http://pubsonline.informs.org/journal/mnsc

MANAGEMENT SCIENCE
Articles in Advance, pp. 1–14
ISSN 0025-1909 (print), ISSN 1526-5501 (online)

When Corporate Social Responsibility Backfires: Evidence from a Natural Field Experiment

John A. List,[a] **Fatemeh Momeni**[a]

[a] Department of Economics, University of Chicago, Chicago, Illinois 60637
Contact: jlist@uchicago.edu (JAL); fmomeni@uchicago.edu, http://orcid.org/0000-0001-5245-837X (FM)

Received: April 26, 2018
Revised: March 22, 2019; June 24, 2019; July 23, 2019
Accepted: August 8, 2019
Published Online in Articles in Advance: June 8, 2020

https://doi.org/10.1287/mnsc.2019.3540

Abstract. This paper uses a natural field experiment to connect corporate social responsibility (CSR) to an important but often neglected behavior: employee misconduct and shirking. Through employing more than 1,500 workers, we find that our use of CSR increases employee misbehavior—24% more employees act detrimentally toward our firm by shirking on their primary job duties when we introduce CSR. Observed data patterns across the treatments are consonant with a model of "moral licensing," whereby the "doing good" nature of CSR induces workers to misbehave on another dimension that is harmful to the firm.

History: Accepted by Yan Chen, decision analysis.

Keywords: corporate social responsibility • natural field experiment • employee • misbehavior

1. Introduction

Corporate social responsibility (CSR) has become a cornerstone of modern business practice, developing from a "why" in the 1960s to a "must" today. Indeed, CSR has become one of the most common business practices in recent decades, as over 90% of G250 companies (the top 250 companies listed in the Fortune Global 500 ranking) now publish annual CSR reports (KPMG 2015). Although such a prevalence in CSR may seem at odds with the idea that the only responsibility of a business is to maximize its profit (Milton Friedman's famous quip in the *New York Times*: "The business of business is business"), recent studies have tried to understand why firms might strategically engage in CSR to increase their profits.[1]

On the demand side, researchers have argued that CSR can be profitable either because it acts as a signal of product quality (Fisman et al. 2008) or because consumers value and are willing to pay higher prices for responsibly produced goods (Sen and Bhattacharya 2001; Sen et al. 2006; Bagnoli and Watts 2003; Besley and Ghatak 2007; Du et al. 2011; Elfenbein et al. 2012; Lii and Lee 2012). On the supply side, workers might have preferences for being employed by a CSR company, working harder when their employer engages in CSR. In this spirit, companies might be able to increase profits by reducing labor costs through CSR (Turban and Greening 1997, Greening and Turban 2000, Backhaus et al. 2002, Besley and Ghatak 2005, Brekke and Nyborg 2008). Yet the empirical evidence on the supply-side profitability of CSR has solely focused on direct productivity effects (Tonin and Vlassapoulos 2014; Burbano 2016a, b; Hedblom et al. 2016; Cassar and Meier 2017; Gubler et al. 2018).

In this paper, we explore another supply-side channel through which CSR can affect profitability: The impact on employee misbehavior on the job. Employee misbehavior is a common and costly problem facing businesses and organizations: Estimates indicate that companies lose about 5% of their annual revenues to various forms of internal fraud (Association of Certified Fraud Examiners 2016). Furthermore, the National Retail Federation reports that in the retail industry alone, employee theft amounted to $15 billion (over a third of the total inventory shrinkage) in 2014 (National Retail Federation 2015). Other sources report that 1 in every 27 employees in the retail industry was apprehended for theft from their employers in 2016.[2]

Given the magnitude of the cost that organizations bear from various forms of employee misbehavior, understanding whether and how common CSR practices interact with employee misbehavior holds great import. We focus on two main channels through which CSR can influence worker misbehavior. First, it can serve as a social incentive tool for motivating workers to reduce unethical and counterproductive behavior on the job. Previous studies have shown that, consistent with the standard gift-exchange model, workers reciprocate a higher wage from their employer by reducing misbehavior on the job (Ockenfels et al. 2015, Flory et al. 2016). Similar to the way monetary incentives reduce worker misbehavior, social incentives in the form of CSR may reduce misbehavior by triggering reciprocity toward the employer.

Therefore, CSR may decrease worker misbehavior through the gift-exchange channel.

A second channel is that CSR can increase worker misbehavior through moral licensing. Prosocial behavior is promoted in part by self- and social-image motives: People act prosocially, in part, to signal to themselves (and to others) that they are good and moral individuals (Bénabou and Tirole 2006, 2011). Although prosocial deeds can boost individual self-image, unethical behavior can lower self-image. Because our moral standards are constantly challenged in multiple dimensions, good behavior in one domain may liberate us to behave unethically in another domain. Such a moral-licensing dynamic in individual behavior has been documented in the social psychology literature. Relatedly, moral licensing has been raised as a potential dark side of CSR's appeal to image concerns because "people who have recently 'done good' in one dimension may feel immunized against negative (social or self) inferences, and thus later on act less morally constrained" (Bénabou and Tirole 2010, p. 6).

The two channels work as opposing forces, opening up the potential for empirical work to provide important measurement of not only the direction but also the magnitude of the effect of CSR on misbehavior. To provide empirical content to the relative importance of each channel, we conducted a natural field experiment with over 1,500 workers who we hired ourselves. As an employer in an online labor market platform, we invited interested workers to our website to perform a short transcription task for payment. The task and the payment structure were designed in a manner that provided opportunities for workers to reciprocate or misbehave.

To study how CSR affects worker misbehavior, we randomized workers into one of three treatments across which we varied the CSR incentive. In addition, we explore the potential role of moral licensing by varying the appeal of CSR to workers' self-image by changing the framing of our CSR messages.

Overall, our results suggest that our use of CSR increased misbehavior. We find that the share of workers who shirk their primary job duties increases significantly—by roughly 24% or five percentage points—from the baseline to our CSR treatments. Indeed, CSR not only increases the propensity to misbehave, but it also increases the level of shirking: The average level of cheating by a worker increases by 11%. Complementary data patterns *suggest* that moral licensing may be the underlying reason for these results because good behavior in one domain seems to liberate us to behave unethically in another domain.

Previous empirical studies that investigated the supply-side profitability of CSR have generally reported positive effects on productivity (Burbano 2016a, b; Hedblom et al. 2016). An exception is a contemporaneous study by Cassar and Meier (2017) that documents a negative supply-side effect of CSR. Cassar and Meier (2017) focus on the effect on worker productivity and find CSR to decrease productivity if the firm's perceived intention is to benefit from engaging in CSR. Our experiment focuses on worker misbehavior rather than productivity, and our results are consistent with moral licensing as another potential dark side to the supply-side effects of CSR. Indeed, our findings have several implications for managers and employers. First, our results suggest that in an environment where worker misbehavior can be particularly costly to the employer, the organization should be careful to weigh the benefits of CSR against its costs because of potential moral-licensing effects. Second, we find evidence that the manner in which CSR is communicated to workers plays an important role in the extent to which it leads to moral licensing. Specifically, we provide suggestive evidence that CSR is especially harmful when it is framed as a benevolent act on behalf of workers.

Our study also makes important contributions to the recently growing literature on moral licensing. A large body of experimental evidence suggests that reflecting on prior good behavior (Sachdeva et al. 2009; Mazar and Zhong 2010; Merritt et al. 2010; Jordan et al. 2011; Conway and Peetz 2012; Effron et al. 2012, 2013; Clot et al. 2014) and planning to engage in future good acts (Urbszat et al. 2002, Cascio and Plant 2015) can license fewer moral, more antisocial, and more self-indulgent choices. Moral licensing has been documented in a variety of domains, such as consumer behavior (Mazar and Zhong 2010; Kouchaki and Jami 2018), racial bias (Effron et al. 2012, Cascio and Plant 2015), diet and health-related decisions (Urbszat et al. 2002, Chiou et al. 2011, Effron et al. 2013), charitable giving (Sachdeva et al. 2009, Conway and Peetz 2012), and energy conservation (Tiefenbeck et al. 2013). A closely related study in this literature is by Kouchaki and Jami (2018), who explore the moral-licensing effect of CSR on consumer behavior. Using a series of experiments, the authors find that when the CSR message praises and credits the consumer as opposed to the company for the good deed, it is more likely to result in self-interested and self-indulgent choices by the consumers. Our study is different from that of Kouchaki and Jami (2018) in two major ways. First, Kouchaki and Jami (2018) explore the moral-licensing effect of CSR on *consumers*, whereas our experiment focuses on the effect on *workers*. Second, the moral-licensing effect documented by Kouchaki and Jami (2018) is not harmful to the company that engages in CSR and does not affect the firm's profit. By contrast, our study documents

a moral-licensing effect of CSR that directly affects a firm's profitability by increasing worker misbehavior on the job. In other words, whereas Kouchaki and Jami (2018) deal with a rather normative question of "should we praise people for their good deeds in the hope that they do even more good deeds?" our paper deals with a positive firm-level question of "how does CSR affect a company's profitability through influencing worker misbehavior?" Therefore, our findings are directly relevant to a purely profit-maximizing employer. Although previous evidence on moral licensing is largely from the laboratory, our study is one of the few that documents moral licensing in the field, where subjects are not aware that they are taking part in the study.[3] The scale of our field experiment is another feature that distinguishes our study from previously documented evidence on moral licensing. Compared with previous studies in the moral-licensing literature with an average sample size of 81.3 participants per study (Blanken et al. 2015), our experiment with a sample size of over 1,500 subjects is one of the largest-scale studies in which moral licensing has been documented.

