Journal of Business Ethics (2022) 175:59–73
https://doi.org/10.1007/s10551-020-04669-0

ORIGINAL PAPER

Inducing Corporate Social Responsibility: Should Investors Reward the Responsible or Punish the Irresponsible?

Tyson B. Mackey[1] · Alison Mackey[1] · Lisa Jones Christensen[2] · Jason J. Lepore[3]

Received: 21 March 2020 / Accepted: 31 October 2020 / Published online: 16 November 2020

Abstract

Investors with a pro-social or sustainability agenda increasingly attempt to influence firm managers to adopt socially responsible behavior, either through positive/reward tactics or negative/punishment tactics. This paper considers how investors can use each approach to differentially influence managers to make more CSR investments. The paper uses game theory with an all-pay contest structure to model how a large institutional investor could reward firms for CSR activities by creating a socially responsible investment fund (reward contest) or punish firms via shareholder activism (punishment contest). We identify conditions under which the punishment contest induces a higher level of CSR activity among firms compared to the reward contest. Managers bearing substantial private costs stemming from the activism is one such condition. Spillover effects are seen as the other managers in the economy engage in CSR to avoid being punished by the investor's activism. This level of engagement is not the case when rewards are used—only those managers with an expectation of being rewarded increase their CSR activity in that scenario. This suggests, for example, that incorporating thresholds or tiers (e.g. gold, silver, and bronze-level winners) can increase the effectiveness of reward contests. Implications for designing both positive and negative CSR inducements are explored. We also identify the ethical dilemmas that relate to such influence attempts.

Keywords Corporate social responsibility · CSR · Socially responsible investing · SRI · Share-holder activism · Stakeholders · Investors · Inducements · All-pay contests · Rewards · Punishments · Managerial incentives

Introduction

Stakeholders increasingly attempt to induce firm managers to adopt socially or environmentally responsible behavior (Bhattacharyya et al. 2008; Flammer et al. 2019; Reid and Toffel 2009). Stakeholders who desire corporate social responsibility (CSR) utilize either positively or negatively framed inducement tactics to influence firm managers to modify their behavior (Reid and Toffel 2009; Chatterji and Toffel 2010; Wowak et al. 2017). Positively framed inducements include designing public ratings and rankings (Chatterji and Toffel 2010) or offering targeted infusions of capital (Hill et al. 2007) and emphasize the promise of reward to induce more CSR. In contrast, negatively framed inducements include orchestrating boycotts or protests (King 2008), running negative public relations campaigns (Bartley and Child 2011), or lodging formal shareholder complaints (Glac 2014; Rehbein et al. 2004) emphasize the threat of punishment to induce more CSR (Flammer et al. 2017, 2019). Understanding the impact of these varied tactics on managerial behavior helps stakeholders select or refine influence strategies.

However, the scholarship on pro-CSR stakeholder inducements has progressed in separate or parallel literature streams (Anas et al. 2015; Bhattacharyya et al. 2008; Kurtz 2008). While both types of tactics produce results, the relative merits of one versus the other remain understudied (see Chatterji and Toffel 2010; Elsbach and Kramer 1996; Espeland and Sauder 2007; Jin and Leslie 2003; Martins 2005). In addition to direct effects, other stakeholders may experience indirect effects, introducing

✉ Alison Mackey
amackey@clarkson.edu

Tyson B. Mackey
tmackey@clarkson.edu

[1] Organization and Consumer Studies, Clarkson University, 8 Clarkson Ave., Potsdam, NY 13699, USA

[2] Management, Brigham Young University, Provo, UT, USA

[3] Economics, California Polytechnic State University, San Luis Obispo, CA, USA

questions about the ethics of the externalities of each approach. In a world where stakeholders have limited resources, the question of when each tactic has a greater likelihood of inducing CSR remains vital to strategy selection. Work directly comparing these approaches can provide valuable information regarding which results in the most efficient and effective use of funds. The same scholarship can also provide data on how best to design inducements that drive managerial action, minimize ethical dilemmas, and ultimately maximize CSR in the economy.

Thus, the purpose of this paper is to study two specific CSR inducement strategies (one positively framed and one negatively framed) in order to compare the efficacy of each in influencing firm managers to make CSR investments. In short, and as indicated by the title of the paper, this research asks whether pro-CSR stakeholders interested in inducing corporate social responsibility should focus limited funds on rewarding the responsible or punishing the irresponsible. The paper addresses this question by using game theory and an all-pay contest structure to model the choice facing pro-CSR stakeholders. Specifically, a large institutional investor (representing the pro-CSR stakeholder) considers either engaging in socially responsible investment (SRI) or deploying the threat of shareholder activism to induce firms to engage in CSR. If the investor chooses SRI, she will announce the creation of an SRI fund with which she commits to buy shares in the best CSR performing firms. The emergence of such a fund prompts firms to vie to be selected (and enjoy an increase in market value) (Kurtz 2008). In contrast, if the investor chooses the threat of shareholder activism, she announces intentions to acquire the worst performing CSR firm and fire the offending manager. The presence of such a threat causes firm managers to engage in CSR to avoid being the worst performer at risk for this targeted activism. Equilibrium solutions comparing these choices (as well as more generalized forms of reward and punishment tactics) are explored to assess the net CSR outcomes from each.

A sufficient condition is derived under which punishing the most irresponsible firm generates more CSR activity in the economy than does rewarding the responsible firms. The private costs borne by managers as well as the distribution of firm abilities to engage in CSR activities are instrumental in giving the punishment approach its power. Interestingly, in the punishment approach, all firms exert at least some effort towards CSR; whereas with the reward scenario, only firms close to crossing the threshold for earning the reward do so. Additionally, firms already well over the "winning" threshold in reward settings may reduce CSR activity in the face of stakeholder intentions to reward (only) the most responsible firms. As such, this finding extends prior work in which the pro-CSR stakeholder's influence is limited to only those firms receiving a direct investment (cf., Mackey et al. 2007).

Punishment tactics are not without their risks and complications as they are more sensitive to the idiosyncratic preferences of managers as well as the capabilities of the firms targeted by the strategy. Thus, broadly, the paper considers whether pro-CSR stakeholders interested in increasing CSR activity may be better served shifting resources away from enticing firms with rewards and towards threatening punishments (or vice versa).

The paper proceeds as follows: first, the paper discusses why and how firms typically invest in CSR and the role that pro-CSR stakeholder inducements play in that decision. Specifically, prior work studying SRI and shareholder activism as effective inducements to CSR activity is reviewed. From there, the model is presented, and we explain how structuring the model as two competing all-pay contests allows comparison of the efficacy of the two contrasting tactics. The paper closes by describing the predictions of the model as well as model extensions for other forms of reward and punishment that pro-CSR stakeholders may utilize. Finally, implications for stakeholders, managers, and scholars—including the ethical concerns of investors using potentially coercive and/or threatening tactics—are discussed.

