely it was given if and only if the worker would create at least three more slogans (for a total of six slogans), or ii) unconditional on performance, namely it was given independently on whether the worker would create three more slogans. Importantly, to keep the reference-point of expected slogans fixed across treatments, also within the unconditional treatments we mentioned that we would appreciate if they could do at least three more slogans.

We also created two versions for each of those treatments that varied whether the workers were given an explanation for the choice of the incentives or not. In the conditional incentives treatments the explanation emphasized the instrumental nature of the conditional incentive by adding the following sentence: “Why do we give you (make) the bonus (donation)? The bonus (donation) is profitable for us: you doing at least three extra slogans for a wage (donation) of 75 cents is less costly for us than hiring another worker to do at least three slogans for a wage of \$1.50.”<sup>9</sup> In the unconditional incentives treatments the explanation emphasized the non-instrumental reason for offering an unconditional incentive: “Why do we give you (make) the bonus (donation)? We are strongly committed to be an employee-friendly company (socially responsible company—e.g. helping the larger community), even if this implies sacrificing some profit.” In the no-explanation treatments nothing was stated regarding the motive why the company offered a bonus or a donation. Finally, we had a baseline treatment with no additional incentives.

Table 1 summarizes the nine treatments that vary whether the treatment had a monetary or charitable incentive, whether the incentive was conditional on performance or not and whether an explanation for the choice of incentive was provided. In most of the analysis, we pool the treatments with and without explanation together. However, we do explore whether giving an explanation makes a difference.

After seeing the description of the task and the incentive scheme, workers started working on the slogans. The products were shown in sequence and only if workers finished one slogan, they were shown the next product. After they had created three slogans we thanked them for completing the task and, depending on the treatment, we summarized the payments (plus the bonus or donation) and then we asked everybody: “Would you like to do any more slogans? We would really appreciate if you could create at least three more slogans. Yes OR No” If they clicked ”yes” in the previous page, they were shown again the slogans in sequence. After each

---

<sup>9</sup>Given our behavioral prediction, it might seem unnatural (and unwise) for a firm to explain an action as profit-maximizing. Some firms not even publicize their efforts (e.g. Carlos and Lewis, 2017). However, such an open communication strategy could also be seen as very transparent and potentially minimizing suspicion (Johnson et al., 2016).

new slogan created, they were shown a screen with sentence: “Thank you for creating this extra slogan. Would you like to create one more?” Once they clicked “No” the experiment was over. The participants in the charitable incentives treatments received a link in which we uploaded the receipt of the donations generated during the intervention. We ended with a questionnaire in which we elicited their perceptions about PharmaGIC and few personal characteristics. The subjects only got paid after we approved the tasks. This feature made sure that workers actually entered real slogans related to the products. On average, workers used 81 seconds per slogan.

Table 1: Summary of Treatments

Treatments	Monetary	Charitable	Conditional	Explanation	Number of Observations	
					Treated	Completed
1	✓				335	317
2	✓			✓	335	306
3	✓		✓		334	299
4	✓		✓	✓	334	309
5		✓			334	301
6		✓		✓	333	307
7		✓	✓		334	287
8		✓	✓	✓	331	293
Baseline					334	300
Total Number of Observations					3,004	2,719

*Notes:* Table shows the feature of the nine treatments. The column “Treated” shows the number of workers assigned to the different treatments. Column “Completed” shows the number of workers who completed the task (including the questionnaire).

### 3.2. Sample size and Attrition

We pre-registered the design of the experiment on the AEA RCT Registry as AEARCTR-0001962 (“Response to variation in pro-social incentives”) where we also specified the rule for the sample size. We aimed at recruiting ideally 2,700 subjects and at least 1,800 subjects based on a power analysis which is part of the registrations.<sup>10</sup>

<sup>10</sup>In the registration it says: “We ran a small pilot to ensure that our protocol was working and that we were able to generate variation in the creation of slogans. The pilot also allowed us to get an estimate for what the standard deviation will be in our task. Based on 230 pilot participants, the standard deviation of created slogans was around 2.23. Assuming that this is approximately the standard deviation of each treatment in the

The study was launched on a Monday in January, 2017 and was open from 9am to 9pm EST each day. We specified on M-Turk that we want to collect 2,700 observations and we were able to collect 2,719 in less than two full days. Important for our study is the process of how workers are recruited and remain in the task: Workers see the advertisement for the job on M-Turk. When they ‘accept’ the job, the workers gets re-directed to a Qualtrics survey which explains the task and then randomizes workers into the different treatments so that there is the same number of observation per treatment. After they are informed about all the details (especially the incentive schemes) of the job, workers start working or stop doing it. We record who started seeing the instructions and who worked and completed the task.

Table 1 shows that 3,004 workers started the job and saw the treatments. Not surprisingly, Qualtrics is able to randomize workers equally into the nine treatments. 285 (9.49 %) of workers decide not to complete the task (they either did no slogan or stop after a couple of slogans without completing the survey or submitting a completion code on M-Turk) after seeing the instructions and the incentive schemes. While the attrition rate is relatively low, it is important to realize that there is differential attrition by treatment. Most striking (and consistent with our behavioral prediction) is the fact that attrition is higher in the two treatments with conditional charitable incentives, (12.78%), than in the treatment with unconditional charitable incentives (8.85%). The difference is significant in a Mann-Whitney test ( $Z = -2.312$ ,  $p < 0.05$ ). For monetary incentives, there is no significant difference between conditional, 8.98%, and unconditional incentives, 7.01%, ( $Z = -1.326$ ,  $p = 0.185$ ).

In the main analysis we include all the 3,004 workers that were treated, i.e. that saw the incentive scheme and then decided to put in either no effort or a positive effort. For some of the analysis, however, we do not have important information from workers that leave the job.

We discuss how this attrition can affect our results.

---

experiment and assuming a sample size of 1,800 (200 per treatment), there is thus an 80% power to reject the null hypothesis of zero difference in average slogans between two treatments when the actual difference between the two treatments is 0.64 slogans. Assuming instead a sample size of 2,700 (300 per treatment), there is then an 80% power to reject the null hypothesis of zero difference when the actual difference is 0.51 slogans. Based on our pilot, different treatments can create differences in average number of slogans by as much as 0.7-1.88 slogans, a difference of which can easily be detected statistically given the preceding calculations.”

### 3.3. Outcome variables and sub-group definition

We will focus on two main outcome variables: 1) The main outcome variable is the proportion of workers who created more than three slogans (the extensive margin). 2) We also analyze the number of created slogans (intensive margin). On average, 64.2% of workers create more than three slogans and the average number of slogan is 5.11 (s.d. of 2.89).

The brief survey at the end elicited workers’ perception of the firm, PharmaGIC (see Appendix A.4 for all the questions in the questionnaire). The main question asked workers to answer on a 10-point scale “Please rate PharmaGIC on a scale from 0 ‘Not at all socially responsible’ to 10 ‘Very socially responsible’”. We use this measure to test whether different incentives affect workers’ perception of PharmaGIC’s pro-sociality as predicted by our model. In addition, we also asked “How attractive would our company be as a potential employee?”, “How satisfied were you with the incentives that we provided for this task?”, and “How calculated do you think was our choice of incentives?”

Finally, we elicited worker’s pro-sociality towards charities, as a proxy for  $\theta_A$  in our model. The literature has shown that people typically contribute to the social good by making donations and/or by volunteering, and that both dimensions capture workers’ pro-sociality. In fact, some people will have a preference for donating money, other for volunteering, and others still will prefer a bit of both (see, e.g., Dur and Lent, 2016). Therefore, we asked subjects a) “How often do you donate money to a charitable organization?”, and b) “How often do you volunteer for a good cause?” To both questions, subjects could answer with “never”, “rarely”, “sometimes”, “often” and “regularly”. Table 2 shows the distribution of answers to these two questions.

We used the answers to these questions to divide workers into two categories: “non-motivated” (i.e., with sufficiently low  $\theta_A$ ) and “motivated” (i.e., with sufficiently high  $\theta_A$ ). In order to ensure an approximately equal number of people in each category, we categorized as motivated all subjects who either donate or volunteer “*often-regularly*” or who both donate and volunteer “*sometimes*” (bold in Table 2). This categorization leads to 1,390 motivated

Table 2: Workers’ pro-sociality

		Volunteer			Total
		<i>“never or rarely”</i>	<i>“sometimes”</i>	<i>“often or regularly”</i>	
Donation	<i>“never or rarely”</i>	706	251	<b>50</b>	1,007
	<i>“sometimes”</i>	372	<b>694</b>	<b>153</b>	1,219
	<i>“often or regularly”</i>	<b>84</b>	<b>196</b>	<b>213</b>	493
Total		1,162	1,141	416	2,719

*Notes:* The table shows the number of workers in terms of their volunteering ( *“How often do you volunteer for a good cause?”*) and donating ( *“How often do you donate money to a charitable organization?”*). The workers indicated in bold are labeled “motivated” workers and the rest are “non-motivated”.

types and to 1,329 non-motivated. The results are robust to another categorization of workers into “non-motivated” and “motivated”, which only categorizes workers who either donate or volunteer *“often-regularly”* as motivated, while the 694 workers who both donate and volunteer sometime were now categorized as non-motivated. In the analysis, we will also discuss how to think about how differential attrition of types by treatment can influence our results. As discussed before, in the conditional donation treatments, attrition is higher and one could expect that “non-motivated” types are more likely to decide not to complete the task. Table A.1 in the Appendix does show that indeed in the conditional donation treatments, the proportion of “motivated” workers is higher (i.e. more non-motivated workers did not complete the task). For all the other treatments, there is no significant difference.

## 4. Results

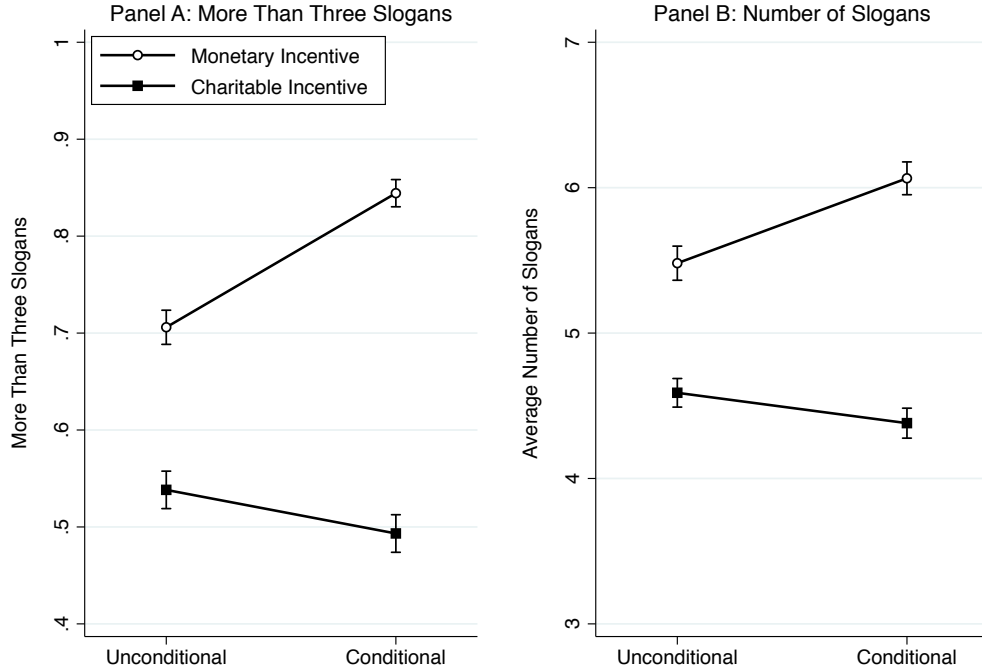
We present the results in three steps: first, we investigate the effect of performance-based versus unconditional incentives for both monetary and prosocial incentives. In doing so, the section explores a) the impact on effort, b) whether providing an explanation for the choice of incentive has a differential effect, and c) how the choice of incentive impacts the workers’ perception of the firm’s motive. Second, we explore whether non-motivated and motivated workers react

differently to the two types of incentives as predicted by our model. Third, we compare the two types of incentives to a situation in which we do not offer an incentive at all.