The remainder of our paper is organized as follows. Section 2 discusses the channels through which CSR affects worker misbehavior. Section 3 describes our experimental design, and Section 4 presents our findings. Section 5 discusses our results and concludes.

2. Gift Exchange vs. Moral Licensing

Consider a worker who commits effort for a firm that pays a wage of W and invests in CSR. While working, the worker might also engage in various forms of misconduct that are harmful to the employer, such as stealing from or lying to the employer.[4] We assume that the worker cares about his or her self-image. Unethical behavior lowers the worker's self-image, and CSR, especially if the behavior is perceived as a prosocial act by the worker, boosts it. Alternatively, we could think of self-image as an inverse function of guilt. The worker's feeling of guilt increases with misbehavior and decreases with CSR. Therefore, CSR may generate a moral-licensing effect and *increase* misbehavior by boosting the worker's self-image and licensing the unethical act.

The worker also potentially exhibits gift-exchange motives toward the employer and cheats less if he or she perceives the employer to be "kind." Both the wage and CSR can increase the worker's perception of the employer's level of kindness, and an employer is perceived to be kinder if it pays a higher wage or spends more on CSR. Previous research has shown that financial and nonfinancial incentives can induce gift exchange in workers (e.g., Kosfeld and Neckermann 2011, Kube et al. 2012, Ockenfels et al. 2015, Bradler et al. 2016, Flory et al. 2016). Therefore, CSR can *decrease* misbehavior by generating a gift-exchange effect. The two channels of gift exchange and moral licensing operate in opposing directions, leaving the overall effect of CSR on worker misbehavior ambiguous. In the next section, we present the design of our experiment, which provides empirical evidence on the relative importance of the two channels.

3. Experimental Design

We conducted our empirical study as a natural field experiment (Harrison and List 2004), using workers from Amazon's Mechanical Turk (MTurk). MTurk is an online labor market platform where businesses and individuals can post tasks, and workers can complete those tasks for payment. We were the employer and recruited workers for a short task to be performed on our website. The task was designed such that it provided opportunities for workers to cheat without losing payment or risking their reputation on MTurk. We invited interested workers to our website, where we offered them a short contract. At this point, to uncover how CSR affects cheating on the job, we randomized workers into treatments across which we varied the CSR incentive.

3.1. Recruitment on MTurk

The MTurk environment provides an appropriate platform to study worker misbehavior. The relative anonymity and the remote nature of the job make it possible to design a task with potential for cheating. We recruited workers by posting a human intelligence task (HIT) on MTurk and invited interested workers to our website, where they learned more about the task and payment and worked on the task. The HIT provided a general description of the task and instructions on how to proceed. Interested workers were instructed to click on a link, which took them to our website, where a unique identification (ID) code was generated for each worker. Workers were then required to submit their code through MTurk. We later used these ID codes to link workers on our website to their MTurk account, through which we made payments. The codes also allowed us to ensure that a worker was not exposed to multiple treatments.

3.2. The Contract

Upon landing on our website, workers were randomized into one of our treatments and were provided with a "contract" with detailed information about the task and payment. The contract indicated that 10% of the total wage would be paid upfront, upon accepting the contract, and the remaining 90% would be paid to workers after they completed the task.

After reading this information, workers decided whether to accept the contract. They were explicitly asked to accept the contract only if they intended to complete the task. All workers who accepted our contract were immediately paid the upfront money and were sent a message through MTurk notifying them about the upfront payment. Accepting the contract (and consequently receiving the upfront payment) without actually finishing the task serves as one of our measures of misbehavior. The contract required workers to complete the task within 2 hours of acceptance.

3.3. Treatment Groups

Workers were randomized into one of the three treatment groups (Table 1) and received a wage of \$1.20 for completing the task.[5] In addition to the wage, workers in treatments $T_{FirmMsg}$ and $T_{WorkerMsg}$ also received a social incentive (CSR). Our CSR initiative was in the form of donating cash to a prespecified nonprofit organization. Donation to charities and nonprofit organizations is a form of CSR that is frequently advertised and practiced by companies.[6] Companies often report their charity donations as the equivalent of a percentage of their profit. However, because we did not make profits from our experiment, in order to avoid deception, we framed our donation as the equivalent of a percentage of our total wage bill. The language in our message was carefully designed to make it clear that the donation was not going to come at a cost to the workers and would not reduce the wage promised on the contract. We made the donation payment after completion of the experiment.

As previously discussed, if the moral-licensing effect is strong enough, CSR can increase cheating. Moral licensing operates via concerns for self-image. If workers view themselves as a part of the prosocial act in which their employer engages, CSR can improve workers' self-image and license fewer subsequent ethical choices. The ability to take credit for the prosocial act of the employer is an important determinant of whether CSR influences the worker's self-image and therefore is key to whether it leads to moral licensing.

The manner in which CSR is framed and communicated to employees can influence the extent to which workers' self-image is exploited by CSR. To shed light on the role of moral licensing in the relationship between CSR and misbehavior, we manipulated the appeal of CSR to workers' self-image by changing the framing of our CSR message across treatments. While for some workers ($T_{FirmMsg}$) we simply framed CSR as a "donation to the charity," for others ($T_{WorkerMsg}$) we framed it as a "donation to the charity on behalf of the workers."[7] If CSR is to increase cheating through moral licensing, we should expect the later framing to generate a stronger effect. Kouchaki and Jami (2018) studied the moral-licensing effect of CSR on *consumer* behavior and found that CSR was more likely to generate a moral-licensing effect when the CSR message praised the consumer as opposed to the firm for the good deed. In line with this finding, we expect that if CSR is to generate a moral-licensing effect in workers' choices, the message that frames CSR as an act on behalf of the worker will cause a stronger effect. The following paragraph presents the CSR message for $T_{FirmMsg}$:

> Our firm is committed to give back in meaningful ways. We are passionate about encouraging education for the next generation. We do our part by donating money to influential non-profit organizations that support education for children from low socioeconomic backgrounds. In keeping with our philanthropic mission, we donate the equivalent of 5% of our wage bill in cash on behalf of all workers who help us with this project to UNICEF Education Programs. UNICEF works tirelessly to ensure that every child—regardless of gender, ethnicity or circumstances—has access to a quality education. You may find out more about UNICEF Education Programs at: UNICEF.

In $T_{WorkerMsg}$, we replaced "we donate the equivalent of 5% of our wage bill in cash to . . ." with "we donate the equivalent of 5% of our wage bill in cash on behalf of all workers who help us with this project to. . . ." Appendix C presents the information provided to workers at each stage of the experiment.

The information about CSR was given to workers *after* they accepted the contract and before they started to work on the task. Thus, our treatments remove the possibility of sorting, because workers in the CSR treatments could not select themselves into working for the CSR-type employer.[8] We also conducted two additional treatments, which together with $T_{FirmMsg}$ allow us to explore the trade-off between allocating employer's resources between CSR and wage. Across these three treatments, we held the total employers' budget per worker (wage and CSR donation) fixed and varied the share of the budget spent on CSR. Appendix D discusses these treatments in more detail.

Table 1. Treatments

	$T_{Baseline}$ (1)	$T_{FirmMsg}$ (2)	$T_{WorkerMsg}$ (3)
Wage	\$1.20	\$1.20	\$1.20
CSR as % of wage bill	0%	5%	5%
CSR framing	—	Firm	Worker

3.4. The Task

We used a transcription task, which is one of the most common tasks posted by employers on MTurk. We provided each worker with a batch of 10 images of

short German texts, scanned from old books, to transcribe. On average, each image was composed of around 30 words or 183 characters. To make it harder for workers to use transcription software, which automatically transcribes images of texts, before scanning each text, we crinkled the papers on which the image was printed to slightly deform the shape of each character. We used German texts to make the task more tedious and less enjoyable for our predominantly non-German workers. Workers who submitted all 10 images received a full wage of \$1.20.