Background: CSR and Firm Responses to Stakeholder Rewards and Punishments

Defining CSR

Corporate social responsibility (CSR) describes deliberate firm actions designed to improve social or environmental conditions (Aquilera et al. 2007; Waddock 2004; Wood and Jones 1995). For example, when firm managers make CSR investments they may spend money in one or more of the following ways: make donations to charitable causes (Godfrey 2005; Cuypers et al. 2016; Wang and Qian 2011), voluntarily reduce emissions and environmental waste (Bansal 2005; Bansal and Roth 2000; Hart 1995; Shrivastava 1995), provide improved working conditions for employees (Jawahar and McLaughlin 2001; Turban and Greening 1997; McWilliams and Siegel 2001), or otherwise support vulnerable populations (Madsen and Rodgers, 2015). Such spending increases aggregate CSR activity in the economy as more and more firms make these investments. Managers have many options to select from and likely have various motivations should they decide to undertake more CSR activity (e.g. legitimacy, values and morals, competitive advantage) (Bansal and Roth 2000).

Pro-CSR Stakeholders

Given the steady rise in consumer and investor interest in the firm-community interface (Bhattacharyya et al. 2008; Jones Christensen et al. 2014) managers face increasing pressure from communities, customers, employees, governments to get involved with addressing social and environmental problems (Arenas et al. 2009). Such "pro-CSR" stakeholders act individually or convene as groups with similar objectives (Ditlev-Simonsen and Wenstop 2013). Whether pro-CSR stakeholders act alone or in collectives, the influence attempts they use represent positively-(reward) or negatively-(punishment) framed inducement tactics. We explore the mechanisms of each approach in turn.

Reward Tactics and Socially Responsible Investing

Reward-based strategies include offering premium pricing, preferential advisory services, infusions of capital, higher rankings, or other reputational capital associated with positive PR (Chatterji and Toffel 2010; Lewis and Carlos 2019) and offer some promise of benefit when managers respond as desired (Fung et al. 2007). Managers may accrue increased executive pay (Kang 2016), takeover protection (Kacperczyk 2009), retirement-related positive legacies (Kang 2016), and image protection (Chatterji and Toffel 2010).

One common type of reward strategy involves socially responsible investing (SRI) (US SIF 2018) in which pro-CSR stakeholders dedicate investment funds to firms that meet pro-CSR criteria (Adam and Shavit 2008). An SRI strategy involves investing money in managed pools of capital that screen out and/or specifically include firms based on factors related to CSR activities (Kurtz 2008). These strategies often revolve around how firms perform according to sustainability or CSR indexes (Lewis and Carlos 2019; US SIF 2018). Alternatively, SRI fund managers may select which firms to include in the fund based on firm product offerings or other similar criteria. The efficacy of SRI funds to effect changes in aggregate CSR stems, to some degree, from the premise that the positive attention and increased funds associated with being "chosen" effectively increase the share price of the included firms. These firms then enjoy an associated increase in market value, which can generate a virtuous cycle that motivates firm managers to direct (or keep directing) resources to CSR activities.

In order to be maximally effective, SRI funds publicize the social and environmental performance data and the associated corporate ratings (Chatterji and Toffel 2010; Paul and Lydenberg 1992; Sadowski et al. 2010). The public data aids fund managers and investors in directing capital to the highest or best performing firms. As such, the goal of being rated or ranked can become an end in itself for firms. Empirical findings link such enticements to increased resources being directed to CSR activities (Elsbach and Kramer 1996; Fung et al. 2007; Jin and Leslie 2003). Scholars estimate that over 12 trillion U.S. dollars is currently invested in SRI funds (US SIF 2018). Given the frequency with which pro-CSR stakeholders engage in socially responsible investing tactics (Kurtz 2008), this work uses SRI investments in our model to represent positively framed inducement options.

Punishment Tactics and Investor Activism

In direct contrast to the positively framed inducements, punishment tactics attempt to generate more CSR by either encouraging firms to stop engaging in "irresponsible" behavior or by encouraging firms to increase CSR to avoid negative consequences. Common CSR-related punishment tactics include crafting shareholder resolutions; staging protests, boycotts, negative press, and mass dumping of shares; or other activism aimed at driving down earnings or stock price of the offending firms (cf. den Hond and de Bakker 2007; King 2008). The efficacy of activism to effect changes in aggregate CSR stems, to some degree, from the expectation that myriad other stakeholders follow and emulate the activists and "punish" the targeted firm (Fung et al. 2007; Vasi and King 2012). Punishments trigger lower sales/revenues, or generate negative employee-related outcomes (turnover, lower morale, loss of discretion). Other CSR-related negative outcomes tied to threats include increased involvement from regulators for the targets of the activism or increased personal costs to managers in the offending firms (Baron et al. 2011; Flammer et al. 2017, 2019; Reid and Toffel 2009).

Shareholder activist tactics effectively reduce managerial discretion or remove managerial control for managers in poorly performing CSR firms (Baron 2001, 2009; Binder and Neumayer 2005; Gillan and Starks 1998; Lenox and Eesley 2009). Such tactics essentially "force the hand" of managers, particularly when activists broadcast undesirable activities or inadequacies in CSR performance. Specifically, with a shareholder activism strategy, influence comes from a variety of places such as shareholder proposals and resolutions (created and/or voted on) as well as from activists selling shares over a particular unresolved issue (Mathiasen 1994). Often, these proposals link corporate governance and CSR issues (e.g. tying executive pay to social or environmental performance, or to changes to racial/gender composition of board) (O'Rourke 2003; Flammer et al. 2017, 2019).

Empirical findings on such tactics indicates that they are effective, with the chances for success rising the more severe and disruptive the tactics are in design (King 2011). Accordingly, fears over declines in stock price (Davidson et al. 1995; Pruitt and Friedman 1986; Pruitt et al. 1988), lost or lowered wages (Parthiban et al. 1998), negative legacy creation (Kang 2016), or a tarnished public image

(Chatterji and Toffel 2010)–in addition to shareholder activism aimed at removing and replacing management (Kacperczyk 2009)–drive managerial strategic choices in relation to CSR. Not surprisingly, the use of activist tactics continues to rise (David et al. 2007; Flammer et al. 2017; Johnson and Greening 1999; Rehbein et al. 2004). Given the frequency with which pro-CSR stakeholders engage in activism aimed at imposing costs on firms and managers, this work uses shareholder activism in our model to represent negatively framed inducement options. Specifically, the pro-CSR stakeholder threatens to acquire the worst CSR performing firm and fire the manager. Takeover threats (even if unsuccessful) can result in managerial dismissal (Denis and Serrano 1996) and often spillover to competitor firms who preemptively dismiss their own manager hoping to avoid a takeover threat themselves (Walsh and Kosnik 1993).[1]

Context for Comparing Rewards and Punishments

Pro-CSR stakeholders rarely attempt to protest and partner with an organization at the same time (see Rondinelli and London 2003 for exceptions). Thus, stakeholders generally choose to either reward the responsible or punish the irresponsible. Further, in a world of scarce capital, it is important for stakeholders to understand the most efficient and effective use of efforts to induce managers towards CSR activity. As noted, our model compares two specific reward and punishment tactics in order to provide guidance on how to spend limited funds in pursuit of maximum aggregate CSR. In the next section, we introduce a model in which a pro-CSR stakeholder (as either a high-net worth individual or an institutional investor, hereafter referred to as "investor") chooses between rewarding the responsible or punishing the irresponsible in her efforts to maximize the amount of CSR activity in the economy. Of course, as with any stakeholder, this investor has resource constraints and is therefore most interested in multiplying her dollar-for-dollar influence to reach as many firms as possible with inducement efforts.