#### 4.1. Conditional vs. Unconditional (Charitable and Monetary) Incentives

Figure 1 shows the effect of performance-based (“Conditional”) and not performance-based (“Unconditional”) incentives (both monetary and prosocial incentives) on effort. Panel A shows the effect on the proportion of workers who do more than three slogans and Panel B shows the effect on the average number of slogans.

Figure 1: Difference between Prosocial and Monetary Incentives



*Notes:* The graph shows the effect of making monetary or charitable incentives performance-based (“Conditional”) or not performance-based (“Unconditional”). Proportion of workers who did more than three slogans (in Panel A) and average number of slogans (in Panel B). Bars shows standard errors of the mean.

Figure 1 clearly shows that making incentives conditional on performance has very different effects on effort depending on whether the incentive is private and monetary or charitable. As expected, making monetary incentives conditional on performance significantly increases effort: the proportion of workers who are willing to create more than three slogans increases

13 percentage points when monetary incentives are made conditional on effort. The difference is statistically significant in a Mann-Whitney test ( $Z = -6.06$ ,  $p < 0.01$ ).<sup>11</sup> Also, the average number of slogans increases around 11 percent when monetary incentives are conditional on effort compared to when monetary incentives are unconditional ( $Z = -6.67$ ,  $p < 0.01$ ).

The effect of making charitable incentives performance-based is quite different from the results for monetary incentives. Making charitable incentives performance-based *lowers* effort: The proportion of workers who are willing to do more than three slogans decreases from 54 percent to 49 percent ( $Z = 1.64$ ,  $p = 0.10$ ). It also has a slight negative effect on the number of slogans ( $Z = 1.46$ ,  $p = 0.14$ ).<sup>12</sup>

Table 3 confirms the results from Figure 1 in a regression framework. The regressions control for whether the incentive was charitable or monetary, whether the incentive was conditional on effort or not, and for the interaction of the two, i.e. whether the effect of ‘conditional’ is different for charitable incentives. The dependent variable is a dummy for whether the worker does more than three slogans in column (1) and the number of slogans in column (2). The results show: a) Charitable incentives result in lower effort compared to monetary incentives (based on the negative and significant coefficient of “Charitable Incentive”). This is consistent with Imas (2014); Charness et al. (2014) for high incentives. b) Making monetary incentives conditional on effort increases effort as seen in the positive and statistically significant coefficient of the dummy “Conditional”. c) Making charitable incentives conditional on effort has a lower effect on effort than when monetary incentives are made performance-based. In both models (in column (1) and (2)) the coefficient of the interaction is highly statistically significant ( $p < 0.01$ ). Hence, conditional incentives work much better for monetary incentives than for charitable incentives. If anything, the size of the effects indicate that conditional “charitable” incentives may actually backfire: in both models the coefficient of “Conditional” minus the coefficient of “Charitable×Conditional” is negative. For example, the results in column (1) shows that

---

<sup>11</sup>We will use Mann-Whitney tests to compare means – unless otherwise noted.

<sup>12</sup>Interestingly, Tonin and Vlassopoulos (2015) find in their between-subject exercise (Table 4, Columns 7 and 8) that a lump-some donation of £10 increases productivity more than a pay-for-performance charitable pay of 5-10p per task. This is very consistent with our results.



Table 3: Monetary and Pro-social Incentives

Dependent Variable:	(1) > 3 slogans	(2) # slogans	(3) > 3 slogans	(4) # slogans	(5) Soc. resp.
Charitable Incentive (=1)	−0.168*** (0.026)	−0.891*** (0.153)	−0.198*** (0.037)	−1.154*** (0.218)	0.405*** (0.102)
Conditional Incentive (=1)	0.138*** (0.023)	0.584*** (0.163)	0.107*** (0.032)	0.331 (0.239)	−0.343*** (0.105)
Charitable Incentive×Conditional	−0.183*** (0.035)	−0.793*** (0.216)	−0.131*** (0.050)	−0.469 (0.311)	−0.003 (0.149)
Explanation (=1)			−0.039 (0.035)	−0.364 (0.234)	
Explanation×Charitable Incentive			0.061 (0.052)	0.525* (0.306)	
Explanation×Conditional			0.063 (0.045)	0.505 (0.325)	
Explanation×Charitable×Conditional			−0.105 (0.071)	−0.647 (0.433)	
Constant	0.706*** (0.018)	5.481*** (0.117)	0.725*** (0.024)	5.663*** (0.169)	7.239*** (0.072)
<i>F</i> -test: “Cond.”+“Charit.×Cond.”=0	<i>p</i> =0.10	<i>p</i> =0.14	<i>p</i> =0.54	<i>p</i> =0.49	<i>p</i> <0.00
Adj. <i>R</i> <sup>2</sup>	0.084	0.055	0.083	0.055	0.020
Observations	2,670	2,670	2,670	2,670	2,418

*Notes:* OLS regressions with robust standard errors. In regressions (1) and (3) the dependent variable is a dummy variable equal 1 if the worker created more than 3 slogans, whereas in regressions (2) and (3) the dependent variable is the number of created slogans. Column (5) has as dependent variable the answers to the question: ‘Please rate PharmaGIC on a scale from 0 “Not at all socially responsible” to 10 “Very socially responsible”’. The dummy variable “charitable incentives” takes value 1 for the treatments with charitable incentives and 0 for the treatments with monetary incentives; the dummy variable “conditional” takes value 1 for the treatments with conditional incentives and 0 for the treatments with unconditional incentives; the dummy variable “explanation” takes value 1 for the treatments with explanation and 0 for the treatments without explanation. Significance levels: \*\*\* *p*<.01, \*\* *p*<.05, \* *p*<.1.

workers are 4 percentage points less likely to generate more than three slogans when offered a conditional than an unconditional prosocial incentive. However, in both models the joint effect is at most statistically significant on the 10%-level (F-test:  $p = 0.10$  in Column (1) and  $p = 0.14$  in Column (2)).<sup>13</sup>

The results support Prediction 4 in Section 2: While performance-based monetary incentives are very powerful to increase worker’s effort, performance-based prosocial incentives are not effective and may even backfire. The negative effect is specific to charitable incentives and cannot be explained by workers generally reacting negatively to conditional incentives.

In the following, we investigate whether the above result depends on whether the firm provides an explanation for the choice of incentive. In particular, we are interested whether explicitly explaining that offering a conditional incentive is in the best interest of the firm, has a stronger (negative) effect on effort. Remember that in the conditional treatments, we mention to the worker that it is cheaper for the firm to offer \$0.75 in bonus or charitable giving for three more slogans than hiring a new worker to do three more slogans for \$1.5. Directionally the effect of explaining the rationale behind offering conditional charitable incentives does lower effort more than without explanation (see Figures A.1 in the Appendix for a graphical representation). For prosocial incentives, the proportion of workers decrease by 7% when giving an explanation ( $Z = 1.71$ ,  $p = 0.09$ ) and decreases by 3% ( $Z = 0.62$ ,  $p = 0.54$ ) without giving an explanation. The difference between the difference is, however, not statistically significant. Regression results presented in Column (3) and (4) of Table 3 support that providing an explanation, “Explanation (=1)” has a limited and insignificant effect on effort.

The prediction presented in Section 2 on why conditional monetary incentives is more effective than conditional charitable incentives is based on the hypothesis that workers interpret a conditional incentive (both monetary and prosocial) as a signal that the employer is less altruistic than a firm that offers unconditional incentives. The questionnaire asked workers “Please rate PharmaGIC on a scale from 0 “Not at all socially responsible” to 10 “Very socially

---

<sup>13</sup>We find the same results in terms of the sign and significance of the coefficients if we use a logit regression for the models with a dummy for whether the worker has done more than three slogans as the dependent variable (see Table A.2 in the Appendix).

responsible”. The results are consistent with our hypothesis: the firm is indeed perceived as significantly less socially responsible under the conditional treatments than under the unconditional treatments and this holds true for both monetary and charitable incentives ( $p < 0.01$  for both cases).<sup>14</sup> Column (5) in Table 3 show a model with the answers to the question about the social responsibility of the firm as dependent variable. The results confirm that when the firm offers conditional incentives, the worker perceives the company as less socially responsible ( $p < 0.01$ ).<sup>15</sup> We explore the perception as a channel more explicitly below.

In sum, the result support our behavioral prediction in that performance-based incentives work much better when the incentive is a private bonus than when it is a prosocial incentives – in the form of a charitable incentive. The results also support our hypothesis that performance-based incentives have a negative effect on perceived employer’s pro-sociality compared to unconditional incentives. Hence, consistent with our behavioral predictions, it seems that performance-based incentives have two countervailing effects: on the one hand, they motivate workers to exert higher effort in order to earn more money as a bonus or to increase the donation for the charity. On the other hand, they discourage effort by negatively impacting how workers perceive the employer’s motives. While the incentive effect clearly dominates for monetary incentives, for pro-social incentives it seems that both effects counterbalance on aggregate. Directionally, making prosocial incentives performance-based even backfires on average. However, our theory predicts that there should be heterogeneous effects of performance-based pro-social incentives depending on the pro-sociality of the workers. We investigate this question in the next section.

## 4.2. Non-motivated vs. Motivated Workers

Our model makes clear predictions in how performance-based incentives should affect non-motivated and motivated workers. Workers who care about charities (who we label ‘motivated’)

---

<sup>14</sup>Workers, on average, gave the firm a 7.64 (on a 10-point scale) for unconditional charitable incentive, 7.30 for conditional charitable, 7.24 for unconditional monetary incentives, and 6.90 for conditional monetary.

<sup>15</sup>Table A.3 show the results for the other relevant survey questions. Throughout, the results show that directionally workers perceive conditional incentives as less attractive, less satisfying and more calculated.

face the two countervailing effects of a performance-based prosocial incentives: performance-based prosocial incentives make them work harder because they want to give to charity. However, also for them the conditionality of the incentive negatively affects their view of the firm. The net effect will be unclear. However, the incentive effect should be largely absent for workers who do not care about charities (who we label ‘non-motivated’). Hence, according to our theory, for non-motivated workers conditional charitable incentives should unambiguously reduce effort compared to unconditional charitable incentives.

Figure 2 shows the effect of charitable and monetary incentives on the proportion of non-motivated (Panel A) and motivated (Panel B) workers who create more than three slogans. The Figure shows that consistent with our model prediction, the effect of conditional compared to unconditional charitable incentives largely depends on the pro-sociality of the workers.<sup>16</sup> Non-motivated workers are 11 percentage points *less* likely to write more than three slogans under conditional than unconditional pro-social incentives ( $Z = 2.62$ ,  $p < 0.01$ )<sup>17</sup>, whereas motivated workers are 3 percentage points *more* likely to write more than three slogans under the conditional donation than under the unconditional donation treatment but the difference is not significant ( $Z = -0.79$ ,  $p = 0.43$ ). The effect is also reflected in the average number of slogans: Motivated workers create on average 2 percent more slogans under conditional than unconditional charitable incentives, although not significantly so ( $Z = -1.02$ ,  $p = 0.31$ ). On the contrary, non-motivated workers create on average 6 percent fewer slogans under conditional than unconditional charitable incentives. This difference is statistically significant ( $Z = 2.18$ ,  $p < 0.05$ ).