Before starting to work on any given image, workers were required to report whether the image was legible or not. One of the 10 images was made too blurry to be readable, and the remaining images were easily readable.[9] Legibility of images was assessed by two undergraduate research assistants who, prior to the experiment, transcribed images into text. All workers transcribed the same 10 images, but the order in which the images appeared differed across workers and was determined randomly. If an image was reported as unreadable, the worker skipped that image and moved on to the next. The instructions made it clear that it was possible for some images to be illegible and that reporting those as unreadable was acceptable and did not lower the payment. In this manner, offering fixed payments, regardless of the number of images actually transcribed, provided opportunities for workers to misreport perfectly readable images as unreadable and avoid transcribing them.

We consider such behavior as our main measure of misbehavior on the job. Thus, we analyze and compare two measures of misbehavior across treatments: *type 1 cheating*, which refers to misreporting readable images as unreadable, a form of shirking, and *type 2 cheating*, which refers to accepting the contract (and receiving the upfront payment) without completing the job—essentially "taking the money and running."

Type 1 cheating is more harmful to the employer than type 2 cheating for several reasons. First, a worker who commits type 2 cheating would only receive the 10% upfront payment (because, by definition, he or she has not completed the task), whereas a type 1 cheater, who lies about the legibility of images, receives the full payment provided the he or she submits all images. Moreover, if a worker does not transcribe a readable image simply by not submitting the image (type 2), the employer can still receive the output by hiring another worker to transcribe that image. However, in the absence of monitoring, if a worker skips transcribing an image by misreporting it as unreadable (type 1), the employer would not receive the output because he or she has received misleading information about the image. Thus, type 1 cheating is more likely to be perceived as immoral and induce guilt because it harms the employer more, not only through higher wage costs but also by deceiving the employer and providing him or her with corrupt information.

3.5. Payment

Employers on MTurk have two means of making payments to workers. Each employer is required to set a "reward" for each assignment he or she posts on MTurk, which is a prespecified level of payment that will be paid to workers if (and as soon as) the employer approves the submitted HIT. In addition to the reward, an employer can make extra payments in the form of a "bonus." We set the reward for our assignments to \$0.10, which was paid to all workers who submitted the HIT before they were randomized into treatments and were provided with the contract. The wage that was specified in the contract was paid to the workers who accepted the contract via bonus payments. Appendix C presents the information provided to workers at each stage of the experiment.

3.6. Reputational Concerns

Once a worker completes a task on MTurk, the employer may review the work and either approve or reject the submission. Future employers may restrict their tasks to workers whose lifetime approval rate is above a certain desired level. The approval rate thus may serve as a signal of a worker's quality and can be used by employers as a requirement that potential workers have to meet. In order to eliminate potential reputational concerns and increase the baseline cheating rate, we automatically approved all the HITs that were submitted by workers before offering them the contract and notified workers that their submission was approved via automated messages.

In summary, the timeline of the experiment was as follows: (1) recruitment on MTurk (link to our website), (2) contract offered on our website, (3) decision to accept or reject the contract, (4) payment of 10% of the wage if the contract was accepted, (5) task instructions (and CSR message in treatments $T_{FirmMsg}$ and $T_{WorkerMsg}$), (6) transcription task, and (7) payment of the remaining wage.

4. Experimental Results

We launched the experiment in October 2016 and collected data for 12 weeks until we gathered at least 500 subjects in each treatment cell who accepted the contract.[10] In total, 1,590 workers landed on our website and were randomized into one of the following three treatments. Of these workers, 1,514 accepted the contract, and on average, they spent between 15 and 20 minutes on our website.

To shed light on how CSR affects cheating on the job, we start by comparing the propensity to accept the contract across treatments. Note that because, up to the point of making the decision on whether to accept the contract, all workers are exposed to the exact same information, we do not expect differences in acceptance rates across treatments. We find the acceptance rates to be very high and to lie between 94.59% and 96.01% (Table 2) and, as expected, to not significantly differ across treatments (p-values from a two-sample test of proportions between pairs of treatments: $p_{1-2} = 0.71$, $p_{1-3} = 0.45$, $p_{2-3} = 0.27$). Next, we compare cheating behaviors.

4.1. Propensity to Cheat

Does CSR affect the propensity to cheat? To answer this question, we compare the share of cheaters in the baseline with the share in the CSR treatments. We refer to a worker as *cheater* if he or she commits either type 1 or type 2 cheating at least once, that is, if he or she either misreports a readable image as unreadable or accepts the contract and receives the upfront payment without completing the job.

As shown in Table 3, we find that the mention of CSR increases the propensity to cheat by about 24%. The share of workers who cheat increases from 0.209 (standard.error (SE) = 0.018) in the baseline to 0.259 (SE = 0.014) in the two CSR treatments ($p = 0.034$).[11]

Consistent with moral licensing, framing CSR as a benevolent act *on behalf of the workers* leads to a weakly larger increase in the propensity to cheat than simply framing it as an employer's donation. The share of cheaters increases from 0.209 in the baseline to 0.247 (SE = 0.019) in $T_{FirmMsg}$ ($p = 0.157$), an 18% increase. This is considerably smaller than the 30% increase we find when comparing the baseline with $T_{WorkerMsg}$. This 30% difference is statistically significant at the $p < 0.05$ level. Yet we note that this should be considered weakly informative of the underlying mechanism because the difference in the share of cheaters across the two CSR frames is not large enough to be statistically significant at conventional levels ($p = 0.37$). Appendix A shows that the documented increase in the propensity to cheat with CSR is robust to restricting our analysis to workers who completed the task and thus received full payments.

Table 2. Acceptance Rates

Number of workers	$T_{Baseline}$ (1)	$T_{FirmMsg}$ (2)	$T_{WorkerMsg}$ (3)
N_{Total}	528	536	526
$N_{Accepted}$	502	507	505
Acceptance rate	95.08%	94.59%	96.01%

Notes. The first row presents the number of workers who landed on our website and were randomized into treatments. The second row presents the number of workers who accepted the contract they were offered. And the last row presents the rate of accepting the contract across treatments.

Table 3. Propensity to Cheat

Measure of cheating	$T_{Baseline}$ (1)	$T_{FirmMsg}$ (2)	$T_{WorkerMsg}$ (3)
$S_{Cheater}$	0.209	0.247	0.271
	(0.018)	(0.019)	(0.020)
S_{Type1}	0.056	0.073	0.101
	(0.010)	(0.012)	(0.013)
S_{Type2}	0.159	0.181	0.174
	(0.016)	(0.017)	(0.016)
Obs.	502	507	505

Notes. The first row presents the share of workers who commit type 1 or type 2 cheating, among those who accepted the contract. The second row presents the share of type 1 cheaters (i.e., workers who misreported at least one readable image as illegible), and the third row presents the share of type 2 cheaters (i.e., workers who accepted the contract without completing the task). Standard errors are reported in parentheses.

Next, we explore the effect of CSR on the propensity to engage in type 1 and type 2 cheating separately. The share of type 1 cheaters (i.e., workers who misreport one or more readable image as illegible) significantly increases by over 55% from 0.056 (SE = 0.010) in the baseline to 0.087 (SE = 0.009) in the two CSR treatments ($p = 0.032$). Exploring the effect of each framing separately, we find CSR to increase the share of type 1 cheaters by over 30% to 0.073 (SE = 0.012) in $T_{FirmMsg}$ and further to 0.101 (SE = 0.013) in $T_{WorkerMsg}$. Although the increase from $T_{Baseline}$ to $T_{FirmMsg}$ is insignificant at conventional levels ($p = 0.266$), the increase from baseline to $T_{WorkerMsg}$ is large and statistically significant ($p = 0.008$). Interestingly, framing CSR as a prosocial act on behalf of workers nearly doubles the number of type 1 cheaters. The difference in the share of type 1 cheaters between the two framings of the CSR message is marginally insignificant ($p = 0.11$). Although the share of type 2 cheaters also increases with CSR from 0.159 (SE = 0.016) in $T_{Baseline}$ to 0.178 (SE = 0.012) in $T_{FirmMsg}$ and $T_{WorkerMsg}$, the increase is not significant at conventional levels ($p = 0.367$).