Model

Consider an economy with a set of firms $N=\{1,\ldots, n\}$. In this economy, firms choose a portfolio of CSR activities $s_i \geq 0$ that they will perform. These CSR activities could take various forms as listed above and are only specific in that they carry costs to the firm ($a_i > 0$ for all i).[2] Reflecting reality, the CSR can vary anywhere along these lines, and extant analogs include, but are not limited to spending on social or environmental projects unrelated to product portfolio, discontinuing currently profitable product lines to address social or environmental concerns, engaging in cause-related philanthropy.

Firms vary in their ability to conceive of and implement socially responsible activities (Barney 1991; Peteraf and Barney 2003). These differences in capabilities are captured in the cost component in the model—that is, firms can be indexed (from best to worst) based on the heterogeneous costs of the CSR activities such that $a_1 > a_2 > \ldots > a_n$. Note that the marginal cost of these activities is defined as $1/a_i$ per unit of CSR activities. Thus, the "best" CSR performing firm is the firm with the lowest marginal cost of engaging in CSR. The model also assumes that the difference between the cost of socially responsible activities of the lowest and highest cost firm is bounded such that $a_1/n < a_n$. This last assumption is for mathematical convenience and does not impact the conclusions of the model.

Each firm in the economy is run by a utility-maximizing manager. Each manager seeks to maximize his compensation at the firm. That payoff is defined as $u_i(s) = w + \alpha y_i$, where w is the compensation that can be lost if the manager is fired (i.e. future compensation, both cash and stock grants), y_i is the expected market value of the firm, and α is the percentage of shares owned by the manager such that $0 < \alpha < 1$. This second term, αy_i, is the compensation that cannot be lost if the manager is fired as it represents the monetary value of shareholdings already owned by the manager.

As noted, in this economy there is a pro-CSR stakeholder (either a high-net worth individual or an institutional investor, hereafter referred to as "investor") with interests that go beyond maximizing personal financial well-being: the

[1] Given the rise in the number and type of newer organizational forms, such as B-Corps, social purpose firms, and low-profit LLCs—types of firms which have some level of social benefit "baked into" their organizational governance, could increase the likelihood of negative influence tactics rising. Specifically, as more firms make legally binding decisions about CSR topics, they run the risk that some stakeholders may become frustrated with how much CSR the firm does or does not pursue. Thus, the generalizability of punishment contests may be increasing to the extent that more firms embrace these newer organizational forms.

[2] Defining CSR to be costly enables us to isolate the effect from the investor's inducement tactic from other potential stakeholder inducements (cf. Flammer 2015). For example, customers can induce managers to adopt CSR activities when they are willing to pay premium prices for products. Here, for parsimony, we assume the net present impact of the CSR activities is negative so the only inducement a manager feels stems directly from the investor's tactics (cf., Mackey et al. 2007). Costly CSR also introduces a theoretically interesting nuance about the kind of CSR that many stakeholders likely prefer or deem more "authentic" (Benabou and Tirole 2006; Cording et al. 2014; Godfrey 2005; Makov and Newman 2016; Newman and Cain 2014).

investor also wants to maximize CSR activities across the economy. However, this investor has a fixed budget $b>0$. Thus, the goal of the investor is to maximize the expected CSR activities within her budget $V(s)=E[s_1+\ldots+s_n]$ using her budget b. To accomplish these ends, the investor will use her investment budget to motivate the managers of some firms in the economy into increasing their firms' CSR activities.

The investor can announce one of two contests to motivate managers to direct firm resources toward CSR activity.[3] The contests are structured as "all-pay," meaning that all players must make investments or exert effort before any one player can know whether or not he or she "wins." (cf. Siegel 2009). For example, an R&D race can also be modeled as an all-pay contest in which firms must make their investments in new product development before any one firm can know whether they will emerge as the technological standard (or some other type of winner). In this case, creating an SRI fund is modeled as an all-pay contest in which firms will decide the CSR activities they will engage in before they know which firms are selected for the SRI fund. Further, shareholder activism is modeled as an all-pay contest in which firms will decide their CSR activity before they know which firm is targeted by the investor for a hostile takeover and replacement of management. The investor is agnostic about what tactic to employ (e.g. she has no bias for or against rewards or punishments); the investor simply wants to select the option that leads to the largest net increase in firm socially responsible behavior.

The Reward-Type Contest: Including the Best CSR Firms in an SRI Fund

Under the first contest, the investor will create an SRI portfolio (akin to a mutual fund) to buy shares in the top m CSR performing firms. These top firms each receive a boost to their market value of b/m—that is, the portion of the investor's budget that is invested in the firm from being selected into the fund. The increase will depend, of course, on how large the investor's budget is and how many firms are selected into the SRI fund.

Each firm's market value is based on three components: (1) the present value of cash flows, (2) the expected value of a "boost" from being selected into the SRI fund and (3) the total cost of the socially responsible activities. More technically, the firm's market value is $y_i = v_i + p_i(s)\frac{b}{m} - \frac{s_i}{a_i}$, where v_i is the present value of a firm's cash flows, $p_i(s)b/m$ is the expected boost from winning the contest, and s_i/a_i is the total cost of their socially responsible activities.

A firm's probability of being selected for the fund would depend on how many firms are being selected and whether a particular firm engages in more CSR than enough other firms to make the cutoff. Thus, given a profile of CSR actions s_i, firm i's probability of getting an increase in demand for its stock, denoted by $p_i(s)$, satisfies

$$p_i(s) = \begin{cases} 0 & \text{if } s_j > s_i \text{ for } m \text{ players } j \neq i \\ 1 & \text{if } s_j < s_i \text{ for } n-m \text{ players } j \neq i \\ \text{any value in } [0,1] & \text{otherwise,} \end{cases} \tag{1}$$

such that $\sum_{i\in N} p_i(s) = m$.

The intuition being that if there are m firms engaging in higher levels of socially responsible behaviors than the focal firm, the probability of winning the contest will be zero. Likewise, if there are less than m firms with higher levels of socially responsible performance, the firm wins the contest with a probability of one.

Hence, in this contest, for a firm with a market value of $v_i + p_i(s)b/m - s_i/a_i$, the payoff of manager i will be

$$u_i(s) = \alpha\left(v_i + p_i(s)\frac{b}{m} - \frac{s_i}{a_i}\right) + w. \tag{2}$$

Again, the manager's utility is a function of current stock holdings αy_i and expected future compensation w. The value of the current stock holdings depends on the outcome of the contest (i.e. which firms are the top performing CSR firms included in the SRI). Since the manager's utility is a function of his other holdings in the firm, the managers selected for the SRI fund will accrue personal benefits from "winning" the contest/being included in the fund.