Interestingly, the difference between conditional and unconditional monetary incentives depends also on the worker’s pro-sociality. For non-motivated workers, conditional monetary incentives are more effective in increasing the proportion of workers creating more than three

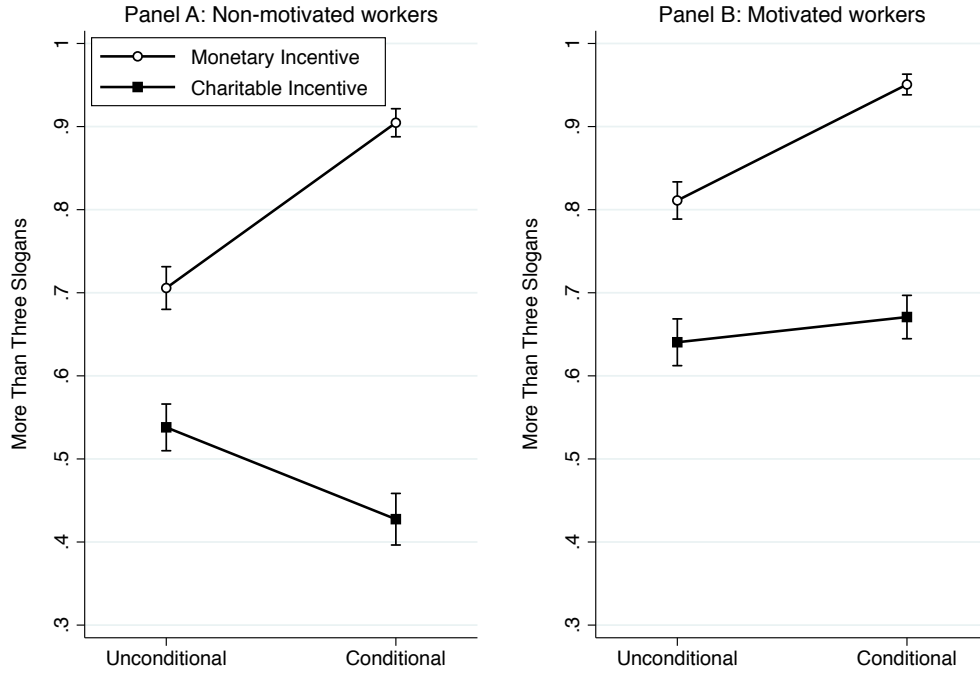
---

<sup>16</sup>The Figure also shows that more motivated workers (64%) are creating more than three slogans than non-motivated agents (54%) ( $Z = -2.56$ ,  $p = 0.01$ ) in the unconditional charitable donation treatment. This is not surprising and just indicates that our classification of motivated and non-motivated agents is associated (as it should) with the behavior in the experiment.

<sup>17</sup>Even when analyzing just treatments without explanation, the proportion of non-motivated agents doing more than three slogans decreases by 7 percentage points ( $p < 0.05$ ).

slogans compared to unconditional monetary incentives than for motivated workers. This seems to be due to the fact that motivated workers create more slogans under the unconditional bonus treatment, presumably because they also care more about the employer’s utility (they may have a higher  $\lambda$ .) Therefore the scope for improvement of conditional monetary incentives is more limited for motivated workers.

Figure 2: Non-motivated vs. Motivated Workers



*Notes:* The graph shows the effect of making monetary or charitable incentives performance-based (“Conditional”) or not performance-based (“Unconditional”) on the proportion of workers who created more than three slogans. The panels show the effect for non-motivated workers in Panel A and motivated workers in Panel B. Bars show standard errors of the mean.

Table 4 show the results for the two subsamples in a regression framework. Columns (1) and (3) focus on non-motivated workers while columns (2) and (4) focus on the motivated workers. As can be seen, for the non-motivated workers the interaction between the conditional and the charitable incentive treatment in each regression is negative and highly significant ( $p < 0.01$  in both regressions). Furthermore, an F-test shows that the proportion of workers who create more than three slogans is significantly lower under conditional than unconditional charitable incentives—the coefficient of the interaction term is significantly higher in absolute terms than the coefficient on the conditional variable—revealing an overall significant negative effect (F-test:

$p < 0.01$ ).

The results in Column (2) and (4) show that the effect is different for motivated workers. The interaction term “Charitable Incentives×Conditional” is negative but only significant in regression (2) ( $p = 0.02$ ). F-test additionally show that the probability and the number of slogans are the same for conditional and unconditional charitable incentives (F-test:  $p = 0.43$  and  $0.60$  respectively). Column (5) and (6) show the difference between non-motivated and motivated workers in a single model for each outcome variable. The important coefficient on “Motivated×Charitable Incentive×Conditional” is highly significant – indicating that motivated workers react quite differently to the conditionality of prosocial incentives than non-motivated workers.

The analysis in this section is based on workers who answered the survey questions at the end of the task (as only for those we know whether they are “motivated” or not). As we discussed in Section 3.2, about 9.5% of workers did not complete the task. We do not have information about whether those workers are motivated or non-motivated, but we do know that they put in very little effort (if at all). While attrition is relatively low, the analysis could be slightly biased because of differential attrition by treatment. Table A.1 in the Appendix shows that differential attrition lead to different distribution of motivated and non-motivated workers in the different treatments. It looks like that more non-motivated agents decided not to work after being informed about the conditional prosocial incentives (there is no significant difference between all the other treatments). Therefore, if we assume that all the workers who decided not to work are non-motivated, the effects that we report are biased downwards. However, even if we assume that all of them are motivated, our results still hold that non-motivated workers react negatively to conditional prosocial incentives (see Table A.4 in the Appendix).

As a final piece, we can investigate the influence of the perception of the firm’s intention on the backfiring of ‘instrumental’ prosocial behavior. According to our behavioral prediction, effort decreases for conditional prosocial incentives compared to unconditional incentives because workers perceive the firm as less prosocial. In particular for non-motivated agents, our theory predicts that (conditional) prosocial incentives affect effort only through their perception of

the firm’s intentions. For those agents, we can use the treatment as an instrument for their perception and then estimate the effect on effort (for a similar approach, see Haggag and Pope, 2016). For motivated agents, we cannot use the treatment as an instrument as effort is directly affected by conditional prosocial incentives. Those agents work harder because this will increase the donation to a cause they support. Therefore, the exclusion restriction is only met when  $\theta_A$  is close to zero. The result (see Appendix Table A.5) confirms our result that making prosocial incentives conditional on performance reduces the perception of firm’s prosociality. Using the instrumented perception variable indicates that an 1-point change in perception reduces the probability of doing more than three slogans by 0.274 points ( $p < 0.05$ ) and the number of slogans by 0.746 ( $p = 0.11$ ).

In sum, our results support the theoretical prediction that performance-based prosocial incentives have heterogeneous effects on effort depending on the pro-sociality of the worker. For motivated workers we observe a net effect of two countervailing forces: an incentive effect that increases workers motivation to work more (and donate more as a result) and the negative effect on effort because performance-based prosocial incentives reflects negatively on the firm’s motive. In our setting, the two effects seem to cancel each other out. For non-motivated workers, however, the positive incentive effect is largely absent. The negative effect of making prosocial incentives performance-based dominates. As a result, performance-based prosocial incentives backfire. Using an instrumental variable approach, we also show that for non-motivated agents the channel of backfiring comes from the change in the perception of the firm’s prosociality.

### 4.3. Comparison to Baseline

The results so far show that making charitable incentives performance-based can have negative effects on motivation and effort. According to our theory and consistent with the survey evidence and subgroup analysis, this is because performance-based incentives negatively affect the perceived employer’s pro-sociality. Workers use the choice of incentives as a signal about the intention and type of the firm.

The theory and the empirical test so far compared conditional with unconditional incen-

Table 4: Motivated versus Non-motivated Workers

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variables:	> 3 slogans	> 3 slogans	# slogans	# slogans	> 3 slogans	# slogans
Subsample:	Non-motivated	Motivated	Non-motivated	Motivated	All	All
Charitable Incentive (=1)	-0.168*** (0.038)	-0.171*** (0.036)	-0.611*** (0.165)	-1.100*** (0.234)	-0.168*** (0.038)	-0.611*** (0.165)
Conditional Incentive (=1)	0.199*** (0.031)	0.140*** (0.026)	1.073*** (0.184)	0.474** (0.227)	0.199*** (0.031)	1.073*** (0.184)
Charitable Incentive×Conditional	-0.309*** (0.052)	-0.109** (0.046)	-1.368*** (0.248)	-0.372 (0.301)	-0.309*** (0.052)	-1.368*** (0.248)
Motivated Type (=1)					0.105*** (0.034)	1.049*** (0.221)
Motivated×Charitable					-0.003 (0.052)	-0.489* (0.287)
Motivated×Conditional					-0.059 (0.040)	-0.599** (0.292)
Motivated×Charit.×Cond.					0.200*** (0.069)	0.996** (0.390)
Constant	0.706*** (0.026)	0.811*** (0.022)	5.345*** (0.124)	6.394*** (0.183)	0.706*** (0.026)	5.345*** (0.124)
<i>F</i> -test:						
“Cond.” + “Charit.×Cond.”=0	$p < 0.01$	$p = 0.43$	$p = 0.08$	$p = 0.60$		
Adj. $R^2$	0.136	0.082	0.107	0.058	0.126	0.096
Observations	1,191	1,228	1,191	1,228	2,419	2,419

*Notes:* OLS regressions with robust standard errors. In regressions (1), (2) and (5) the dependent variable is the number of created slogans, whereas in regressions (3), (4) and (6) the dependent variable is a dummy variable equal 1 if the worker created more than 3 slogans. The dummy variable “donation” takes value 1 for the treatments with charitable incentives and 0 for the treatments with monetary incentives; the dummy variable “conditional” takes value 1 for the treatments with conditional incentives and 0 for the treatments with unconditional incentives; the dummy variable “motivated” takes value 1 if the worker was categorized as pro-social depending on the frequency with which he donates or volunteers. Significance levels: \*\*\*  $p < .01$ , \*\*  $p < .05$ , \*  $p < .1$ .

tives – both monetary and prosocial. A natural question is how workers interpret unconditional incentives relative to a scheme with no additional incentives. Note that we did not make any specific prediction on how pro-social incentives would perform compared to a baseline with no incentives. Importantly, even unconditional charitable incentives could, in principle, backfire as it depends on the workers’ perception of the firm type when the latter offers unconditional charitable incentives compared to when it does not offer any incentive. If the use of unconditional charitable incentives is perceived as instrumental to increase effort, it could well be that workers reduce rather than increase their effort compared to the baseline.

We implemented a treatment in which we did not offer any additional incentive for creating three more slogans. Table 5 compares effort in this treatment to the four other treatments



(pooling treatments with and without explanation).<sup>18</sup> The results show a number of interesting patterns: First, monetary incentives do significantly better than no incentives. This is not surprising for conditional monetary incentives, but the positive effect of unconditional monetary incentives indicate that workers reciprocate the *unconditional* incentive with higher effort. The proportion of workers doing more than three slogans increases by 9 percentage points when offering an unconditional bonus.<sup>19</sup>

The second finding is, however, more important for our study as it shows that both unconditional and conditional pro-social incentives *reduce* effort in terms of proportion of workers who create more than three slogans and the number of slogans. While the effect is smaller for unconditional charitable incentives than for conditional (and estimated with less precision), even unconditional incentives backfire. This result may seem in contrast to some papers that show that prosocial incentives increase effort (see literature in introduction). However, there are a number of papers that found small (e.g., Tonin and Vlassopoulos, 2010; DellaVigna and Pope, 2016) or no effects (e.g., Fehrler and Kosfeld, 2014) of prosocial incentives.