Result 1. *CSR significantly increases the propensity to cheat. Our data provide suggestive evidence that the mechanism through which CSR increases cheating is moral licensing.*

4.2. Intensity of Cheating

Our extensive margin analysis suggested that CSR turns more noncheaters into cheaters. We now turn to the intensity of cheating and investigate whether CSR also increases the intensity of cheating by a worker. To answer this question, we compare the average level of cheating by a worker across treatments, measured by

the number of images that a worker either misreported as unreadable (type 1) or did not submit (type 2).

As summarized on Table 4, we find CSR to increase the average intensity of cheating per worker by about 11%: From 1.540 (SE = 0.150) images in the baseline to 1.716 (SE = 0.107) in the CSR treatments ($p = 0.06$).[12] Exploring the effect of each type of CSR messaging separately, we find the intensity of cheating to increase from the baseline by about 13% to 1.740 in $T_{FirmMsg}$ ($p = 0.171$) and by about 10% to 1.693 in $T_{WorkerMsg}$ ($p = 0.061$). Note that while the increase in cheating in $T_{FirmMsg}$ is larger in magnitude, the increase is only (marginally) significant in $T_{WorkerMsg}$.

Focusing on the effect of CSR on the intensity of type 1 cheating, we find a 43% increase from 0.145 (SE = 0.038) to 0.207 (SE = 0.028) images per worker from the baseline to the CSR treatments ($p = 0.031$). Examining the effect of each type of messaging separately, we find the intensity of type 1 cheating to increase from 0.145 in the baseline to 0.162 (SE = 0.034) in $T_{FirmMsg}$ ($p = 0.267$) and to 0.251 (SE = 0.045) in $T_{WorkerMsg}$ ($p = 0.007$). Note that while this higher level of type 1 cheating in $T_{WorkerMsg}$ compared with $T_{FirmMsg}$ is in line with moral licensing because the difference across the two framings is not significant at conventional levels, we should take it as only *suggestive* evidence for moral licensing. Although the intensity of type 2 cheating also increases with CSR, from 1.394 in the baseline to 1.578 (SE = 0.154) and 1.441 (SE = 0.145) in $T_{FirmMsg}$ and $T_{WorkerMsg}$, these increases are not significant. Therefore, the increase in the intensity is mainly driven by type 1 cheating. As we discuss in Appendix A, our conclusions are robust to restricting the sample to workers who finished the task and received full payment.[13] Result 2 summarizes our findings on the effect of CSR on the intensity of cheating by workers.

Table 4. The Intensity of Cheating

Measure of cheating	$T_{Baseline}$ (1)	$T_{FirmMsg}$ (2)	$T_{WorkerMsg}$ (3)
C_{Total}	1.540	1.740	1.693
	(0.150)	(0.156)	(0.148)
C_{Type1}	0.145	0.162	0.251
	(0.038)	(0.034)	(0.045)
C_{Type2}	1.394	1.578	1.441
	(0.147)	(0.154)	(0.145)
Obs.	502	507	505

Notes. The first row presents the average number of images a worker cheated on. The second row presents the average number of readable images that a worker misreports as unreadable (type 1). And the third row shows the average number of images a worker skipped transcribing by not finishing the task. Note that these averages reflect data from both cheaters and noncheaters. Standard errors are reported in parentheses.

Result 2. *CSR increases the average intensity of cheating. Our data provide suggestive evidence that the mechanism through which CSR increases cheating is moral licensing.*

5. Discussion

One of the unprecedented global economic movements among firms in the past 50 years is their newfound appreciation for their own social responsibility. Indeed, it is difficult to find more than a handful of leading firms that do not have a CSR program, and most have entire departments devoted to CSR. Although this change in landscape is not debatable, what is noteworthy is the dearth of empirical evidence on the consequences of such efforts.

We make a contribution in this area using a natural field experiment with over 1,500 workers to explore the relationship between CSR and employee misbehavior. Our empirical analyses suggest that CSR can actually increase worker misbehavior. Workers who received our CSR message were more likely to cheat and cheated more intensively than those in the baseline group. The overall data patterns are consistent with a model of moral licensing. We find the propensity for misbehavior to be the highest when we frame CSR as a prosocial act on behalf of workers compared with when the CSR is on behalf of firms. That is, when we communicated CSR as a benevolent act from the employer, CSR does not significantly impact misbehavior compared with the control group. This result is consistent with the findings of Kouchaki and Jami (2018), who document a higher level of moral licensing when consumers are exposed to a CSR message that praises the consumers compared with a message that praises the company for CSR.

Although we view our results as one of the first tests of moral licensing in a natural field experiment, another factor that distinguishes our work from previous evidence on moral licensing is the sample size. In a meta-analysis of 91 studies in moral licensing, Blanken et al. (2015) find the sample size in these studies to be relatively small, with an average of only 81.3 participants per study. With a sample size of over 1,500 subjects, our experiment is one of the largest-scale ones in which moral licensing has been documented.[14]

Beyond these advances, our findings have important implications for managers and practitioners on the use of CSR as a tool for motivating workers. Previous studies that explored the supply-side effect of CSR have focused on the effect on productivity. Our results suggest that in an environment where worker misbehavior is costly to the employer, CSR may have counterproductive effects by increasing misbehavior though moral licensing. Of course, the degree to which CSR generates moral licensing also depends on how it is communicated to workers.

Because moral licensing works by exploiting image concerns, we believe that framing CSR in a way that triggers worker self-image is more likely to increase misbehavior on the job.

Moreover, our findings provide important insights with regard to the decision that an employer faces in allocating his or her limited financial resources between private incentives (wage) and social incentives in the form of CSR. Combing the previously documented effect of wage in reducing worker misbehavior (Ockenfels et al. 2015, Flory et al. 2016) with our finding that CSR increases misbehavior, one can infer that an employer who seeks to minimize worker misconduct would be better off by spending his or her financial resources in the form of private incentives (wage). This is indeed what our complementary treatments reveal. As we substitute wage with CSR, the share of workers who misbehave and the average intensity of cheating per worker increase. Appendix D presents these findings in more detail.

We believe that our findings are likely to be generalizable to the dimensions of a worker's counterproductive behavior that damages his or her self-image. For CSR to increase misbehavior via moral licensing, engaging in misbehavior should induce some form of guilt in the worker. In other words, an individual does not need a license to misbehave if that behavior does not provoke a feeling of guilt and does not damage his or her self-image. Therefore, we expect counterproductive workplace behaviors that are perceived to be unethical (such as stealing or lying) to be more likely to increase with CSR.

We recognize a number of limitations in the extent to which our findings are generalizable to other settings that we believe future research should seek to address. First, to evaluate whether CSR generates positive supply-side effects, a firm needs to take into account two major determinants: The selection effect and the treatment effect. The selection effect speaks to whether a firm who engages in CSR attracts better (more productive/ethical) workers, whereas the treatment effect refers to the impact of adding CSR incentives for an existing pool of workers. Our treatments shut down the selection effect by design, and our results reflect the pure treatment effect of CSR. Therefore, our findings will be most relevant for an employer with an existing pool of workers who is trying to determine whether to invest in CSR. It thus remains an open question whether, in an environment with the possibility of selection, firms that engage in CSR would attract more ethical workers, or they would attract less honest workers who use the moral-licensing effect of CSR to justify their misconduct.

Second, our experiment was conducted as a one-shot interaction with contingent workers who have limited connection to their employer. The temporary nature of the relationship of employers with contingent workers surely affects the extent to which workers identify with the organization for which they work and, consequently, with the employer's CSR identity. Thus, the extent to which our findings generalize in longer-term employment relationships remains to be explored in future research.

Finally, although our evidence is consistent with moral licensing, because the difference in cheating across the two framings of CSR is not large enough to be significant at the conventional levels, we can only take it as *suggestive* evidence for moral licensing. Indeed, CSR may induce negative supply-side effects via other channels as well. For example, CSR might induce negative reciprocity because of the perceived intention of the employer to use CSR instrumentally. As discussed in Cassar and Meier (2017), the efficacy of social incentives in inducing positive reciprocity might depend on the perceived intention of the employer, and if perceived as used instrumentally to maximize profits, prosocial incentives may backfire. Alternatively, the CSR message might signal an employer's benevolence. CSR can increase misbehavior if workers believe that the employer's benevolence is negatively correlated with possible punishments in case misbehavior is detected. Finally, one could argue that spending resources on social causes might simply upset workers (and induce negative reciprocity) because it might imply that fewer resources will be allocated for wage payment. We do not believe such a concern plays a major role in our setting because (1) a firm that invests in CSR can keep wage fixed and have reduced profits, and (2) our instructions mitigated such concerns by clearly stating a worker's payoff in the contract. Although we do not believe that such effects play a key role in generating our results, we cannot completely rule out their presence in our experiment.