All the firms in the economy are aware of the investor, her intentions for creating the SRI fund, and her preferences for CSR activities. What happens in such a scenario is that the boost for the manager from being included in the SRI fund must be greater than the cost of engaging in socially responsible activities to motivate the manager to action. Said more technically, in such a contest, Siegel (2009) has shown that an equilibrium solution exists and at least one equilibrium has the expected payoff of the players equal to $E[u_i^*] = \alpha v_i + w + \alpha \max\{b/m - ba_{m+1}/ma_i, 0\}$. The intuition for this is that, all else equal, CSR activities lower the manager's compensation because the firm has lower profits (costly CSR activities) and the manager's compensation depends on the firm's financial performance. Thus, the boost from being selected for the SRI fund must be greater than the costs associated with the CSR activities for the utility-maximizing manager to choose to engage in CSR (conditions

[3] The proceeding analysis relies heavily on results established in Siegel (2009) and Lepore et. al (2012). Baye et al. (1996), Clark and Riis (1998) and Gonzalez-Diaz (2012) all make important contributions to understanding equilibria of all-pay contests with perfect information.

under which this does and does not hold are explored in a later section).

The Punishment Contest: Targeting the Worst CSR Firm with a Hostile Takeover

Alternatively, the investor could take the opposite approach. Instead of inducing managerial action with the lure of increased firm value and associated managerial compensation, the investor could instead decide to announce a different contest to acquire the firm with the lowest level of CSR (i.e. the most socially irresponsible firm) and then to fire the offending manager of that firm.[4]

Given a profile of costly CSR actions s_i, firm i's probability of not getting acquired is denoted by $\pi_i(s)$, which satisfies

$$\pi_i(s) = \begin{cases} 0 & \text{if } s_j > s_i \text{ for } n - \tilde{m} \text{ players } j \neq i \\ 1 & \text{if } s_j < s_i \text{ for } \tilde{m} \text{ players } j \neq i \\ \text{any value in } [0,1] & \text{otherwise,} \end{cases} \tag{3}$$

We define $\tilde{m} = 1$ due to budget constraints of the investor–the investor may not have a budget sufficiently large to buy more than one firm. Said differently, the probability of losing this contest and being the "punished" firm depends upon the level of CSR activity that the firm puts forth–the firm with the lowest CSR activity is acquired with a probability of 1. Of course, for the firm that gets acquired, the fired manager will lose cash compensation and future stock grants, but gain his reservation wage r, where $w > r > 0$, while retaining equity in the firm αy_i.

Hence, in this second contest, the manager's payoff is

$$\tilde{u}_i(s) = \alpha\left(v_i + \big(1 - \pi_i(s)\big)b - \frac{s_i}{a_i}\right) - \big(1 - \pi_i(s)\big)\Delta, \tag{4}$$

where the compensation lost for the manager is defined as $\Delta = w - r$. The manager's utility, in this scenario, is a function of the change in stock value and the compensation lost if acquired. The utility for all managers in this economy is lower if the firm is acquired (and they are fired). This induces all managers to exert at least some effort (i.e. increase their firm's activity in CSR) to avoid being acquired.

This expression in (4) can be rewritten, for mathematical convenience, as

$$\tilde{u}_i(s) = \underbrace{\alpha v_i - \Delta + \alpha b}_{\tilde{v}_i} + \pi_i(s)(\Delta - \alpha b) - \alpha\frac{s_i}{a_i}. \tag{5}$$

In such a contest, Siegel (2009) shows that an equilibrium exists and at least one equilibrium has the expected payoff of the players such that $E\left[\tilde{u}_i^*\right] = \alpha v_i - \Delta + \alpha b_i + \max\{(\Delta - \alpha b) - \frac{a_n(\Delta - \alpha b)}{a_i}, 0\}$. Simply stated, the intuition of this equilibrium is that the loss to the manager from her/his firm being acquired and her/him subsequently being fired is greater than the boost to the value of the manager's stock holdings from getting acquired. This is what motivates managers to engage in CSR in this scenario–preservation of employment.

As noted previously, investing capital in either contest gives the same expected financial payoff to the investor. That is, neither investment is inherently financially better than the other. This means that the decision to announce one contest or the other is wholly determined by which investment maximizes the total expected effort into costly CSR activities, $E[s_1 + \ldots + s_n]$. Thus, the investor seeks the choice that maximizes the aggregate CSR in the economy.

Comparing Reward versus Punishment in Driving Manager CSR Behavior

In comparing the solutions to both contests, we derive a sufficient condition for when the compensation losses the manager faces from losing his job (due to the hostile takeover) are greater than the gains to stockholdings caused by being included in the SRI fund—that is, we derive a sufficient condition for when the threat of punishment is more effective than the promise of reward. We do this by setting up an inequality in which a manager's utility is greater under the contest in which a socially irresponsible firm is acquired than it is under the contest in which stock in the socially responsible firms is purchased. Proposition 1 defines this sufficient condition (see Appendix 1 for formal proof of the condition).[5]

Proposition 1: *Acquiring the firm with the lowest level of socially responsible activities induces higher levels of social responsibility across the economy than does investing in the stock of the highest performers if*

$$\Delta \geq \alpha b + \alpha b\frac{a_2}{na_n - a_1} \tag{6}$$

As noted, this sufficient condition holds when the manager's wage loss from being punished is greater than the acquisition premium she would gain from her firm being taken over by the investor. Certainly, the punishment

[4] Importantly, Proposition 1 is a sufficient condition. Simply stated, if Proposition 1 does not hold, it does not then necessarily imply that the reward-based contest is better than the punishment-based contest.

[5] Notice from Proposition 1 that the heterogeneity term, $(a_2/(na_n - a_1))$ is multiplied by the manager's stock holdings α, so in the case of low stockholdings, heterogeneity in CSR abilities will not as easily be a factor against obtaining the sufficient condition.

design needs to be very punishing such that all the managers in the contest really want to avoid being punished (cf., King 2011). When punishment is designed to target the manager's job security/compensation, the manager will stand to lose more when (1) the compensation in the current job increases relative to the manager's next best option and/or (2) when the manager has less valuable shareholdings to take with her to "soften the blow" of losing the job (shareholdings can also decrease from having fewer shares or those shares being individually less valuable). For example, some managers may be extracting rents in their current position while others have limited outside options for employment (e.g. binding non-compete contracts, working in extremely path-dependent industries/narrow niche, etc.) such that they cannot find ready and comparable substitute employment. It is the private costs to the manager, tied to his or her job security and compensation, which causes the negative, punishment-based contest to dominate the positive, reward-based contest in this context. Facing substantial personal costs, the self-interested manager may be motivated to increase a firm's CSR activity not because it will enhance the value of the firm, but only to save his or her job.

The sufficient condition also depends on the degree of heterogeneity across firms in terms of the ability to conceive of and implement CSR activities. In some respects, in a heterogenous mix, the worst performers will be the first to go. A salient analogy relates to adages about life in the wilderness, where the slowest or worst performing prey will be consumed by the predator- and all that the other potential prey need to accomplish is to be faster or better performing than the slowest others. If the "predator" is the investor, then as heterogeneity in firm CSR abilities decreases, more effort is required to outrun the weakest firm to escape, and the punishment contest becomes stronger. Thus, in cases in which firms are more homogenous in abilities to implement CSR, the sufficient condition is more likely to hold.