There are a number of potential explanations for this result, all of which can be seen as reinforcing our main argument that prosocial incentives are used as a signal for the intention of the firm. First, note that in our experiment, the baseline payment of \$1.50 for *three slogans* was already very generous relative to the payment that M-Turk workers receive in general. As a comparison, DellaVigna and Pope (2016) paid a baseline payment of \$1 for a tedious 10-minute task and state “This pay is quite generous given that average pay on MTurk is \$1.40 per hour according to Horton and Chilton (2010)” (p. 13). Including the bonus or prosocial incentives, our workers got, on average, \$2.1 (median: \$2.25) and it took them, on average, 11.2

---

<sup>18</sup>Table A.6 shows comparisons between all the treatments – broken up in treatments with and without explanations.

<sup>19</sup>The result that unconditional monetary bonus increases effort is interesting for the debate about gift-exchange (Akerlof, 1982; Fehr et al., 1993). While the evidence from laboratory studies show quite strong reciprocity to monetary gifts, the evidence from the field is rather mixed (e.g., Gneezy and List, 2006; Hennig-Schmidt et al., 2010). Our result could be interpreted as gift exchange in the field. However, we want to be careful in this interpretation as workers might have had reemployment concerns (see the nice discussion of multiple confounds in gift exchange experiments in Esteves-Sorenson, 2017). While the firm never mentioned any reemployment (and in fact never thought about hiring more workers), the experiment which was not focused on gift exchange also never explicitly excluded it.

minutes for the task including survey (median: 9.3 minutes). Furthermore, a number of workers commented on how fun and creative the task was and how satisfied they were with the payment. This very generous baseline could have affected the workers' perception of the firm as being of the good type. Adding a charitable incentive in such a situation, even if unconditional, might have been perceived as instrumental. Why would the firm otherwise mention its donations to the workers if not to generate a reaction? Note, indeed, that not being offered a prosocial incentive in the baseline treatment does not imply that the firm is not making any donation. The workers might well be thinking that the firm is donating but not revealing it in this context. In fact, in an extension of our model (Section A.2 in the Appendix) that allows for multiples types and for the possibility of not revealing one's donations, we argue that there exists an equilibrium in which the most prosocial employers do not reveal their charitable donations in order to (counter)signal a higher prosociality compared to the medium prosocial employers.

Second, it could also be that in the baseline treatment the workers are just not thinking about the firm's potential prosocial initiatives, while in the charitable incentive treatments, these initiatives are very salient and the participants might have even preferred to receive the charitable incentive as a private bonus and, therefore, reacted negatively to its introduction. This negative reaction cannot explain the difference between unconditional and conditional charitable incentives explored in the main part of the paper as the presence of charitable incentives is kept constant, but it could explain the difference between baseline and charitable incentive treatments. While we did not design our study to focus on the comparison between incentives – monetary or charitable – and not offering an additional incentive, our results are important on that dimension in that they indicate that one has to be careful in offering prosocial incentives. The perceived intentions of why the prosocial incentives are offered – and revealed – will be important for their success.

Table 5: Comparison to Baseline

	(1)	(2)
Dependent Variables:	> 3 slogans	# slogans
Constant (Baseline)	0.614*** (0.027)	4.934*** (0.166)
Unconditional Monetary	0.092*** (0.032)	0.546*** (0.203)
Conditional Monetary	0.231*** (0.030)	1.130*** (0.200)
Unconditional Charitable	-0.076** (0.033)	-0.345* (0.193)
Conditional Charitable	-0.121*** (0.033)	-0.554*** (0.195)
Adj. $R^2$	0.074	0.048
Observations	3,004	3,004

*Notes:* OLS regressions with robust standard errors. Dependent variables: dummy variable equal 1 if the worker created more than 3 slogans and 0 otherwise in (1), and the number of created slogans in (2). Significance levels: \*\*\*  $p < .01$ , \*\*  $p < .05$ , \*  $p < .1$ .

## 5. Instrumental CSR

Our results show that workers’ perception of an employer’s pro-sociality matters for their effort and this perception is affected by the type of incentives that the employer decides to offer: compared to unconditional incentives, performance-based incentives send a stronger signal that the employer’s intention is to extract effort from the worker. They reduce the workers’ perception about their employer’s pro-sociality. We then show that this effect has important implications for the use of pro-social incentives: if workers are not particularly motivated to generate a donation, the use of performance-based pro-social incentives is likely to backfire against the employer.

While this paper focused on charitable incentives as one particular aspect of CSR initiatives, the same argument could be applied to any type of investment in CSR: if workers perceive CSR to be done strategically to increase profit (e.g. by improving the company’s image) it may not be as effective or may even backfire. If this is the case, the insights learned from our model and our empirical findings are much broader and extend beyond the use of pro-social incentives in the sense of performance-based charitable giving. As a first step, we explore the effect of instrumental CSR using a short survey study. We test whether just doing a market research study about the potential financial benefit of a CSR program (i.e. checking whether the program is profitable before committing to it) is enough to affect the perception of potential workers of the firm’s intentions – keeping the social impact of the CSR initiatives the same.

We conducted a vignette study on M-Turk with 300 subjects (for all the details of the study and the instructions, see A.5 in the Appendix). Subjects were presented with a scenario in which a firm decided to make an annual donation to a charity as part of their CSR program. The 3×2 design varied: 1) Three potential benefits to the firm in terms of a) to improve the firm’s image and brand; b) to increase workers’ effort; c) to attract workers for a lower wage. 2) For each possible benefit, we then varied whether the investment in CSR was made strategically, i.e. *“only after making a market analysis whether the donation increases the firm’s profit because of [one of the three benefits]”* or non-strategically (unconditionally), i.e. *“without*

*making a market analysis whether the donation increases the firm’s profit because of [one of the three benefits]”*<sup>20</sup> Our theory would predict that for each possible benefit, the worker’s perception of the firm’s image would be better, their hypothetical effort would be higher and their hypothetical probability of accepting a lower wage would be higher if the CSR invest was made unconditionally rather than strategically. Subjects revealed their view of the firm and their motivation to work for the firm on multiple dimensions using 10-point scales.

Table 6 shows the effect of strategic vs. non-strategic CSR on multiple dimensions of the worker’s perception or motivation.<sup>21</sup> The results support our predictions that when CSR was made strategically: a) the firm was rated less attractive; b) the donation was perceived as less generous; c) the firm was perceived as less socially responsible ; c) the donation was perceived as less effective in motivating effort and d) workers would be less likely to accept a lower wage, than when CSR was made unconditionally. Hence, the results of the vignette study suggest that using CSR instrumentally to increase profits might destroy the very benefits it was hoping to achieve.

Table 6: Strategic and Non-strategic CSR

Dimensions:	Strategic	Non-strategic	Difference
Attractiveness	5.74	7.99	$p < 0.001$
Generous donation	5.07	8.53	$p < 0.001$
Socially responsible	6.08	9.05	$p < 0.001$
Acceptance lower wage	0.14	0.22	$p < 0.06$
Motivation	5.95	7.33	$p < 0.001$

*Notes:* Table shows average answers for multiple questions/dimensions on subjects’ perception of the firm and their motivation to work for the firm. Subjects used 10-point scales. Columns show between-subject differences in whether the firm was ‘strategic’ or ‘non-strategic’ about their CSR.  $p$ -value of Mann-Whitney test of difference. N=300. For all the details about the study, see A.5 in the Appendix.

<sup>20</sup>For more details see the full version of the six scenario in the Appendix. The survey also implemented a within-subject design by asking workers about the another scenario. In the analysis reported here, we only analyze the first scenario that the subjects saw, i.e. effectively looking at between-subject differences.

<sup>21</sup>Table 6 shows the results pooling across the three potential benefits. Table A.7 in the Appendix shows the survey responses for all six treatments.

## 6. Conclusions

This paper investigates the limits of prosocial incentives by providing a mechanism through which prosocial incentives can backfire. The important ingredient of our argument is that agents use the choice of incentives as a signal about the firm’s type and its prosociality. When incentives (monetary and prosocial) are used instrumentally, workers adjust their perception of the firm. The net effect is then determined by a positive incentive effect and a negative effect due to the reduced positive image of the firm. This net effect is more likely to be negative for prosocial incentives because non-motivated agents only care about the signaling value of the prosocial incentive and are not incentivized by the effect their effort has for the mission goal.

The results of an experiment with around 3,000 workers support the theory that monetary and prosocial incentives work very differently. While monetary incentive that are profitable for the firm by making them performance-based has a positive effect on effort, making prosocial incentives instrumental, i.e. profitable for the firm, backfires. We can show that especially for non-motivated workers there is a substantial backlash to using prosocial incentives instrumentally.

This paper has important implications for research and practice, and points to interesting open questions and next steps.

The paper shows that prosocial incentives have substantial limitations and might be potentially more delicate than monetary incentives. More research needs to investigate the limits of prosocial incentives. Our paper makes the important point that future research has to take seriously the signaling value of prosocial incentives. Workers use the choice of incentives as a signal about firms – as they should because the firm explicitly chooses one incentive scheme over another. While a number of theories and empirical approaches already focus on the signaling value of *monetary* incentives (e.g., Bénabou and Tirole, 2003; Sliwka, 2007; Ellingsen and Johannesson, 2008), there should be more research investigating how *prosocial* incentives and CSR activities are perceived by the agents. As we show in this paper, non-monetary incentives have different dynamics from monetary incentives that should be further explored.

The effect of prosocial incentives will depend on the pool of workers in a firm. Past research stressed rightly so that motivated workers are crucial for the working of prosocial incentives and select into firms with strong CSR initiatives (see, e.g., Nyborg, 2014). Past research, however, disregarded the non-motivated agents as they expected no effect of prosocial incentives for those workers. We show that focusing only on motivated workers for the working of prosocial incentives is missing that non-motivated workers (workers with a low  $\theta_A$ ) are as important to study as they care about the signaling value of prosocial incentives. Our model and experiment did not allow for sorting. Selection by motivated and non-motivated workers into different firms is, however, an important issue that should be explored further.

We also argue that the more workers care about the type of firm, i.e. they have a higher  $\lambda$ , instrumental prosocial incentives are more detrimental. Our paper cannot test this predictions, but future studies should investigate whether the effect is higher in settings or industry with more high  $\lambda$  workers.

Our paper argues that the perceived intention of the firm is important for the working of non-monetary incentives. We used one particular form of prosocial incentives and one particular way (i.e. using ‘performance-pay’) to make the incentive be perceived as instrumental, i.e. in the best interest of the firm. There are, of course, other ways in which (prosocial) incentives are instrumental (e.g. by implementing them only after a market research study confirms a positive net effect on profits as in our vignette study). The main identification of how instrumental prosocial incentives could backfire comes from comparing prosocial incentives to monetary incentives but we also show that even unconditional prosocial incentives do worse than no incentive at all. Future research should investigate how different types of prosocial incentives are interpreted by workers. In our setting, unconditional charitable incentives might have backfired compared to no incentives because workers in the donation treatment perceived the incentive scheme as instrumental and would have rather got the bonus. And in the baseline treatment they not even thought about getting a higher incentive as our pay was already very generous. While our theory and test was not designed to compare prosocial incentives (instrumental or not) and no incentives, future work should investigate the conditions under

which any prosocial incentive might backfire.