Acknowledgments

The authors thank Joseph Seidel for his excellent research assistance.

Appendix A. Robustness to Sample Restriction

In this appendix, we restrict our sample to only those workers who completed the task and thus received full payment and show that our findings are robust to this restriction. Note that, by definition, those who completed the task could have only committed type 1 cheating, that is, misreporting a readable image as unreadable. As Table A.1 presents, the propensity to cheat increases with CSR for this subsample of workers. The share of cheaters among those who submitted all 10 images significantly increases from 0.059 (SE = 0.012) in $T_{Baseline}$ to 0.099 (SE = 0.010) in the CSR treatments ($p = 0.019$), and the increase is the largest when CSR is framed as a benevolent act on behalf of workers.

Likewise, the intensity of cheating among workers who completed the task increases with CSR. CSR significantly

Table A.1. Cheating Among Workers Who Completed the Task

Measure of cheating	$T_{Baseline}$ (1)	$T_{FirmMsg}$ (2)	$T_{WorkerMsg}$ (3)
$S_{Cheater}$	0.059	0.080	0.118
	(0.012)	(0.013)	(0.016)
C_{Total}	0.159	0.181	0.290
	(0.044)	(0.041)	(0.053)
Obs.	422	415	417

Notes. The first row presents the share of cheaters among those who completed the task. The second row presents the average intensity of cheating by workers who completed the task. The last row shows the number of workers who submitted all 10 images and therefore received full payment. Standard errors are reported in parentheses.

increases the intensity of overall cheating by about 48% from 0.159 in the baseline to 0.236 (SE = 0.033) in the CSR treatments ($p = 0.018$). Decomposing the effect of each type of framing, we find a 14% increase in cheating from the baseline to 0.181 (SE = 0.041) in $T_{FirmMsg}$ ($p = 0.249$) and an 82% increase from the baseline to 0.290 (SE = 0.053) in $T_{WorkerMsg}$ ($p = 0.003$).

Appendix B. Quality of Transcription

Our findings suggest that CSR increases cheating: Workers who are exposed to the CSR messages are more likely to either lie about the legibility of images or receive the upfront payment but not complete the task. Here we explore whether CSR affects the quality of work for the images that workers *do* transcribe. We measure the quality of transcription using the Levenshtein distance between the texts that were transcribed by workers and the actual texts on the images. The *Levenshtein distance* is defined as the minimum number of single-character edits (i.e., insertions, deletions, or substitutions) required to go from one string of text to another. One can think of this measure as the number of errors/typos that a worker makes while transcribing. Therefore, a lower Levenshtein distance corresponds to a higher quality of transcription.

Table B.1. Quality of Transcription

Measure of quality	$T_{Baseline}$ (1)	$T_{FirmMsg}$ (2)	$T_{WorkerMsg}$ (3)
L_{All}	22.147	22.795	23.641
	(0.985)	(0.984)	(1.124)
Obs.	448	446	457
$L_{Noncheater}$	20.864	21.585	21.957
	(0.940)	(0.893)	(0.978)
Obs.	397	382	368
$L_{Cheater}$	42.414	34.853	37.455
	(6.197)	(3.658)	(4.253)
Obs.	51	64	89

Notes. The first row presents the average Levenstein distance across treatments. The second and third rows present the average Levenstein distance for noncheaters and for cheaters separately. Standard errors are reported in parentheses.

As illustrated on Table B.1, the average Levenstein score in the baseline is 22.147, which increases to 22.795 and 23.641 in $T_{FirmMsg}$ and $T_{WorkerMsg}$. None of these increases is significant. Comparing the quality of output produced by cheaters and noncheaters, one can see that cheaters' work is of a much lower quality across all treatments.[15] The number of errors per image made by a cheater is on average between 43% and 110% higher than the errors made by a noncheater across treatments.

Appendix C. Information on the HIT and Our Website

C.1. A Snapshot of the HIT on Amazon MTurk

Figure C.1 presents a snapshot of our HIT that we posted on MTurk to recruit workers.[16]

C.2. Information on the First Page

Your response to our MTurk HIT was just auto-approved and your $0.10 reward was paid on your Amazon Payment account. Please read the following information about the task and decide if you are willing to work for us for a few minutes and earn extra bonus!

Figure C.1. (Color online) A Snapshot of the Human Intelligence Task (HIT) on Amazon MTurk

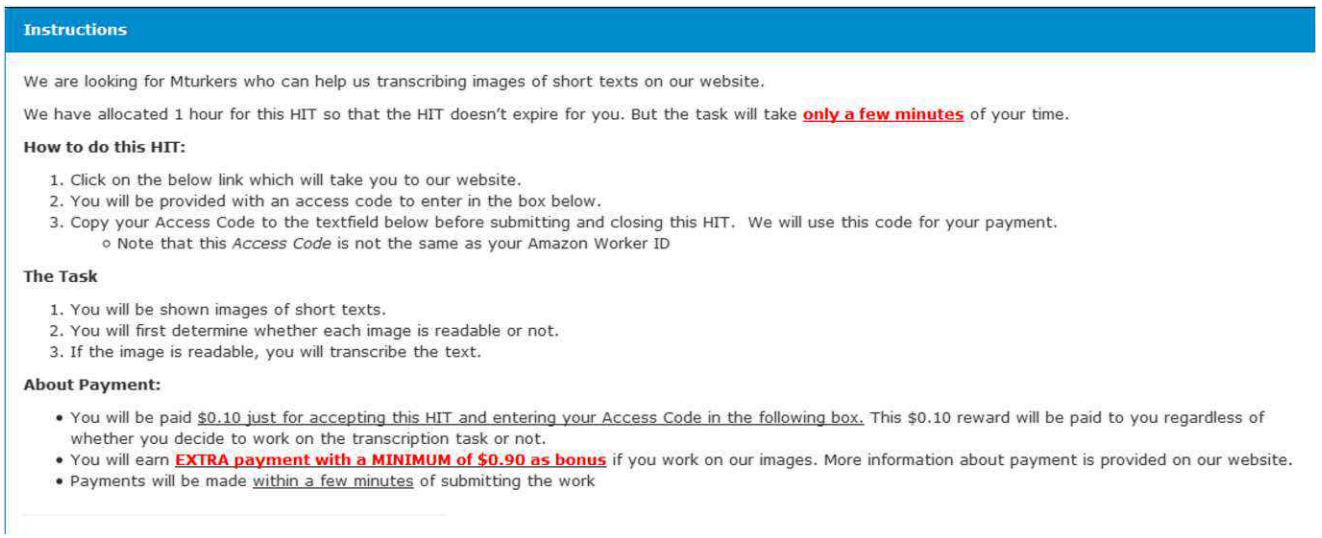
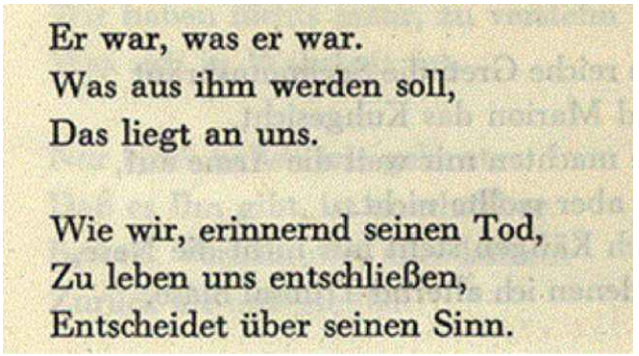


Figure C.2. (Color online) An Example of the Image, Which Was Included in the Experimental Instructions


Er war, was er war.
Was aus ihm werden soll,
Das liegt an uns.

Wie wir, erinnernd seinen Tod,
Zu leben uns entschließen,
Entscheidet über seinen Sinn.

This HIT requires you to transcribe short texts that have been scanned from German documents.

- You will receive a total of $1.20 bonus for working on 10 images (similar to the one you see below), which will be paid as a bonus on your Amazon Payment account.
- The task is short and will not take more than a few minutes.
- If you chose to work on 10 images, we will immediately pay 10% of the total bonus to you on your Amazon Payment account upfront, and the remaining 90% will be paid as soon as you complete the task (i.e., finish all 10 images)
- It is possible that some of the images are too blurry to be readable. Reporting those as unreadable is acceptable and will NOT reduce your bonus payment.
- You will need to finish the task within the next 2:00 hours in order to receive payments.