Of note, in comparing the equilibrium conditions of the two contests, all firms in the economy exert some effort to increase CSR engagement in the punishment contest, hoping to avoid takeover (as alluded to in the adage of the running from the predator); whereas, in the reward contest, only those firms with a realistic expectation of being selected into the fund (those closest to the cutoff) put forth effort into increasing their CSR activities (see Lewis and Carlos 2019). Indeed, it is likely that in the reward-contest, $n-m-1$ firms will expend no CSR effort, as these firms do not have reasonable expectations of winning the contest. So, depending on the investor's preferences, they may prefer a punishment regime to get more firms involved in CSR (even if that initial involvement is modest) to a reward regime in which a small set of firms get involved at higher levels of CSR. Certainly, some investors may prefer this latter approach, perhaps believing that getting a few trend leaders involved in CSR will cause those behaviors to diffuse through the economy.

Given the importance of exploring boundary conditions for theory development, we next introduce a more generalized version of the reward and punishment contests to consider how the conditions derived above may still apply to reward and punishments contests more generally. We also relax some of the restrictive assumptions about firms and managers that were used in the original setting.

Generalizing Reward and Punishment Contests

We now generalize the reward and punishment contests by not restricting the way that the investor's budget b transfers to the utility of the winner (or loser) of the contest—that is, we do not restrict the reward contest to be inclusion in an SRI fund nor do we restrict the punishment contest to be the hostile takeover of the worst social performer. Both contests still impact the financial value of the firm (as opposed to some other non-financial incentive) and the manager's utility ($u_i = w + \alpha y_i$) is still tied to that value. So, in the case of the reward contest, the firm's market value is defined as $y_i = v_i + p_i(s)g_i(b) - s_i/a_i$. Note the use of the general function $g_i(b)$ that translates b dollars of expenditure to units of gain (either direct or indirect) for each specific manager i (i.e. gains/costs are idiosyncratic to each manager). Thus, the scenario of the SRI fund as a reward regime is just a special case of this generalized function in which $g_i(b) = b/m$ for every manager i.

Likewise, in the case of the punishment contest, the firm's market value is defined as $y_i = v_i - \left(1 - \pi_i(s)\right)c_i(b) - s_i/a_i$. Analogous to $g_i(b)$, the general function $c_i(b)$ translates b dollars of expenditure to units of cost for a manager i. Again, this could be indirect or direct costs. And, again, the scenario of the hostile takeover is just a special case of this generalized function for which $c_i(b) = (\Delta/\alpha) - b$. For notational parsimony we omit the b from the notation and write g_i and c_i.

In the reward contest we label $\gamma_i = a_i g_i$. This is manager i's benefit of the prize divided by her marginal cost of CSR effort. Without loss of generality, we order the players such that $\gamma_1 \geq \gamma_2 \geq \ldots \geq \gamma_n$. The payoff to the manager in the reward contest is

$$u_i(s) = \alpha\left(v_i + p_i(s)\frac{g_i}{m} - \frac{s_i}{a_i}\right) + w. \tag{7}$$

In the punishment contest we label $\xi_i = a_i\, c_i$. This is manager i's cost of the punishment divided by her marginal cost of CSR effort. Without loss of generality, we order the players such that $\xi_1 \geq \xi_2 \geq \ldots \geq \xi_n$. The payoff to the manager in the punishment contest is

$$\tilde{u}_i(s) = \alpha\left(v_i - (1-\pi(s))c_i - \frac{s_i}{a_i}\right). \quad (8)$$

Notice that we are not ordering by a (costs/abilities of CSR activities) but rather we are ordering on γ and ξ. Note that the two indices (γ and ξ) could be different—that is, the ranking of best to worst firms could be different for the reward contest compared to the punishment contest. Again, in comparing the solutions to both contests, we derive a sufficient condition (see Appendix 2 for formal proof) for when the threat of punishment is more effective than the promise of reward:

Proposition 2: *In general, punishing the firm with the lowest level of socially responsible activities induces higher levels of social responsibility across the economy than rewarding the highest performers if*

$$n\xi_n - \xi_1 \geq \gamma_2 \quad (9)$$

In considering the implications of this sufficient condition, recall that originally the model assumed that managers faced the same incentives to vie for the reward or avoid the punishment—that is, managers had the same shareholdings, future compensation, and the same outside employment options. We also assumed that the financial impact on firm value of a given reward or a punishment tactic would be the same for each manager (firm). These assumptions were made for tractability in the model. When we relax these assumptions and allow these factors to vary, findings hold but nuances emerge.

First, suppose that for some manager the reservation payment (r) is larger than the expected future compensation (w) for working in the firm. This could happen because this manager has a very large golden parachute exit package (Ujah and Okafor 2019). While this manager cannot quit and obtain the parachute, she may prefer to be fired under the punishment contest and would therefore put forth zero CSR activity on behalf of the firm. Further, if other firms in the contest knew that one firm was certain to put forth zero CSR activity, the other firms could put forth only a token level of CSR effort. The negative, punishment-based contest would be thus be stripped of effectiveness in motivating CSR activity.

Second, suppose there is heterogeneity in the stockholdings of managers (α), such that one manager has a very large stake in the firm (α is large), a very low salary, and a very low reservation payout (w and r near zero). This could happen if the firm is owned primarily by its founder. Assuming that the founder-manager only (or mostly) derives utility from the wealth produced by the firm–that is, no psychic benefits from running a firm or from CSR activity–this manager would be perfectly happy being acquired by the investor at full market value. Thus, he would put forth no effort toward CSR activity. Just as in the previous example, if the other firms in the punishment-based contest knew that this firm would put forth zero effort, the contest would be bereft of any ability to induce more than a superficial amount of CSR activity.

Third, with generalized reward and punishment functions, g_i and c_i, we can consider punishments and rewards in all-pay contests beyond takeovers and investments, so long as those punishments and rewards influence the value of the firm (and thereby the manager's compensation). These could include, for example, boycotts/protests, which may reduce the firm's value (unlike the takeover, which *increases* firm value) or positive PR, which may increase the firm's value. We can also consider reward and punishment contests with heterogeneous effects on firms. For example, some firms may have a sufficiently large stockpile of public goodwill as "insurance" against boycotts or bad PR (Godfrey 2005). With sufficiently strong insurance, an attempted boycott may fail, and therefore not be punishing, which would limit the threat of the punishment contest for such firms. Alternatively, a firm may have such a poor reputation that good PR would not be able to increase the firm's value. Firms of this type would be unwilling to participate in the reward contest.

Across these scenarios, the degree to which the investor can "hurt" the managers in the contest varies, underscoring an important boundary condition for the model—that is, the pro-CSR stakeholder needs to construct the contest to exclude managers (and firms) that may have no incentive to engage in CSR, even in the face of threatening punishments. Failure to exclude these managers (firms) would limit the power of such tactics to influence other firms to respond to the contest. And in the extreme case that at least one manager does not care at *all* about the punishment (and thus puts forth zero effort) *and* this is widely known, the other managers in the contest would only need to put forth a small effort to avoid punishment. Such a scenario would render the punishment contest very ineffective. For this and the other reasons given above, the investor may need to include some sort of bonding mechanism to ensure that the punishment is a credible threat. Inasmuch as the bonding mechanism represents an incomplete contract, the investor may be able to adjust the terms mid-contest to cull those firms who have managers who are relatively indifferent to being fired.