Workers in our setting were assigned to one incentive scheme and had to infer the intention/type of the employer from this particular scheme. We decided to not make the intention of the firm more salient - except in the explanation treatment and the vignette study in which we indicated that the firm decided actively for or against market research to evaluate the benefit of a CSR campaign on the firm's profit. It is possible (and should be tested further) that the effect we document in this paper would be more extreme if the workers were informed that the firm could have offered the incentive unconditionally but decided to make the pay performance-based.

While the experiment focuses on agent's effort, the research also has important implications for firms. Our result indicate that firms cannot just instrumentally use non-monetary incentives or CSR initiatives as another tool in their HR strategy.<sup>22</sup> Workers do care about the genuine use of prosocial incentives and will react negatively to firms using social initiatives in a calculated manner to increase profits. This signaling problem makes it more difficult for firms to profit from CSR. Walmart, for example, gets probably little benefit in terms of worker productivity or even public image by having the largest hybrid truck fleet because this operational choice not only helps the environment but is seen as being cost-effective for Walmart. Our research indicates that prosocial incentives that are profitable for the firm might backfire. Future research should investigate further whether firms need to choose costly, i.e. unprofitable, CSR activities in order to get the most out of it and which firms ultimately will undertake CSR activities and which firms will not.

---

<sup>22</sup>In this paper, we focus on the effect of instrumental prosocial incentives on workers. We would expect – but need future research to confirm – that instrumental CSR also negatively affects consumers' willingness-to-pay.



## References

- Akerlof, George A**, “Labor contracts as partial gift exchange,” *The quarterly journal of economics*, 1982, 97 (4), 543–569.
- Ambuehl, Sandro, Muriel Niederle, and Alvin E Roth**, “More Money, More Problems? Can High Pay be Coercive and Repugnant?,” *The American Economic Review*, 2015, 105 (5), 357–360.
- Andreoni, James**, “Giving with Impure Altruism: Applications to Charity and Ricardian Equivalence,” *Journal of Political Economy*, 1989, 97 (6), 1447–58.
- , “Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving,” *The Economic Journal*, 1990, 100 (401), pp. 464–477.
- Armouti-Hansen, Jesper, Lea Cassar, and Anna Dereky**, “Efficiency wages with motivated agents,” *Working Paper*, 2017.
- Ashraf, Nava, Oriana Bandiera, and B Kelsey Jack**, “No margin, no mission? A field experiment on incentives for public service delivery,” *Journal of Public Economics*, 2014, 120, 1–17.
- Bartling, Björn, Roberto A Weber, and Lan Yao**, “Do markets erode social responsibility?,” *The Quarterly Journal of Economics*, 2015, 130 (1), 219–266.
- Bénabou, Roland and Jean Tirole**, “Intrinsic and Extrinsic Motivation,” *The Review of Economic Studies*, 2003, 70 (3), 489–520.
- and ———, “Individual and corporate social responsibility,” *Economica*, 2010, 77 (305), 1–19.
- Berman, Jonathan Z, Emma E Levine, Alixandra Barasch, and Deborah A Small**, “The braggart’s dilemma: On the social rewards and penalties of advertising prosocial behavior,” *Journal of Marketing Research*, 2015, 52 (1), 90–104.
- Besley, Timothy and Maitreesh Ghatak**, “Competition and Incentives with Motivated Agents,” *American Economic Review*, 2005, 95 (3), 616–636.
- Bhattacharya, Chitra B, Sankar Sen, and Daniel Korschun**, “Using corporate social responsibility to win the war for talent,” *MIT Sloan Management Review*, 2008, (2), 37–44.
- Bode, Christiane, Jasjit Singh, and Michelle Rogan**, “Corporate social initiatives and employee retention,” *Organization Science*, 2015, 26 (6), 1702–1720.
- Bradler, Christiane, Robert Dur, Susanne Neckermann, and Arjan Non**, “Employee recognition and performance: A field experiment,” *Management Science*, 2016, 62 (11), 3085–3099.
- Brekke, Kjell Arne and Karine Nyborg**, “Attracting responsible employees: Green production as labor market screening,” *Resource and Energy Economics*, 2008, 30 (4), 509–526.
- Burbano, Vanessa**, “Corporate Social Responsibility and Firm Performance: Field Experimental Evidence from Online Labor Marketplaces on the Role of Employee Salary Requirements,” *Organization Science*, 2016, (4), 1010–1028.
- Carlos, W Chad and Ben W Lewis**, “Strategic Silence: Withholding Certification Status as a Hypocrisy Avoidance Tactic,” *Administrative Science Quarterly*, 2017, p. Forthcoming.

- Carnahan, Seth, David Kryscynski, and Daniel Olson**, “When Does Corporate Social Responsibility Reduce Employee Turnover? Evidence from Attorneys Before and After 9/11,” *Academy of Management Journal*, 2016, pp. amj–2015.
- Carpenter, Jeffrey and David Dolifka**, “Exploitation Aversion: When Financial Incentives Fail to Motivate Agents,” *IZA Discussion Papers*, 2013.
- and **Erick Gong**, “Motivating Agents: How much does the mission matter?,” *Journal of Labor Economics*, 2016, 34 (1), 211–236.
- Cassar, Lea**, “Job Mission as a Substitute for Monetary Incentives: Benefits and Limits,” *Management Science*, 2017, p. Forthcoming.
- , “Optimal Contracting with Endogenous Project Mission,” *CESifo WP 6181*, 2017.
- and **Stephan Meier**, “Non-monetary Incentives and the Quest for Work Meaning,” *Working Paper*, 2016.
- Charness, Gary, Ramon Cobo-Reyes, and Angela Sanchez**, “The effect of charitable giving on workers’ performance. Experimental evidence,” 2014. Working Paper.
- DellaVigna, Stefano and Devin Pope**, “What Motivates Effort? Evidence and Expert Forecasts,” *Review of Economic Studies*, 2016, p. Forthcoming.
- Delmas, Magali A and Vanessa Cuerel Burbano**, “The drivers of greenwashing,” *California Management Review*, 2011, 54 (1), 64–87.
- Dur, Robert**, “Gift exchange in the workplace: Money or attention?,” *Journal of the European Economic Association*, 2009, 7 (2-3), 550–560.
- and **Max van Lent**, “Serving the Public Interest in Several Ways: Theory and Empirics,” *Working Paper*, 2016.
- Elfenbein, Daniel W, Ray Fisman, and Brian McManus**, “Charity as a substitute for reputation: Evidence from an online marketplace,” *The Review of Economic Studies*, 2012, 79 (4), 1441–1468.
- Ellingsen, Tore and Magnus Johannesson**, “Pride and prejudice: The human side of incentive theory,” *The American Economic Review*, 2008, 98 (3), 990–1008.
- Esteves-Sorenson, Constanca**, “Gift exchange in the workplace: Addressing the conflicting evidence with a careful test,” *Management Science*, 2017.
- Falk, Armin and Michael Kosfeld**, “The Hidden Costs of Control,” *The American Economic Review*, 2006, 96 (5), pp. 1611–1630.
- Fehr, Ernst, Georg Kirchsteiger, and Arno Riedl**, “Does fairness prevent market clearing? An experimental investigation,” *The quarterly journal of economics*, 1993, 108 (2), 437–459.
- Fehrler, Sebastian and Michael Kosfeld**, “Pro-Social Missions and Worker Motivation: An Experimental Study,” *Journal of Economics Behavior and Organization*, 2014, 100, 99 – 110.
- and **Wojtek Przepiorka**, “Charitable giving as a signal of trustworthiness: Disentangling the signaling benefits of altruistic acts,” *Evolution and Human Behavior*, 2013, 34 (2), 139 – 145.
- and ———, “Choosing a partner for social exchange: Charitable giving as a signal of trustworthiness,” *Journal of Economic Behavior & Organization*, 2016, 129, 157–171.
- Feltovich, Nicholas J, R. Harbaugh, and T. To**, “Too Cool for School? Signalling and

- Countersignalling,” *The RAND Journal of Economics*, 11 2002, 33 (4), 630–649.
- Flammer, Caroline and Jiao Luo**, “Corporate social responsibility as an employee governance tool: Evidence from a quasi-experiment,” *Strategic Management Journal*, 2017, 38 (2), 163–183.
- Frey, Bruno and Jana Gallus**, “Towards an economics of awards,” *Journal of Economic Surveys*, 2017, 31 (1), 190–200.
- Gallus, Jana**, “Fostering Public Good Contributions with Symbolic Awards: A Large-Scale Natural Field Experiment at Wikipedia,” *Management Science*, 2016, p. Forthcoming.
- Gartenberg, Claudine, Andrea Prat, and George Serafeim**, “Corporate Purpose and Financial Performance,” *Working Paper*, 2016.
- Gneezy, Uri and Aldo Rustichini**, “Pay enough or don’t pay at all,” *The Quarterly Journal of Economics*, 2000, 115 (3), 791–810.
- **and John A List**, “Putting behavioral economics to work: Testing for gift exchange in labor markets using field experiments,” *Econometrica*, 2006, 74 (5), 1365–1384.
- , **Stephan Meier, and Pedro Rey-Biel**, “When and why incentives (don’t) work to modify behavior,” *The Journal of Economic Perspectives*, 2011, 25 (4), 191–209.
- Goodman, Joseph K, Cynthia E Cryder, and Amar Cheema**, “Data collection in a flat world: The strengths and weaknesses of Mechanical Turk samples,” *Journal of Behavioral Decision Making*, 2013, 26 (3), 213–224.
- Haggag, Kareem and Devin G Pope**, “Attribution Bias in Consumer Choice,” *Working Paper*, 2016.
- Hedblom, Daniel, Brent R Hickman, and John A List**, “Toward an Understanding of Corporate Social Responsibility: Theory and Field Experimental Evidence,” 2016.
- Henderson, Rebecca and Eric Van den Steen**, “Why Do Firms Have “Purpose”? The Firm’s Role as a Carrier of Identity and Reputation,” *The American Economic Review*, 2015, 105 (5), 326–330.
- Hennig-Schmidt, Heike, Abdolkarim Sadrieh, and Bettina Rockenbach**, “In search of workers’ real effort reciprocity? a field and a laboratory experiment,” *Journal of the European Economic Association*, 2010, 8 (4), 817–837.
- Horton, John Joseph and Lydia B Chilton**, “The labor economics of paid crowdsourcing,” *Proceedings of the 11th ACM conference on Electronic commerce*, 2010, pp. 209–218.
- Imas, Alex**, “Working for the “warm glow”: On the benefits and limits of prosocial incentives,” *Journal of Public Economics*, 2014, 114, 14 – 18.
- , “The realization effect: Risk-taking after realized versus paper losses,” *The American Economic Review*, 2016, 106 (8), 2086–2109.
- , **Michael Kuhn, and Vera Mironova**, “Waiting to Choose,” *Working Paper*, 2016.
- Johnson, Eric, Stephan Meier, and Olivier Toubia**, “What’s The Catch? Exploring sluggish mortgage refinancing using administrative data, surveys, and field experiments,” *Working Paper*, 2016.
- Kajackaite, Agne and Dirk Sliwka**, “Social responsibility and incentives in the lab: Why do agents exert more effort when principals donate?,” *Journal of Economic Behavior & Organization*, 2017, 142, 482 – 493.
- Koppel, Hannes and Tobias Regner**, “Corporate Social Responsibility in the work place,”