You now have the option to choose if you want to only earn your $0.10 reward for accepting the HIT and exit, or you would like to transcribe images for us and earn an extra bonus.

NOTE: 10% ($0.12) of the total bonus will be paid to you immediately after choosing the option to work on all 10 images, and the remaining 90% ($1.08) will be paid within a few minutes after you complete the task.

Please choose one of the following options:

□ I want to transcribe all 10 images.

□ I do not want to transcribe images.

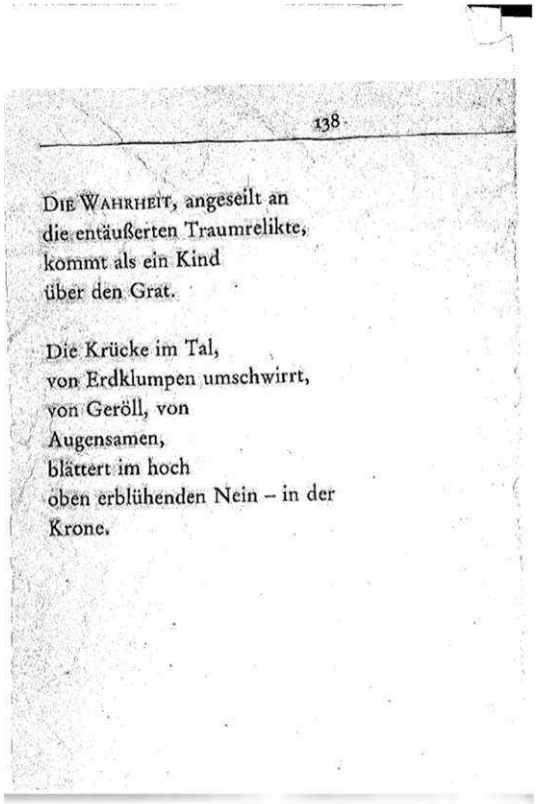
C.3. Information on the Second Page (Shown to Subjects After They Accepted the Contract)

Thank you for helping us with this transcription task. We just made an upfront payment of $0.12 as a bonus on your Amazon Payment account. You will receive the remaining bonus upon finishing the task.

A few words about our firm's philanthropic mission:

Our firm is committed to give back in meaningful ways. We are passionate about encouraging education for the next generation. We do our part by donating money to influential non-profit organizations that support education for children from low socioeconomic backgrounds. In keeping with our philanthropic mission, we donate the equivalent of 5% of our wage bill in cash to UNICEF Education Programs. UNICEF works tirelessly to ensure that every child—regardless of gender, ethnicity or

Figure C.3. Examples of Readable and Unreadable Images


138

Die Wahrheit, angeseilt an
die entäußerten Traumrelikte,
kommt als ein Kind
über den Grat.

Die Krücke im Tal,
von Erdklumpen umschwirrt,
von Geröll, von
Augensamen,
blättert im hoch
oben erblühenden Nein – in der
Krone.

circumstances—has access to a quality education. You may find out more about UNICEF Education Programs at UNICEF.[17]

Instructions.

1. You will be shown 10 images similar to the one seen on this page.
2. For each image, you will first need to determine whether it is readable or not.
3. It is possible that some of the images are too blurry to be readable. Reporting those as unreadable is acceptable and will NOT reduce your bonus payment.
4. If you report the image as readable, you will need to transcribe the text.
5. Use the following rules for non-standard characters:
   - Transcribe ä as ae, Ä as Ae.
   - Transcribe ö as oe, Ö as Oe.
   - Transcribe ü as ue, Ü as Ue.
   - Transcribe β as ss.
6. Do NOT transcribe page numbers.
7. Do NOT preserve the stanzas as you transcribe. Click to see an example.
8. If you cannot read some characters, replace them with an underscore.
9. Press "Submit" after you finished transcribing.

You will be able to review these instructions at any time.

We need this work to be completed quickly. Please complete the task in the next 2:00 hours in order to receive the remaining 90% of your payment.

C.4. Examples of Images

See Figure C.2 for an example of the image. See Figure C.3 for examples of readable and unreadable images.

Appendix D. Pay for Wage or Pay for CSR?

We conducted two additional treatments ($T^{HWage}_{Baseline}$ and $T^{LWage}_{FirmMsg}$) that together with $T_{FirmMsg}$ allow us to explore how misbehavior changes as an employer with a fixed budget substitutes wage with CSR. Across these treatments, we fixed our total expenditure on wage and CSR and reduced the share of the budget spent on wage as we moved from $T^{HWage}_{Baseline}$ to $T_{FirmMsg}$ and $T^{LWage}_{FirmMsg}$. In $T^{HWage}_{Baseline}$, the entire budget of \$1.26 was allocated to wage, whereas in $T_{FirmMsg}$ ($T^{LWage}_{FirmMsg}$), \$1.20 (\$0.90) of the budget per worker was spent on wage payments, and the remaining \$0.06 (\$0.30) was spend on CSR. Table D.1 summarizes these treatments.

Previous experiments have documented that, consistent with the prediction of the gift-exchange model, increasing wage reduces misbehavior on the job (Ockenfels et al. 2015, Flory et al. 2016). Putting this finding together with our finding that CSR increases cheating, one would expect misbehavior to increase as we allocate resources away from wage and into CSR. This is indeed what our empirical analysis reveals. Below we report our findings in terms of both the propensity to cheat and the average intensity of cheating.[18]

Table D.1. Wage vs. CSR Treatments

Treatment-specific variable	$T^{HWage}_{Baseline}$ (1)	$T_{FirmMsg}$ (2)	$T^{LWage}_{FirmMsg}$ (3)
Wage	\$1.26	\$1.20	\$0.90
CSR as % of wage bill	0	5%	40%
CSR (\$)	\$0	\$0.06	\$0.30
CSR message	—	Firm	Firm

D.1. Propensity to Cheat

As wage falls from \$1.26 to \$1.20 and \$0.90 and are substituted with CSR, we find that the propensity to cheat increases. The share of cheaters increases by about 25%, from 0.198 in $T^{HWage}_{Baseline}$ to 0.247 in $T_{FirmMsg}$ ($p = 0.066$) and by over 50% from $T^{HWage}_{Baseline}$ to 0.300 in $T^{LWage}_{FirmMsg}$ ($p = 0.000$). The increase in the share of cheaters from $T_{FirmMsg}$ to $T^{LWage}_{FirmMsg}$ is also significant ($p = 0.058$).

A similar pattern emerges when we explore the propensity to engage in type 1 and type 2 cheating separately. The share of type 1 cheaters increases from 0.062 in the $T^{HWage}_{Baseline}$ to 0.073 in $T_{FirmMsg}$ and to 0.097 in $T^{LWage}_{FirmMsg}$. Although the increase in the share of type 1 cheaters from $T^{HWage}_{Baseline}$ to $T^{LWage}_{FirmMsg}$ is significant ($p = 0.036$), the jumps between $T^{HWage}_{Baseline}$ and $T_{FirmMsg}$ ($p = 0.47$) and between $T_{FirmMsg}$ and $T^{LWage}_{FirmMsg}$ ($p = 0.167$) are not large enough to be significant at the conventional levels. The propensity to engage in type 2 cheating also increases as we substitute wage with CSR. The share of type 2 cheaters increases from 0.147 in the baseline to 0.181 and 0.220 in $T_{FirmMsg}$ and $T^{LWage}_{FirmMsg}$. (The p-value from $T^{HWage}_{Baseline}$ to $T_{FirmMsg}$ is 0.137, the p-value from $T_{FirmMsg}$ to $T^{LWage}_{FirmMsg}$ is 0.124, and the p-value from $T^{HWage}_{Baseline}$ to $T^{LWage}_{FirmMsg}$ is 0.003.) Table D.2 summarizes these findings.[19] We thus conclude that substituting wage with CSR increases cheating in the extensive margin. Reducing a worker's wage by 5% and substituting it by CSR increases the propensity to cheat by 25%. A further 25% decrease in wage and substitution with CSR further increases cheating by another 21%.

D.2. Intensity of Cheating

Our findings on the extensive margin are echoed in the intensity of cheating. As illustrated in Table D.3, the average number of images that a worker cheats on increases from 1.383 in $T^{HWage}_{Baseline}$ to 1.74 and 2.121 in $T_{FirmMsg}$ and $T^{LWage}_{FirmMsg}$. The increases in the intensity of cheating from $T^{HWage}_{Baseline}$ to $T_{FirmMsg}$ ($p = 0.054$), from $T_{FirmMsg}$ to $T^{LWage}_{FirmMsg}$ ($p = 0.066$), and from $T^{HWage}_{Baseline}$ to $T^{LWage}_{FirmMsg}$ ($p = 0.000$) are all significant.