Note that even though this generalized model enables us to identify some important boundary conditions on the power of punishments over rewards, difficulties arise when we broaden the possible rewards and punishments into activities like positive PR and boycotts. First, a dollar spent by the activist on good PR may increase firm value by a greater amount than a dollar spent promoting the boycott would reduce firm value (or vice versa), whereas in our previous cases, the spending had a direct effect on firm value—spending *b/m* on firm stock

increased firm value by *b/m*. In the PR vs. boycotts scenario, the relative productivity of spending on PR vs boycotts may be the main driver of the activist's decision instead of whether the contest was a reward or punishment regime. Additionally, both PR and boycotts are also more likely to succeed when the targeted firms compete in consumer-facing industries; whereas the interventions based on buying firm stock will not have these limitations.

That said, one advantage of PR and boycotts over acquiring shares of firms is that, as firms are allowed to differ in size and market valuation, the size of an activist investor's acquisition is limited by her budget—the largest firms cannot be invited to the contest—but there may be economies of scale in an activist's investment in PR and boycotts. And regardless of whether the contests are modeled as buying stock, PR boosts, or boycotts, the activist's influence to induce CSR extends beyond just the firms receiving the investment. This outcome contrasts with direct investment without a contest structure in which the investor's influence is limited to only those firms which received an investment (cf., Mackey et al. 2007).

Lastly, consider the scenario we alluded to before in which firms do not vary in their ability to conceive of or implement CSR activities ($a_1 = a_2 = \ldots = a_n$) but the ability of the investor to impact the manager varies. In this case, proposition 2 reduces to

$$nc_n - c_1 \geq g_2. \tag{10}$$

To simplify the expression, we normalize $g_2 = 1$. From that simplification, we obtain an inequality (for which punishment is superior to reward) based on the ability to punish the two extreme managers (i.e.. the best and the worst)—that is, a contest design that is motivating to both the best and worst CSR performers, is motivating to the entire range of performers.

$$c_n \geq \frac{1 + c_1}{n} \tag{11}$$

In this case, the investor would only want to include the *n*th manager in the punishment contest if $nc_n > kc_k$ for all $k \in \{1, \ldots, n-1\}$, the idea being if the punishment is not sufficiently punishing to the worst manager *n*, then the investor should only extend the contest to the *k*th manager for which the contest is punishing.

Taken together, there are many factors for pro-CSR stakeholders to consider in designing inducement strategies aimed at increasing CSR across the economy. We turn now to discuss the predictions of our model and applications to practice and policy.

Discussion

Herein, we derived conditions with which a pro-stakeholder induces more CSR activity by punishing the most "irresponsible" firms than he or she induces by rewarding the most responsible. Severe consequences tied to managerial compensation and job security are particularly effective. Perhaps not surprising, in the most extreme condition, the managers earning economic rents *(w–r)* in their jobs are very motivated to not lose the compensation associated with those positions. Similarly, managers with limited outside options are also very motivated to not be fired. Further, the hostile takeover threat is more effective when firms are more homogenous in their abilities to conceive of and implement CSR activities.

We also looked at less restrictive models for rewards and punishments and found that under the threat of punishment, all firms increase CSR activity and not just those on the margins (as happens with reward contests). This finding might be somewhat surprising to stakeholders biased toward positively framed inducements, as well as to those who hold strongly to the feeling behind colloquialisms such as "you catch more flies with honey." However, this paper introduces the idea that some types of CSR, particularly those which have not diffused widely, may be effectively induced using punishment contests to increase total firm engagement. Yet, we also find that punishment contests entail risks, as negative tactics are sensitive to details such as: the kind and degree of punishment, the number of managers included in the contest, and the capabilities of firms for engaging in the CSR activities.

This paper thus introduces several practical and ethical considerations for pro-CSR stakeholders to consider in their efforts to increase CSR activity. Findings open new discussions related to: incorporating personal costs to managers in punishment regimes, the role of tiers in improving reward tactics, the question of why pro-CSR stakeholders would consider contest structures (i.e. the benefits of spillover effects), and ethical concerns with using coercive tactics that may send deceptive signals to stakeholders. We discuss each in turn below.

Designing Punishments to Impose Personal Costs on Managers

As outlined, imposing personal costs on a manager for his or her firm's CSR activity (or lack thereof), and substantial costs at that, is a critical mechanism driving the effectiveness of the punishment contest (cf. Flammer et al. 2017; King 2011). Thus, in extending findings to real world behaviors, it follows that when stakeholders design a punishment-based strategy they should design substantial personal costs

to managers. This decision remains instrumental to the efficacy of the punishment approach. When shareholder activism varies in the extent to which it imposes such costs, the effectiveness of the tactic will likewise vary. Notably, although shareholder activism took the form of a threatened hostile takeover and subsequent removal of offending managers in this model, the model generalizes to any other punishment-based forms of shareholder activism to the extent that the alternate form imposes personal costs on managers (e.g. protests and boycotts, etc.) (King 2008). Punishment contests will likely work best for CSR activities that can be completed during the period of the contest (e.g. helping a country ravaged by a natural disaster), as compared to ongoing activities that the firm may abandon after the contest ends (e.g. reducing carbon emissions or tackling racial bias). Design features matter significantly with punishment regimes.

Making Rewards Work: The Role of Thresholds

Reward-based inducements for CSR will likely work better with tiers (or thresholds) associated with positive incentives. By this we mean having more than one type of "winner." To elaborate, we return to the finding that rewards are limited in effectiveness as they only motivate those firms with a reasonable expectation of winning (Vroom and Jago 1978; Pinder 2014; Zedelus et al. 2014). Given the prevalence of ratings as central to most reward-based strategies (Scalet and Kelly 2010) (e.g. Domini, KLD, FTSE 500, Dow Jones Social Index, Fortune's Most Admired Companies)—this result has significant implications for the entire practice of making third party-ratings. The underlying rationale of rating systems is that the information provided to a firm's investors from these rating or rankings motivates firms to modify behavior (Fung et al. 2007; Paul and Lydenberg 1992); however, in many cases, the ratings/rankings may actually create a *disincentive* for managers who see achieving a positive rating or a ranking as too distal to matter (Lewis and Carlos 2019).

In addition, thresholds are often embedded within CSR rating/ranking systems. For example, there are only 50 winners in Fortune's 50 Most Admired Companies list. Thus, there are countless firms without any expectation of winning and therefore not motivated by the reward. And worse, firms that are well above the threshold may reduce effort, still staying above the threshold, and still winning the reward. This finding from the model is consistent with prior work suggesting that poorly rated firms enact more CSR after the publication of ratings than do highly rated firms (cf. Chatterjee and Toffel 2010; Lewis and Carlos 2019).