- Experimental Economics*, 2014, 17 (3), 347–370.
- Kosfeld, Michael and Susanne Neckermann**, “Getting more work for nothing? Symbolic awards and worker performance,” *American Economic Journal: Microeconomics*, 2011, 3 (3), 86–99.
- , ———, and **Xiaolan Yang**, “Knowing that you matter, matters! The interplay of meaning, monetary incentives, and worker recognition,” *Economic Inquiry*, 2014, p. Forthcoming.
- Levine, David K**, “Modeling altruism and spitefulness in experiments,” *Review of economic dynamics*, 1998, 1 (3), 593–622.
- Neckermann, Susanne, Reto Cueni, and Bruno S Frey**, “Awards at work,” *Labour Economics*, 2014, 31, 205–217.
- Newman, George E and Daylian M Cain**, “Tainted altruism: When doing some good is evaluated as worse than doing no good at all,” *Psychological science*, 2014, 25 (3), 648–655.
- Nyborg, Karine**, “Do responsible employers attract responsible employees?,” *IZA World of Labor*, 2014.
- and **Tao Zhang**, “Is Corporate Social Responsibility Associated with Lower Wages?,” *Environmental and Resource Economics*, 2013, 55 (1), 107–117.
- Porter, Michael E and Mark R Kramer**, “The Link Between Competitive Advantage and Corporate Social Responsibility,” *Harvard business review*, 2007.
- PricewaterhouseCoopers**, “Redefining business success in a changing world: CEO Survey,” Technical Report, *PricewaterhouseCoopers 19th Annual Global CEO Survey* 2016.
- Singh, Jasjit, Nina Teng, and Serguei Netessine**, “Philanthropic Campaigns and Customer Behavior: Field Experiments in an Online Taxi Booking Company,” 2016.
- Sliwka, Dirk**, “Trust as a Signal of a Social Norm and the Hidden Costs of Incentive Schemes,” *American Economic Review*, 2007, 97 (3), 999–1012.
- Tonin, Micro and Michael Vlassopoulos**, “Disentangling the sources of pro-socially motivated effort: A field experiment,” *Journal of Public Economics*, 2010, 94 (11-12), 1086–1092.
- Tonin, Mirco and Michael Vlassopoulos**, “Corporate philanthropy and productivity: Evidence from an online real effort experiment,” *Management Science*, 2015, 61 (8), 1795–1811.

## A. Appendix

### A.1. Proof of Proposition 1

Consider the following worker's belief on the employer's type:  $\hat{q}_{d_C} = E_A(q \mid d = d_C) = 0$  and  $\hat{q}_{d_U} = E_A(q \mid d = d_U) = 1$ . Given these beliefs we can derive the agents' optimal effort levels. If an agent is offered an unconditional donation he exerts effort that maximizes  $w - \lambda e - \frac{1}{2}e^2 + \theta_A \bar{d}$ . This give an optimal effort level that is independent of  $\theta_A$ . Hence, both motivated and non-motivated agents when offered an unconditional donation, exert effort equal to  $\lambda$ , that is  $e_{d_U}^m = e_{d_U}^{nm} = \lambda$ . A non-motivated agent who is offered the conditional donation chooses effort that maximizes  $w - \frac{1}{2}e^2$ , that is  $e_{d_C}^{nm} = 0$ . Finally, a motivated agent who is offered a conditional donation chooses effort that maximizes

$$\begin{cases} w - \frac{1}{2}e^2 + \theta_A \bar{d} & \text{if } e \geq \bar{e} \\ w - \frac{1}{2}e^2 & \text{if } e < \bar{e} \end{cases} \quad (\text{A-1})$$

The first part of the function is maximized at  $e = \bar{e}$  while the second part of the function is maximized at  $e = 0$ . Given the assumption that  $\theta_A > \frac{\bar{e}}{2\bar{d}}$ , it is then easy to see that the agent's optimal level of effort  $e_{d_C}^m$  is equal to  $\bar{e}$ .

Next, for a separating equilibrium to exist, it must be the case that given these agents' equilibrium beliefs and effort levels, the expected utility of sufficiently prosocial employers is higher when offering unconditional donations than conditional donations, while the opposite must be true for selfish employers. The employer's expected utility from offering an unconditional donation is equal to

$$U_P(d_U) = (\theta_P - 1)\bar{d} + p\pi e_{d_U}^m + (1 - p)\pi e_{d_U}^{nm}$$

which, once we substitute for the the agent's equilibrium effort levels gives:

$$U_P(d_U) = (\theta_P - 1)\bar{d} + \pi\lambda \quad (\text{A-2})$$

The employer's expected utility from offering a conditional donation is equal to

$$U_P(d_C) = p\left((\theta_P - 1)\bar{d} + \pi e_{d_C}^m\right) + (1 - p)\pi e_{d_C}^{nm}$$

which, once we substitute for the the agent's equilibrium effort levels gives:

$$U_P(d_C) = p\left((\theta_P - 1)\bar{d} + \pi\bar{e}\right) \quad (\text{A-3})$$

By comparing equations A-2 and A-3 it is then trivial to show that  $U_P(d_U) > U_P(d_C)$  iff  $\theta_P > \bar{\theta}_P \equiv \frac{\pi(\bar{e}-\lambda)}{(1-p)\bar{d}} + 1$ . Hence, as long as  $\bar{\theta}_P > 0$ , sufficiently prosocial principals will donate unconditionally whereas selfish principals (who, recall, have  $\theta_P = 0$ ) will donate conditionally on the agent's performance. Since workers' beliefs are consistent with employer's strategies and employer's strategies make sense given the workers' beliefs, there exists a signalling equilibrium in which sufficiently prosocial employers signal their prosociality by donating unconditionally.

## A.2. Extension of model with countersignalling

The result that effort is higher in the baseline than in the (conditional and unconditional) charitable incentive treatments can be in line with our theory of the signaling value of charitable incentives. To see how this result can be sustained in equilibrium, consider an extension of the illustrative model presented in section 2 that incorporates the possibility of countersignaling (Feltovich et al., 2002). Let there be four types of employers depending on the size of their  $\theta_P$ : with high prosociality ( $H$ ), medium-high prosociality ( $M_h$ ), medium-low prosociality ( $M_l$ ) and low prosociality ( $L$ ). Also, let us enrich the set of strategies available to the employers. The latter can not only decide between offering conditional and unconditional donations but they can also opt for not donating at all. Furthermore, they can decide between revealing their charitable initiatives or not revealing them. The strategy of “not revealing” means that the employer does not provide any information to her employees about her charitable initiatives—which does not mean, however, that the employer is not making the unconditional donation, she may very well donate and keep it secret.<sup>23</sup> Our baseline treatment corresponds to this “not revealing” strategy rather than to the “no charitable incentive” strategy because PharmaGIC does not state that they do not donate, they simply do not mention any donation. Crucial for countersignalling to emerge, let also assume that the worker receives some extra noisy information on the employer’s type. For example, this could be the wage offered and the pleasantness of the task, which are assumed to be exogenous. Employers  $L$  offer low wage and tedious tasks, employers  $H$  offer high wages and motivating tasks, while employers  $M_h$  and  $M_l$  offer high wage and motivating tasks with probability  $h$ . What will the employers do? Consider if the worker believes that only employers  $M_h$  and  $M_l$  reveal their charitable initiatives. Then if these employers don’t reveal their charitable initiatives they take the chance of being perceived as a  $H$ -type with probability  $h$  and as a  $L$ -type with probability  $(1 - h)$ . If  $h$  is sufficiently low or the worker’s disutility of being matched with a  $L$ -type is sufficiently large, it may be too risky not to reveal. Furthermore,  $M_h$  may want to distinguish herself from  $M_l$ , so if the difference between the  $\theta_P$  of  $M_h$  and  $M_l$  is sufficiently large,  $M_h$  will choose to reveal that they donate unconditionally, while  $M_l$  will choose to donate conditional on workers’ performance, which by definition is revealed. For  $H$ -types, however, the situation is different because given their high wage and their pleasant task they do not have to worry about being perceived as an  $L$ -type. They face the clear choice between being perceived as  $M_h$  if they reveal their unconditional donation and as  $H$  if they do not reveal it. Finally for  $L$  it may be too costly to donate so they will not donate and not reveal it. Since workers’ beliefs are consistent with employer’s strategies and employer’s strategies make sense given the workers’ beliefs, there exists a countersignaling equilibrium in which employers  $H$  signal their high prosociality by not revealing their donations.

Given the very generous incentives in the baseline, the workers who were allocated to the baseline treatment and thus were not revealed any charitable initiative, ruled out the possibility that the employer was of a  $L$  type, being ascertained that it must have been a type  $H$ . On the other hand, those workers who were allocated to the unconditional donation treatment must have thought that the employer was of type  $M_h$ , while those workers who were allocated to the conditional donation treatment must have thought that the employer was of type  $M_l$ . This could be the workers’ effort choices in our experiment.

---

<sup>23</sup>Obviously the conditional donation is by definition revealed. So the employers can only keep secret their unconditional donations.

### A.3. Additional Tables and Figures

Table A.1: Non-motivated and Motivated Types by Treatment

Treatment	Non-motivated	Motivated	Total
Unconditional Monetary	163 51.42%	154 48.58%	317
Unconditional Monetary w/ Explanation	153 50%	153 50%	306
Conditional Monetary	149 49.83%	150 50.17%	299
Unconditional Monetary w/ Explanation	155 50.16%	154 49.84%	309
Unconditional Charitable	154 51.16%	147 48.84%	301
Unconditional Charitable w/ Explanation	162 52.77%	145 47.23%	307
Conditional Charitable	127 44.25%	160 55.75%	287
Conditional Charitable w/ Explanation	128 43.69%	165 56.31%	293
Baseline	138 46%	162 54%	300
Total	1,329 48.88%	1,390 51.12%	2,719 100%

Table A.2: Logit Regressions

Sample:	All	Non-Motivated	Motivated
Corresponding Table & Column:	3 (1)	4 (1)	4 (2)
Charitable Incentive (=1)	-0.723*** (0.115)	-0.722*** (0.167)	-0.880*** (0.190)
Conditional Incentive (=1)	0.815*** (0.136)	1.375*** (0.231)	1.501*** (0.302)
Charitable Incentive $\times$ Conditional	-0.995*** (0.175)	-1.819*** (0.287)	-1.367*** (0.347)
Constant	0.876*** (0.085)	0.875*** (0.123)	1.457*** (0.146)
<i>F</i> -test:			
“Cond.” + “Charit. $\times$ Cond.”=0	$p = 0.10$	$p < 0.01$	$p = 0.43$
Pseudo $R^2$	0.0690	0.1169	0.0903
Observations	2,670	1,191	1,228

*Notes:* Coefficients of Logit regressions with robust standard errors in parenthesis. Dependent variable is 1 if worker did more than three slogans and 0 otherwise. Significance levels: \*\*\*  $p < .01$ , \*\*  $p < .05$ , \*  $p < .1$ .



Table A.3: Self-Reported Perception of Firm and Incentives

	Soc. resp.	Attractive	Satisfied	Calculated
Charitable Incentive (=1)	0.405*** (0.102)	-0.255** (0.104)	-0.375*** (0.083)	-0.131 (0.114)
Conditional Incentive (=1)	-0.343*** (0.105)	-0.074 (0.105)	-0.484*** (0.088)	0.139 (0.116)
Charitable Incentive $\times$ Conditional	-0.003 (0.149)	-0.099 (0.151)	0.140 (0.136)	-0.266 (0.168)
Constant	7.239*** (0.072)	7.482*** (0.073)	9.056*** (0.055)	7.741*** (0.083)
<i>F</i> -test:				
“Cond.” + “Charit. $\times$ Cond.”=0	$p < 0.01$	$p = 0.11$	$p < 0.01$	$p = 0.29$
Adj. $R^2$	0.020	0.007	0.023	0.004
Observations	2,418	2,420	2,418	2,418

*Notes:* Coefficients of OLS regressions with robust standard errors in parenthesis. Dependent variables are answers on 10-point scale to four questions about the firm (see Section A.4 for details). “Soc. resp.”: *Please rate PharmaGIC on a scale from 0 “Not at all socially responsible” to 10 “Very socially responsible”*; “Attractive”: *How attractive would our company be as a potential employee?*; “Satisfied”: *How satisfied were you with the incentives that we provided for this task?*; “Calculated”: *How calculated do you think was our choice of incentives?*. Significance levels: \*\*\*  $p < .01$ , \*\*  $p < .05$ , \*  $p < .1$ .