The effects on the intensity of cheating are mainly driven by increases in type 2 cheating across treatments. Although the average intensity of type 1 cheating does not vary as we substitute wage with CSR across treatments, the average intensity of type 2 cheating increases from 1.200 images per worker in $T^{HWage}_{Baseline}$ to 1.578 in $T_{FirmMsg}$ and further up to 1.940 in $T^{LWage}_{FirmMsg}$ (p-values from $T^{HWage}_{Baseline}$ to $T_{FirmMsg} = 0.101$; from $T_{FirmMsg}$ to $T^{LWage}_{FirmMsg} = 0.129$; from $T^{HWage}_{Baseline}$ to $T^{LWage}_{FirmMsg} = 0.002$). We thus conclude that substituting wage with CSR increases the intensity of cheating. Reducing wage by 5% and substituting with CSR increases the average intensity of cheating by over 25%. A further 25% decrease in wage (and substitution with CSR) further increases the intensity of cheating by another 22%.[20]

Table D.2. Wage vs. CSR: Propensity to Cheat

Measure of cheating	$T^{HWage}_{Baseline}$ (1)	$T_{FirmMsg}$ (2)	$T^{LWage}_{FirmMsg}$ (3)
$S_{Cheater}$	0.198 (0.018)	0.247 (0.019)	0.300 (0.020)
S_{Type1}	0.062 (0.011)	0.073 (0.012)	0.097 (0.013)
S_{Type2}	0.147 (0.016)	0.181 (0.017)	0.220 (0.018)
Obs.	504	507	504

Notes. The first row presents the share of workers who commit type 1 or type 2 cheating among those who accepted the contract. The second row presents the share of type 1 cheaters (i.e., workers who misreported at least one readable image as illegible), and the third row presents the share of type 2 cheaters (i.e., workers who accepted the contract without completing the task). Standard errors are reported in parentheses.

Table D.3. Wage vs. CSR: The Intensity of Cheating

Measure of cheating	$T^{HWage}_{Baseline}$ (1)	$T_{FirmMsg}$ (2)	$T^{LWage}_{FirmMsg}$ (3)
C_{Total}	1.383 (0.140)	1.740 (0.156)	2.121 (0.168)
C_{Type1}	0.183 (0.044)	0.162 (0.034)	0.181 (0.033)
C_{Type2}	1.200 (0.135)	1.578 (0.154)	1.940 (0.168)
Obs.	504	507	504

Notes. The first row presents the average number of images a worker cheated on. The second row presents the average number of readable images that a worker misreports as unreadable (type 1). And the third row shows the average number of images a worker skipped transcribing by not finishing the task. Standard errors are reported in parentheses.

In summary, given our previous finding that worker misbehavior increases with CSR, and consistent with the findings from previous studies that misbehavior decreases with wage, we find cheating to increase as we substitute wage with CSR. This result suggests that given the moral-licensing effect that CSR can generate, an employer who is deciding on how to allocate financial resources to motivate workers through private incentive and CSR and seeks to minimize worker misbehavior would be better off by spending his or her entire budget on wage.

Endnotes

[1] See Kitzmueller and Shimshack (2012) for a review of the economic perspective on CSR.

[2] Accessed December 1, 2017, http://hayesinternational.com/news/annual-retail-theft-survey/.

[3] Tiefenbeck et al. (2013) and study 2 in Kouchaki and Jami (2018) are exceptions.

[4] We take the following definition of unethical misbehavior from Kish-Gephart et al. (2010, p. 2): "Behavior that violates widely accepted moral norms such as theft, sabotage, lying to customers and misrepresentation of financial reports.... Other negative, counterproductive or deviant workplace behavior such as lateness, are not included because they do not necessarily violate widely accepted moral norms."

[5] The average hourly earnings on MTurk lie between $1 and $5 (Ross et al. 2010, Paolacci et al. 2010, Berg 2016). Given that an average worker spent about 15–20 minutes on our website, the average hourly earnings of our workers fall within this estimated range.

[6] To name a few examples, in 2015 alone, Walmart, Wells Fargo, and Chevron made about $301 million, $281 million, and $225 million of cash donations to nonprofits (accessed December 1, 2017, http://fortune.com/2016/06/22/fortune-500-most-charitable-companies/).

[7] Engaging in CSR initiatives on behalf of employees (or simply framing CSR messages as such) is a common business practice. See, for example, the Breckinridge 2017 CSR Report (accessed December 1, 2017, https://www.breckinridge.com/corporate-sustainability-report-2017) or the Ansell 2017 CSR Report (accessed December 1, 2017, https://www.ansell.com/us/en/about-us/sustainability).

[8] We also conducted an additional treatment to explore the selection effects. However, the high rate of accepting the contract among workers prevented us from exploring the selection effect. We describe this treatment and discuss the findings in an earlier National Bureau of Economic Research working paper. The interested reader should refer to List and Momeni (2017).

[9] Examples of the blurry versus readable images can be found in Appendix C.4.

[10] We observed around a 21% cheating rate in our pilot. Our power calculation suggested that we would need around 420 subjects in each treatment cell in order to be able to detect a 40% change in cheating (power = 0.80; alpha = 0.05; List et al. 2011).

[11] The p-values reported in all extensive margin analyses are from a two-sample test of proportions against the null hypothesis of equal means.

[12] All the p-values reported in the analysis of the intensity of cheating are from a two-sample Wilcoxon rank-sum (Mann-Whitney) test against the null hypothesis of equal means.

[13] In Appendix B we explore the effect on the quality of output and show that CSR does not affect the quality of data entry.

[14] Our study is also related to the body of literature that documents positive (e.g., Brandts and Cooper 2006, Cason and Gangadharan 2013) and negative (e.g., Angelucci et al. 2016, Banerjee 2016) behavioral spillover effects.

[15] The difference in work quality by cheaters and noncheaters is significant at the $p < 0.001$ level in all treatments.

[16] Note that the $0.10 payment mentioned in the HIT was paid to all workers who visited our website (in the form of an MTurk "reward" regardless of whether they accepted our contract). This $0.10 was paid on top of (and is different from) the 10% upfront payment that was paid to workers who accepted the contract.

[17] The CSR message that was used in treatment $T_{WorkerMsg}$: "Our firm is committed to give back in meaningful ways. We are passionate about encouraging education for the next generation. We do our part by donating money to influential non-profit organizations that support education for children from low socioeconomic backgrounds. In keeping with our philanthropic mission, we donate the equivalent of 5% of our wage bill in cash on behalf of all workers who help us with this project to UNICEF Education Programs. UNICEF works tirelessly to ensure that every child regardless of gender, ethnicity or circumstances has access to a quality education. You may find out more about UNICEF Education Programs at UNICEF."

[18] Although the propensity to accept the contract slightly decreases as the wage offer goes down (from 94.92% in $T^{HWage}_{Baseline}$ to 94.59% in $T_{FirmMsg}$ and further down to 93.33% in $T^{LWage}_{FirmMsg}$), these declines are not significant (p-values from a two-sample test of proportions between pairs of treatments across which we varied wage—Table D.1: $p_{1-2} = 0.81; p_{1-3} = 0.27; p_{2-3} = 0.39$).

[19] The conclusions are robust to restricting our sample to workers who completed the task and received full payment.

[20] We do not find CSR to affect the quality of transcriptions when we substitute wage with CSR.

References

Angelucci M, Prina S, Royer H, Samek A (2016) When incentives backfire: Spillover effects in food choice. Discussion Papers, Research Unit: Market Behavior SP II:2016-205.

Association of Certified Fraud Examiners (2016) *Report to the Nations: Occupational Fraud and Abuse* (Association of Certified Fraud Examiners, Austin, TX).

Backhaus KB, Stone BA, Heiner K (2002) Exploring the relationship between corporate social performance and employer attractiveness. *Bus. Soc.* 41(3):292–318.

Bagnoli M, Watts SG (2003) Selling to socially responsible consumers: Competition and the private provision of public goods. *J. Econom. Management Strategy* 12(3):419–445.

Banerjee R (2016) Corruption, norm violation and decay in social capital. *J. Public Econom.* 137(C):14–27.

Bénabou R, Tirole J (2006) Incentives and prosocial behavior. *Amer. Econom. Rev.* 96(5):1652–1678.