One way to attenuate the negative effects of expectancy on reward-based systems is to have multiple thresholds (or tiers) so that more firms are motivated to exert effort (e.g. resort to "gold/silver/bronze" level prizes). Firms may not be able to win at one level, but they may be able to win at another. In essence, introducing multiple prizes for the reward contest increases the number of firms willing to exert effort. Another way to improve effectiveness is to define the contest to focus on forms of CSR that are more difficult to implement and require more capabilities on the part of the firm. As noted previously, when the distribution of capabilities for CSR activities is significantly dispersed across firms in the economy, then reward systems may outperform punishment-based systems.

Increasing Aggregate CSR in the Economy: The Role of Spillover Effects

Another contribution, and an important benefit of a pro-CSR stakeholder using a contest structure, is that the contest can create spillover effects. That is, the investor gets firms involved in doing CSR without direct investment, effectively increasing her reach. For example, recall that in the punishment contest, all firms in the economy engaged in CSR activities to avoid being the worst social performer. Without a contest, the investor essentially (and only) "pays" for the CSR that she helps to generate in the economy. Central to this outcome is the fact that the investor is not simply buying the "worst" firm just to bring the downfall of one specific offending manager. Instead, the investor is influencing all the other firms in the economy to become more socially responsible. It would be inefficient to just focus on making the least capable CSR-performing firm engage in CSR. One could even argue that is not an efficient use of resources to target such an incompetent firm. However, more capable firms increase their CSR activities to avoid being the "worst" and the target of negative shareholder activism. Likewise, when the investor seeks to create an SRI fund, the intent is not to make the best CSR firm better. Instead, it is to raise the overall level of CSR activity across the economy. An investor could just target the most capable CSR performers and incentivize them to increase efforts; however, the contest approach extends the investor's dollar-for-dollar influence. This suggests to pro-CSR stakeholders that they can influence more firms than their limited resources might otherwise achieve.

Ethical Considerations of Stakeholder Pro-CSR Tactics

Any endeavor from stakeholders to influence firms to engage in CSR may have important ethical implications. This paper suggests that the threat of injurious personal costs to the manager is what gives the punishment regime its effectiveness over the reward regime. The degree to which pro-CSR stakeholders adopt influence tactics that exert questionable,

coercive, or threatening pressures invites potential criticism on ethical grounds. Of course, the effectiveness of any particular influence tactic depends upon managers really wanting to avoid the punishment (or earn the reward, as rewards can also be coercive in nature). This clear incentive on the part of managers could bias pro-CSR stakeholders to prefer using the most motivating, and potentially more coercive tactics. The ethical orientations of the investor (e.g. instrumental versus deontological logics) becomes relevant in considering the ethical implications of our results. Managers who think the end justifies the means may be more prone to punishment tactics. If the findings herein lead more stakeholders to use negatively framed contests, there may be a net increase in threats and coercive activity.

From an ethical standpoint, pro-CSR stakeholders should be free to explain (and expand) their normative stance favoring more CSR in the economy, even going so far as to try and convince firm managers that they should be doing more CSR. However, once these efforts cross the line into badgering, bullying, or intimidating others, the question becomes whether threats and coercion (even positive coercion) have any place in ethical business practices (Otteson 2019). Further, to the extent that accusations of bad faith or coercion generate negative externalities, firms and investors could suffer reputational and financial repercussions. At the extreme, such activities would result in zero-sum exchanges, creating no net value for society.

At a more fundamental level, there may be an even deeper ethical concern behind pro-CSR stakeholders attempting to motivate firms to engage in CSR activities. Implicit to such efforts is the idea that pro-CSR stakeholders need to provide an external (or additional) motivation to induce managers to engage in CSR. Yet, many customers, suppliers, employees, and communities want to support firms with principled commitment to CSR. Such commitment may act as a signal that the manager (firm) will "do right" (in the future) by customers, suppliers, employees, communities, regardless of the financial implications of those choices (Clark et. al. 2020). If a manager is only engaging in CSR activities because she wants an external reward (e.g. inclusion in SRI fund, positive press, rewards/recognitions, more sales volume) or she wants to avoid an external punishment (boycotts, shareholder activism, negative press), then in the eyes of those wanting "authentic" commitment to CSR, such managers could be seen as acting deceptively and (at an extreme) even fraudulently (Clark et al. 2020).

The implications of our work extend beyond the simple question of whether CSR is or is not inherently an ethical issue; an argument which some scholars hold given that that the scholarly study of CSR partially arose from studying applied ethics (Jones Christensen et al. 2007). Rather, the implications listed above provide direction for future research on the ethics of incentive systems in the context of CSR, some of which are outlined below.

Future Research Directions

Scholars and practitioners rely on models to test new ideas and gain insights to apply to field locations and actual firms. Models provide a fundamentally unique perspective in the study of phenomena; they play an integral role in theory-building when they allow simultaneous testing of new or unusual behaviors (Roth and Wilson 2019). While there are dangers in overly relying on model results when designing policy and managerial practice, there is power in using model results to refine research questions and justify new directions for scholars (Roth and Wilson 2019). Herein, modeling allowed this work to compare alternative paths for the pro-CSR stakeholder, finding the key mechanisms that drive investor effectiveness and helping identify design features that improve both types of CSR inducements.

While firms may engage in CSR for myriad reasons (Bansal and Roth 2000), influence from pro-CSR stakeholders is an increasingly important antecedent (Flammer et al. 2017, 2019; Lewis and Carlos 2019). To isolate the impact on CSR from such stakeholders, the model holds constant other drivers of CSR (by restricting CSR to that which is costly to firms) and focuses on the impact inducement strategies can have.

Herein, the paper focused distinctly on how investor influence can best affect market behavior associated with CSR. Of course, other stakeholders, such as employees, customers, communities, can also elect to become an investor that wields positively or negatively framed inducements for CSR (Freeman 1984). Accordingly, these results have implications for multiple stakeholders interested in increasing aggregate CSR in the economy. Such constituents may have different incentives and including them in future models and empirical explorations could build the pro-CSR literature.

Conclusion

At its core, this is a paper about how to most economically induce more firms to engage in CSR as well as to help investors gain maximum value from the investments made to accomplish this end. It is predicated on the assumption that resources are limited for pro-CSR stakeholders and that scholarship can provide guidance about the relative merit of different tactics. This work closely investigates the implications of how investors use rewards and punishments to influence firms to engage in CSR activities. It answers questions of which incentives (or disincentives) to use, how, and when. Specifically, this research suggests that punishments have a larger impact than many pro-CSR stakeholders might

surmise; punishments have substantial spillover effects while rewards based on rankings have built-in boundary conditions that can limit effectiveness (and tiered designs my mitigate this problem).

This work suggests several contextual factors as important. First, "reward" thresholds that separate winners and losers in distinct and limited ways (ratings/rankings) limit the reach of the reward to only those firms with capabilities to win. The implications for stakeholders attached to reward contests are to widen the opportunity for inclusion to mitigate the disincentive. As a second contextual factor, private costs that substantially impact a manager's utility truly catalyze the motivational power of a punishment. Thus, investors and stakeholders interested in increasing aggregate CSR may be better served by keeping an open mind about shifting financial capital away from reward-based approaches and towards punishment-based approaches. Doing so may stimulate more managers to engage in CSR activity.