Table A.4: Non-Motivated vs. Motivated Workers – Robustness

	(1)	(2)	(2)
Classifying workers who drop out:	excluded	all non-motivated	all motivated
Charitable Incentive (=1)	-0.168*** (0.038)	-0.158*** (0.036)	-0.168*** (0.038)
Conditional Incentive (=1)	0.199*** (0.031)	0.138*** (0.034)	0.199*** (0.031)
Charitable Incentive $\times$ Conditional	-0.309*** (0.052)	-0.274*** (0.050)	-0.309*** (0.052)
Motivated Type (=1)	0.105*** (0.034)	0.194*** (0.034)	0.001 (0.035)
Motivated $\times$ Charitable	-0.003 (0.052)	-0.012 (0.051)	-0.000 (0.052)
Motivated $\times$ Conditional	-0.059 (0.040)	0.001 (0.043)	-0.111** (0.045)
Motivated $\times$ Charitable $\times$ Conditional	0.200*** (0.069)	0.164** (0.068)	0.217*** (0.071)
Constant	0.706*** (0.026)	0.617*** (0.026)	0.706*** (0.026)
<i>F</i> -test:			
“Cond.” + “Charit. $\times$ Cond.” = 0	$p < 0.01$	$p < 0.01$	$p < 0.01$
Adj. $R^2$	0.126	0.145	0.089
Observations	2,419	2,670	2,670

*Notes:* Table shows the robustness of our results to different classifications into motivated and non-motivated workers for workers who dropped out before indicating their prosociality. Column (1) shows the main result in the paper that excludes those workers from the analysis. Column (2) assumes that all workers who drop out are non-motivated. Column (3) assumes that all workers who drop out are motivated. Significance levels: \*\*\*  $p < .01$ , \*\*  $p < .05$ , \*  $p < .1$ .

Table A.5: IV Regression

	(1)	(2)
First-stage regression:	DV=Soc. resp.	
Conditional Charitable Incentive	-0.399** (-2.53)	-0.399** (-2.53)
Constant	7.556*** (75.68)	7.556*** (75.68)
First-Stage $R^2$	0.0114	0.0114
First-Stage F-Stat	6.41	6.41
Instrumental variables (2SLS) regression	> 3 slogans	# slogans
Soc. resp.	0.274** (1.98)	0.746 (1.59)
Constant	-1.530 (1.018)	-0.897 -0.26
Wald $\chi^2$	3.94	2.53
Observations	570	570

*Notes:* Table shows results of a 2SLS Regression for non-motivated agents in the charitable incentives treatments. Dependent variable is 1 if worker did more than three slogans and 0 otherwise (in Column 1) and the number of slogans (in Column 2). “Soc. resp.”: *Please rate PharmaGIC on a scale from 0 “Not at all socially responsible” to 10 “Very socially responsible”*. Significance levels: \*\*\* p<.01, \*\* p<.05, \* p<.1.

Table A.6: Comparison to Baseline

Dependent Variables:	(1) > 3 slogans	(2) # slogans
Baseline	Reference Group	
Unconditional Charitable	-0.087** (0.038)	-0.425** (0.216)
Unconditional Charitable with Explanation	-0.064* (0.038)	-0.264 (0.217)
Conditional Charitable	-0.111*** (0.038)	-0.563** (0.220)
Conditional Charitable with Explanation	-0.130*** (0.038)	-0.544** (0.222)
Unconditional Monetary	0.112*** (0.036)	0.729*** (0.237)
Unconditional Monetary with Explanation	0.073** (0.037)	0.364 (0.232)
Conditional Monetary	0.219*** (0.034)	1.060*** (0.236)
Conditional Monetary with Explanation	0.243*** (0.033)	1.201*** (0.224)
Constant	0.614*** (0.027)	4.934*** (0.166)
Adj. $R^2$	0.074	0.048
Observations	3,004	3,004

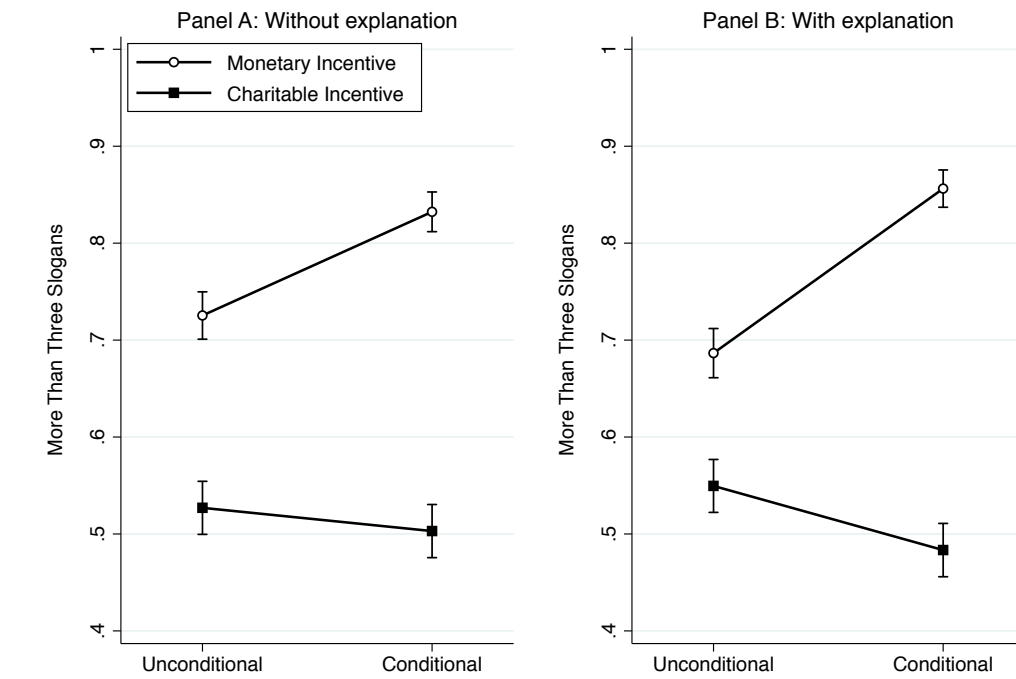
*Notes:* OLS regressions with robust standard errors. Dependent variables: dummy variable equal 1 if the worker created more than 3 slogans and 0 otherwise in (1), and the number of created slogans in (2). Significance levels: \*\*\*  $p < .01$ , \*\*  $p < .05$ , \*  $p < .1$ .

Table A.7: Strategic and non-strategic CSR

Dimensions	Potential ‘Benefit’	Strategic	Non-strategic	Difference
Attractiveness	Image	6.27	8.78	$p < 0.01$
	Effort	6.34	8.24	$p < 0.01$
	Lower Wage	4.6	6.87	$p < 0.01$
Generous donation	Image	5.11	8.61	$p < 0.01$
	Effort	5.5	8.79	$p < 0.01$
	Lower Wage	4.58	8.19	$p < 0.01$
Socially responsible	Image	6.52	9.06	$p < 0.01$
	Effort	6.22	9.22	$p < 0.01$
	Lower Wage	5.48	8.85	$p < 0.01$
Acceptance lower wage	Image	0.11	0.25	$p = 0.07$
	Effort	0.12	0.22	$p = 0.17$
	Lower Wage	0.18	0.19	$p = 0.92$
Motivation	Image	6.48	7.67	$p < 0.01$
	Effort	6.2	7.67	$p < 0.01$
	Lower Wage	5.14	6.62	$p < 0.01$

*Notes:* Table shows responses about different dimensions of a potential employer: “Attractiveness to work for” [on 11-point Likert scale], “Perceived generosity” [on 11-point Likert scale], “Perceived social responsibility of firm” [on 11-point Likert scale], “Willing to accept lower wage” [Yes or No], “Motivated by donation by firm” [on 11-point Likert scale]. Responses are shown for scenarios with different benefits for firm on “Image”, “Worker Effort”, and “Lower Wages”, and for whether they firm was strategic or non-strategic about their CSR activities. See Section A.5 for all the details. The last column shows  $p$ -values of Mann-Whitney tests.

Figure A.1: With and Without Explanation



*Notes:* The graph shows the effect of making monetary or charitable incentives performance-based (“Conditional”) or not performance-based (“Unconditional”) on the proportion of workers creating more than three slogans. The panels show the effect for the treatments without explanation (Panel A) and with explanation (Panel B). Bars shows standard errors of the mean.

## A.4. Instructions

### A.4.1. Introduction text

Thanks for working for us. We are an Italian company, named PharmaGIC srl, that distributes pharmaceutical products. We operate in Italy and abroad. We are currently working on a new English version of our website, through which we promote our products. For this purpose, we collect ideas on possible marketing slogans (taglines) for some of our products in the pharmacy division (mainly childcare products, cosmetics, hygiene, and so on). Slogans typically convey a message about the product, which can be implicit or explicit. A slogan can be a few simple words or an entire phrase. The aim of a slogan is to catch the audiences attention and to make the product more attractive at the eyes of the consumer. Examples of famous slogans include: ‘Just do it’ (Nike), ‘Think different’ (Apple), ‘I’m lovin’ it’ (MacDonald), ‘Connecting people’ (Nokia), ‘The best a Man can get’ (Gillette), ‘Because you’re worth it’ (LOreal) For more information about us please visit our website: [www.pharmagic.net](http://www.pharmagic.net)

We have in total 21 products. Your task is to create at least three slogans. One slogan for each of three different products. But of course, we appreciate if you decide to create slogans for more than three products. The products will be shown to you in sequence. The next product will be shown to you only after you created the slogan for the previous product. For each product you will see a detailed description and a picture. Underneath the product description you will find a blank box. In that blank box you can write the slogan(s). Once you are sure about the slogan please select ‘submit’.

### A.4.2. Treatment Text

- *Treatment 1 (money; unconditional, no explanation)*: The task is considered completed if you create 3 slogans. Upon completion of the task, you will be paid the pre-announced wage of \$1.50. Furthermore, you will receive a bonus of 75 cents. This bonus is given unconditionally on whether you create any extra slogans. However, we would really appreciate if you could create at least 6 slogans.
- *Treatment 2 (money; unconditional, explanation)*: The task is considered completed if you create 3 slogans. Upon completion of the task, you will be paid the pre-announced wage of \$1.50. Furthermore, you will receive a bonus of 75 cents. This bonus is given unconditionally on whether you create any extra slogans. However, we would really appreciate if you could create at least 6 slogans.  
Why do we give you the bonus? We are strongly committed to be an employee-friendly company, even if this implies sacrificing some profit.
- *Treatment 3 (money; conditional, no explanation)*: The task is considered completed if you create 3 slogans. Upon completion of the task, you will be paid the pre-announced wage of \$1.50. Furthermore, if you create (at least) 6 slogans instead of 3, you will receive a bonus of 75 cents. This bonus is given conditionally on you creating at least three extra slogans. Indeed, we would really appreciate if you could create at least 6 slogans.
- *Treatment 4 (money; conditional, explanation)*: The task is considered completed if you create 3 slogans. Upon completion of the task, you will be paid the pre-announced wage of \$1.50. Furthermore, if you create (at least) 6 slogans instead of 3, you will receive a bonus of 75 cents. This bonus is given conditionally on you creating at least three extra slogans. Indeed, we would really appreciate if you could create at least 6 slogans. Why do we give you the bonus? The bonus is profitable for us: you doing (at least) three extra slogans for a wage of 75 cents is less costly for us than hiring another worker to do (at least) three slogans for a wage of \$1.50.
- *Treatment 5 (charitable; unconditional, no explanation)*: The task is considered completed if you create 3 slogans. Upon completion of the task, you will be paid the pre-announced wage of \$1.50. Furthermore, we will make a donation of 75 cents to Doctors without Borders (US). This donation is made unconditionally on whether you create any extra slogans. However, we would really appreciate if you could create at least 6 slogans. We will post the receipt of the donation online for you to verify 14 days after we collected all the slogans (and donations). Totally anonymous.