Bénabou R, Tirole J (2010) Individual and corporate social responsibility. *Economica* 77(305):1–19.

Bénabou R, Tirole J (2011) Identity, morals, and taboos: Beliefs as assets. *Quart. J. Econom.* 126(2):805–855.

Berg J (2016) Income security in the on-demand economy: Findings and policy lessons from a survey of crowdworkers. *Comparative Labor Law Policy J.* 37(3):543–576.

Besley T, Ghatak M (2005) Competition and incentives with motivated agents. *Amer. Econom. Rev.* 95(3):616–636.

Besley T, Ghatak M (2007) Retailing public goods: The economics of corporate social responsibility. *J. Public Econom.* 91(9):1645–1663.

Blanken I, van de Ven N, Zeelenberg M (2015) A meta-analytic review of moral licensing. *Personality Soc. Psych. Bull.* 41(4):540–558.

Bradler C, Dur R, Neckermann S, Non A (2016) Employee recognition and performance: A field experiment. *Management Sci.* 62(11): 3085–3099.

Brandts J, Cooper DJ (2006) A change would do you good… an experimental study on how to overcome coordination failure in organizations. *Amer. Econom. Rev.* 96(3):669–693.

Brekke KA, Nyborg K (2008) Attracting responsible employees: Green production as labor market screening. *Resource Energy Econom.* 30(4):509–526.

Burbano VC (2019) Getting gig workers to do more by doing good: Field experimental evidence from online platform labor marketplaces. *Organ. Environ*, ePub ahead of print May 8, https://doi.org/10.1177/1086026619846455.

Burbano VC (2016b) Social responsibility messages and worker wage requirements: Field experimental evidence from online labor marketplaces. *Organ. Sci.* 27(4):1010–1028.

Cascio J, Plant EA (2015) Prospective moral licensing: Does anticipating doing good later allow you to be bad now? *J. Experiment. Soc. Psych.* 56(January):110–116.

Cason TN, Gangadharan L (2013). Cooperation spillovers and price competition in experimental markets. *Econom. Inquiry* 51(3): 1715–1730.

Cassar L, Meier S (2017) *Intentions for Doing Good Matter for Doing Well: The (Negative) Signaling Value of Prosocial Incentives (No. 11203)* (Institute of Labor Economics, Bonn, Germany).

Chiou WB, Wan CS, Wu WH, Lee KT (2011) A randomized experiment to examine unintended consequences of dietary supplement use among daily smokers: Taking supplements reduces self-regulation of smoking. *Addiction* 106(12):2221–2228.

Clot S, Grolleau G, Ibanez L (2014) Smug alert! Exploring self-licensing behavior in a cheating game. *Econom. Lett.* 123(2):191–194.

Conway P, Peetz J (2012) When does feeling moral actually make you a better person? Conceptual abstraction moderates whether past moral deeds motivate consistency or compensatory behavior. *Personality Soc. Psych. Bull.* 38(7):907–919.

Du S, Bhattacharya CB, Sen S (2011) Corporate social responsibility and competitive advantage: Overcoming the trust barrier. *Management Sci.* 57(9):1528–1545.

Effron DA, Miller DT, Monin B (2012) Inventing racist roads not taken: The licensing effect of immoral counterfactual behaviors. *J. Personality Soc. Psych.* 103(6):916–932.

Effron DA, Monin B, Miller DT (2013) The unhealthy road not taken: Licensing indulgence by exaggerating counterfactual sins. *J. Experiment. Soc. Psych.* 49(3):573–578.

Elfenbein DW, Fisman R, Mcmanus B (2012) Charity as a substitute for reputation: Evidence from an online marketplace. *Rev. Econom. Stud.* 79(4):1441–1468.

Fisman R, Heal G, Nair VB (2008) A model of corporate philanthropy. Working paper, Columbia University, New York.

Flory JA, Leibbrandt A, List JA (2016) The effects of wage contracts on workplace misbehaviors: Evidence from a call center natural field experiment. NBER Working Paper No. w22342, National Bureau of Economic Research, Cambridge, MA.

Greening DW, Turban DB (2000) Corporate social performance as a competitive advantage in attracting a quality workforce. *Bus. Soc.* 39(3):254–280.

Gubler T, Larkin I, Pierce L (2018) Doing well by making well: The impact of corporate wellness programs on employee productivity. *Management Sci.* 64(11):4967–4987.

Harrison GW, List JA (2004) Field experiments. *J. Econom. Literature* 42(4):1009–1055.

Hedblom D, Hickman BR, List JA (2016) Toward an understanding of corporate social responsibility: Theory and field experimental evidence. Working paper, University of Chicago, Chicago.

Jordan J, Mullen E, Murnighan JK (2011) Striving for the moral self: The effects of recalling past moral actions on future moral behavior. *Personality Soc. Psych. Bull.* 37(5):701–713.

Kish-Gephart JJ, Harrison DA, Treviño LK (2010) Bad apples, bad cases, and bad barrels: Meta-analytic evidence about sources of unethical decisions at work. *J. Appl. Psych.* 95(1):1–31.

Kitzmueller M, Shimshack J (2012) Economic perspectives on corporate social responsibility. *J. Econom. Literature* 50(1):51–84.

Kosfeld M, Neckermann S (2011) Getting more work for nothing? Symbolic awards and worker performance. *Amer. Econom. J. Microeconom.* 3(3):86–99.

Kouchaki M, Jami A (2018) Everything we do, you do: The licensing effect of prosocial marketing messages on consumer behavior. *Management Sci.* 64(1):102–111.

KPMG (2015) *KPMG International Survey of Corporate Responsibility Reporting 2015* (KPMG, Amsterdam).

Kube S, Maréchal MA, Puppe C (2012) The currency of reciprocity - gift-exchange in the workplace. *Amer. Econom. Rev.* 102(4): 1644–1662.

Lii YS, Lee M (2012) Doing right leads to doing well: When the type of CSR and reputation interact to affect consumer evaluations of the firm. *J. Bus. Ethics* 105(1):69–81.

List JA, Momeni F (2017) When corporate social responsibility backfires: Theory and evidence from a natural field experiment. NBER Working Paper No. w24169, National Bureau of Economic Research, Cambridge, MA.

List JA, Sadoff S, Wagner M (2011) So you want to run an experiment, now what? Some simple rules of thumb for optimal experimental design. *Experiment. Econom.* 14(4):439–457.

Mazar N, Zhong CB (2010) Do green products make us better people? *Psych. Sci.* 21(4):494–498.

Merritt AC, Effron DA, Monin B (2010) Moral self-licensing: When being good frees us to be bad empirical demonstrations of moral self-licensing. *Soc. Personality Psych. Compass* 4(5): 344–357.

National Retail Federation (2015) National retail security survey 2015. Report, Department of Sociology and the Center for Studies in Criminology and Law, University of Florida, Gainesville, FL.

Ockenfels A, Sliwka D, Werner P (2015) Timing of kindness - Evidence from a field experiment. *J. Econom. Behav. Organ.* 111(March): 79–87.

Paolacci G, Chandler J, Ipeirotis PG (2010) Running experiments on Amazon Mechanical Turk. *Judgment Decision Making* 5(5): 411–419.

Ross J, Irani L, Silberman M, Zaldivar A, Tomlinson B (2010). Who are the crowdworkers?: Shifting demographics in Mechanical Turk. *CHI'10 Extended Abstracts on Human Factors in Computing Systems* (ACM, New York), 2863–2872.

Sachdeva S, Iliev R, Medin DL (2009) Sinning saints and saintly sinners: The paradox of moral self-regulation. *Psych. Sci.* 20(4):523–528.

Sen S, Bhattacharya CB, Korschun D (2006) The role of corporate social responsibility in strengthening multiple stakeholder relationships: A field experiment. *J. Acad. Marketing Sci.* 34(2):158–166.

Tiefenbeck V, Staake T, Roth K, Sachs O (2013) For better or for worse? Empirical evidence of moral licensing in a behavioral energy conservation campaign. *Energy Policy* 57(June):160–171.

Tonin M, Vlassopoulos M (2014) Corporate philanthropy and productivity: Evidence from an online real effort experiment. *Management Sci.* 61(8):1795–1811.

Turban DB, Greening DW (1997) Corporate social performance and organizational attractiveness to prospective employees. *Acad. Management J.* 40(3):658–672.

Urbszat D, Herman CP, Polivy J (2002) Eat, drink, and be merry, for tomorrow we diet: Effects of anticipated deprivation on food intake in restrained and unrestrained eaters. *J. Abnormal Psych.* 111(2):396–401.