- *Treatment 6 (charitable; unconditional, explanation)*: The task is considered completed if you create 3 slogans. Upon completion of the task, you will be paid the pre-announced wage of \$1.50. Furthermore, we will make a donation of 75 cents to Doctors without Borders (US). This donation is made unconditionally on whether you create any extra slogans. However, we would really appreciate if you could create at least 6 slogans. We will post the receipt of the donation online for you to verify 14 days after we collected all the slogans (and donations). Totally anonymous.

Why do we make a donation? We are strongly committed to be a socially responsible company (e.g. helping the larger community), even if this implies sacrificing some profit.

- *Treatment 7 (charitable; conditional, no explanation)*: The task is considered completed if you create 3 slogans. Upon completion of the task, you will be paid the pre-announced wage of \$1.50. Furthermore, if you create (at least) 6 slogans instead of 3, we will make a donation of 75 cents to Doctors without Borders (US). This donation is made conditionally on you creating at least three extra slogans. Indeed, we would really appreciate if you could create at least 6 slogans. We will post the receipt of the donation online for you to verify 14 days after we collected all the slogans (and donations). Totally anonymous.
- *Treatment 8 (charitable; conditional, explanation)*: The task is considered completed if you create 3 slogans. Upon completion of the task, you will be paid the pre-announced wage of \$1.50. Furthermore, if you create (at least) 6 slogans instead of 3, we will make a donation of 75 cents to Doctors without Borders (US). This donation is made conditionally on you creating at least three extra slogans. Indeed, we would really appreciate if you could create at least 6 slogans. We will post the receipt of the donation online for you to verify 14 days after we collected all the slogans (and donations). Totally anonymous.

Why do we make the donation? The donation is profitable for us: you doing (at least) three extra slogans for a donation of 75 cents is less costly for us than hiring another worker to do (at least) three slogans for a wage of \$1.50.

- *Treatment 9 (baseline)*: The Task is considered completed if you create 3 slogans. Upon completion of the task, you will be paid the pre-announced wage of \$1.50. However, we would really appreciate if you could create at least 6 slogans.

How many slogans would you like to create? 3 Slogans or 6 Slogans



### A.4.3. Example Product

## SLOGAN 1

### Nippes - Teel tools made in Solingen since 1923

High range manicure and pedicure steel tools.

Nippes still enhances traditional methods of production and uses only the highest quality materials, taking care of the entire production process through rigorous testing quality system. Nippes uses only hot forged steel to ensure maximum durability of its instruments.



Enter your slogan:

### A.4.4. Survey Questions

In order to understand your experience with us, we kindly ask you to fill a short survey.

- What is your gender?  
*Male, Female, Other*
- How attractive would our company be as a potential employee?  
*“Not attractive at all” 0- 10 “Very attractive”*
- Please rate PharmaGIC on a scale from 0 “Not at all socially responsible” to 10 “Very socially responsible”  
*“Not at all socially responsible” 0- 10 “Very socially responsible”*
- How satisfied were you with the incentives that we provided for this task?  
*“Not at all satisfied” 0- 10 “Very satisfied”*
- What do you think:
  1. How many other M-Turkers did more than 3 slogans? *(0% - 100%)*
  2. How many other M-Turkers did more than 6 slogans? *(0% - 100%)*

3. How many slogans did the others do on average?
- How satisfying are the incentives for other M-Turkers?  
*“Not at all satisfied” 0- 10 “Very satisfied”*
  - How calculated do you think was our choice of incentives?  
*“Not at all calculated” 0- 10 “Very calculated”*
  - How often do you donate money to a charitable organization?  
*Never, Rarely, Sometimes, Often, Regularly*
  - How often do you volunteer for a good cause?  
*Never, Rarely, Sometimes, Often, Regularly*
  - Let us know your opinion about our task and our incentives:

## A.5. Vignette Study

### A.5.1. Details of Vignette Study

In 2015, we conducted a survey on M-Turk with 300 participants. We paid 50 cents for the completion of the survey. The analysis in the paper is based on the between-subject design of the study.

After participants ‘signed’ a consent form, we started the study and they got randomized into six different scenarios.

### A.5.2. Instructions and Vignettes

“What Are We Going to Ask from You?”

You will first be presented with a description of an hypothetical scenario regarding the behavior of a firm. Please read it carefully. Once you are done reading the scenario click continue. You will then be asked to answer a couple of control questions to make sure you understood the scenario correctly. In order to proceed further with the survey, you need to answer all these questions correctly. Finally, once you have answered all these questions right, you will be presented with several questions regarding your perceptions about the scenario. For these questions, there is not a right or wrong answer.

**Participants got randomized into one of six scenarios.**

1. Scenario 1 (better image, without market analysis):

A large, successful, and family-owned firm was considering whether to annually donate part of their profit to a renowned charity that supports education of poor children worldwide. The owner and the top management team decided to make the donation as part of its Corporate Social Responsibility program. The firm explicitly made the decision without making a market analysis whether the donation increases the firm’s profit because of a better image. “We do not view this as an investment in our firm’s prestige and we are not interested whether it has a return in an increased value of our brand”, said the CEO of the company.

2. Scenario 2 (better image, only after market analysis):

A large, successful, and family-owned firm was considering whether to annually donate part of their profit to a renowned charity that supports education of poor children worldwide. The owner and the top management team decided to make the donation as part of its Corporate Social Responsibility program. The firm explicitly made the decision only after making a market analysis whether the donation increases the firm’s profit because of a better image. “We view this mainly as an investment in our firm’s image and we are only interested whether it has a return in an increased value of our brand”, said the CEO of the company.

3. Scenario 3 (workers’ motivation, only after market analysis):

A large, successful, and family-owned firm was considering whether to annually donate part of their profit to a renowned charity that supports education of poor children worldwide. The owner and the top management team decided to make the donation as part of its Corporate Social Responsibility program. The firm explicitly made the decision only after making a market analysis whether the donation increases the firm’s profit because workers would be more motivated and work harder. “We view this mainly as an investment in our firm’s HR strategy and we are only interested whether it has a return in an increased effort and motivation of our workers”, said the CEO of the company.

4. Scenario 4 (workers’ motivation, without market analysis):

A large, successful, and family-owned firm was considering whether to annually donate part of their profit to a renowned charity that supports education of poor children worldwide. The owner and the top management team decided to make the donation as part of its Corporate Social Responsibility program.

The firm explicitly made the decision without making a market analysis whether the donation increases the firm's profit because workers would be more motivated and work harder. "We do not view this as an investment in our firm's HR strategy and we are not interested whether it has a return in an increased effort and motivation of our workers", said the CEO of the company.

5. Scenario 5 (lower wage, only after market analysis):

A large, successful, and family-owned firm was considering whether to annually donate part of their profit to a renowned charity that supports education of poor children worldwide. The owner and the top management team decided to make the donation as part of its Corporate Social Responsibility program. The firm explicitly made the decision only after making a market analysis whether the donation increases the firm's profit because workers would accept lower wages to work for the firm. "We view this mainly as an investment in our firm's HR strategy and we are only interested whether it has a return in that we can pay lower wages to our workers", said the CEO of the company

6. Scenario 6 (lower wage, without market analysis):

A large, successful, and family-owned firm was considering whether to annually donate part of their profit to a renowned charity that supports education of poor children worldwide. The owner and the top management team decided to make the donation as part of its Corporate Social Responsibility program. The firm explicitly made the decision without making a market analysis whether the donation increases the firm's profit because workers would accept lower wages to work for the firm. "We do not view this as an investment in our firm's HR strategy and we are not interested whether it has a return in that we can pay lower wages to our workers", said the CEO of the company

### Control questions

1. Is the firm family-owned? [Yes/No/Not mentioned]
2. Did the firm decide to make the donation? [Yes/No/Not mentioned]
3. Did the firm make a market analysis whether the donation increase profit? [Yes/No/Not mentioned]
4. Is the stated purpose of the donation to increase the firm's profit? [Yes/No/Not mentioned]

### Main questions

1. How attractive would this firm be as a potential employee? [Likert scale: 0-Not attractive at all - 10-Very attractive]
2. How generous do you perceive the donation to be? [Likert scale: 0-Not generous at all - 10-Very generous]
3. Please rate the firm on a scale from 0 "very socially responsible" to 10 "not at all socially responsible": [0-Very socially responsible - 10-Not at all socially responsible]
4. Suppose that you get a job offer at this firm. Would you be willing to accept a lower wage because of the charitable donation? [YES; NO]
5. Suppose that you work at this firm, what would be the effect of the charitable donation on your motivation to work? [-5 Very Discouraging; 0- No Effect; 5-Very Motivating]

### Within-Subject Design

In the following we will ask you a couple of questions about a similar but slightly different scenario. Consider again the previous scenario. Imagine that everything is the same (i.e. a firm is making a donation to a renowned charity that supports education of poor children worldwide). The only difference to the previous scenario is this firm now [add opposite scenario]

1. Would you find this company to be more or less attractive to work for than the other? [-5 Much less attractive; 0- Equally attractive; 5- Much more attractive]
2. How generous do you perceive the donation to be compared to the previous one? [-5 Much less generous; 0- Equally generous; 5- Much more generous]

3. How socially responsible do you perceive this firm to be compared to the previous one? [-5 Much less socially responsible; 0- Equally socially responsible; 5- Much more socially responsible]
4. Which of these sentences reflect better your preferences? Please select ONE:
  - “To accept a job offer at this firm I would need to be paid more than in the previous one”
  - “If the job offer of this firm pays the same as the previous one, I would be indifferent between the two.”
  - “I would be willing accept a job offer at this firm for a lower wage compared to the previous one.”
5. How motivated would you be to work at this firm compared to the previous one? [-5 Much less motivated ; 0- Equally motivated; 5- Much more motivated]

#### **Additional questions**

1. Do you donate? [Never, Rarely, Some times, Often, Regularly]
2. How important do you think it is to support education to poor children? [It's a priority, It's very important, It's as Important as many other social goals, It's less important than many other social goals, It's totally useless]
3. Have you ever volunteered? [Never, Rarely, Some times, Often, Regularly]
4. If you didn't answer 'never' in the previous question, what type of volunteering did you do? [please describe]
5. What do you think about successful firms who make donations to benefit society? [They have a moral obligation to do it, They don't have any moral obligation to do it, so I very much appreciate if they do it, I am indifferent on whether they do it or not, It's not the role of firms to make donations. The money used to make donations should be used to increase wages]
6. Gender [F, M]
7. Age: [please select]
8. Profession: [please write]