Managerial self-interest is clearly an extremely powerful lever to consider in inducement design. Milton Friedman once expressed concern that managers might spend money on CSR, and he considered they might "misdirect" or "squander" an investor's money to further social good (1962,1970). Perhaps Friedman did not ever consider that *investors* might someday attempt to influence *managers* toward such spending (Lewis and Carlos 2019; Flammer et al. 2017, 2019). He likely did not foresee the day that investors would purposely design interventions to put managerial job security and compensation at risk to induce more CSR activity. Pro-CSR stakeholders can use this emerging behavior to their advantage. The models and practical implications herein provide rich encouragement to pro-CSR stakeholders of all types; those who want to improve positively framed inducements and those who previously considered negatively framed inducements as less efficacious than positively framed ones. This work questions the presumed superiority of positively framed inducements and encourages scholars and managers to (1) consider a wider range of influence tactics and (2) to deploy such tactics with closer attention to the alignment of motive and design.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Appendix 1

Proof Denote by $(p_1, p_2, ..., p_n)$ the probability of each player winning the prize in the first contest in some equilibria. Since each player's expected equilibrium payoff is $E[u_i^*] = \alpha v_i + w + \alpha \max\{b/m - ba_{m+1}/ma_i, 0\}$, we know that $E[s_j^*] = p_j a_j b/m + a_{m+1} b/m - a_j b/m$ for $j \in \{1, ..., m\}$ and $E[s_i^*] = p_i a_i b/m$ for all $i \in \{m+1, ..., n\}$. Since mixed strategy Nash equilibria strategies are not correlated, $E\left[\sum_{i=1}^n s_i^*\right] = \sum_{i=1}^n E[s_i^*]$. Estimate an upper bound of the socially beneficial investment from investing in the best social performers:

$$\sum_{i=1}^{n} E[s_i^*] = \frac{b}{m}\left[\sum_{i=1}^{n} p_i a_i + m a_{m+1} - \sum_{i=1}^{m} a_i\right] \tag{12}$$

The probabilities of receiving an investment must equal the number of prizes, $\sum_{i=1}^n p_i = m$ and $p_i \in [0, 1]$ for all i. Therefore, $\sum_{i=1}^n p_i a_i$ is at most $\sum_{i=1}^m a_i$ and it follows that

$$\begin{aligned}\sum_{i=1}^{n} E[s_i^*] &\le \frac{b}{m} m a_{m+1} \\ &\le b a_2.\end{aligned} \tag{13}$$

Since $a_2 \ge a_{m+1}$, the final inequality is true for all $m \ge 1$.

Denote by $(\pi_1, \pi_2, ..., \pi_n)$ the probability of each player winning the prize in the second contest in some equilibrium. Since each player's equilibrium expected payoff is $E[\widetilde{u}_i^*] = \alpha v_i - \Delta + \alpha b_i + \max\{(\Delta - \alpha b) - a_n(\Delta - \alpha b)/a_i, 0\}$ we know that $E[\widetilde{s}_i^*] = \pi_i a_i(\Delta/\alpha - b) + a_n(\Delta/\alpha - b) - a_i(\Delta/\alpha - b)$ for all $i \in \{1, ..., n-1\}$ and $E[\widetilde{s}_n^*] = \pi_n a_n(\Delta/\alpha - b)$. Now we estimate a lower bound of the socially beneficial investment from buying out the worst:

$$\sum_{i=1}^{n} E[\widetilde{s}_i^*] = \left(\frac{\Delta}{\alpha} - b\right) \quad \left[\sum_{i=1}^{n} \pi_i a_i b + (n-1) a_n - \sum_{i=1}^{n-1} a_i\right] \tag{14}$$

The probabilities of not being bought out are equal to $n - 1$, $\sum_{i=1}^n \pi_i = n - 1$ and $\pi_i \in [0, 1]$ for all i. Therefore, $\sum_{i=1}^n \pi_i a_i$ is at least $\sum_{i=2}^n a_i$ and it follows that

$$\begin{aligned}\sum_{i=1}^{n} E[\widetilde{s}_i^*] &\ge \left(\frac{\Delta}{\alpha} - b\right)\left[a_n - a_1 + (n-1)a_n\right] \\ &= \left(\frac{\Delta}{\alpha} - b\right)\left[n a_n - a_1\right]\end{aligned} \tag{15}$$

We obtain a sufficient condition for acquisition of the worst performers to motivate more costly CSR activity than investing in the best performers by setting the lower bound for acquisition of the worst performers to be greater than the upper bound for investing in the best performers.

$$\left(\frac{\Delta}{\alpha} - b\right)\left[na_n - a_1\right] \geq ba_2, \tag{16}$$

which simplifies to

$$\Delta \geq \alpha b + \alpha b \frac{a_2}{na_n - a_1}. \tag{17}$$

□

Appendix 2

Proof Denote by $(p_1, p_2, ..., p_n)$ the probability of each player winning the prize in some equilibria. Since each player's expected equilibrium payoff for the m prize contest is $E[u_i^*] = \alpha v_i + w + \alpha \max\{g_i/m - \gamma_{m+1}/ma_i, 0\}$, we know that $E[s_j^*] = p_j\gamma_j/m + \gamma_{m+1}/m - \gamma_j/m$ for $j \in \{1, ..., m\}$ and $E[s_i^*] = p_i\gamma_i/m$ for all $i \in \{m+1, ..., n\}$. Since mixed strategy Nash equilibria strategies are not correlated, $E\left[\sum_{i=1}^n s_i^*\right] = \sum_{i=1}^n E[s_i^*]$. We estimate an upper bound of the socially beneficial investment from investing in the best social performers:

$$\sum_{i=1}^{n} E[s_i^*] = \frac{1}{m}\left(\sum_{i=1}^{n} p_i\gamma_i + m\gamma_{m+1} - \sum_{j=1}^{m} \gamma_i\right) \tag{18}$$

The probabilities of receiving a reward must equal the number of prizes, $\sum_{i=1}^n p_i = m$ and $p_i \in [0, 1]$ for all i. Therefore, $\sum_{i=1}^n p_i\gamma_i$ is at most $\sum_{j=1}^m \gamma_i$ and it follows that

$$\begin{aligned}\sum_{i=1}^{n} E[s_i^*] &\leq \gamma_{m+1} \\ &\leq \gamma_2\end{aligned} \tag{19}$$

Since $\gamma_2 \geq \gamma_{m+1}$, the final inequality is true for all $m \geq 1$.

Denote by $(\pi_1, \pi_2, ..., \pi_n)$ the probability of each player not receiving the punishment in some equilibrium. Since each player's equilibrium expected payoff is $E[\widetilde{u}_i^*] = \alpha\left(v_i - c_i\right) + \alpha \max\{c_i - \xi_n/a_i, 0\}$, we know that $E[\widetilde{s}_i^*] = \pi_i\xi_i + \xi_n - \xi_i$ for all $i \in \{1, ..., n-1\}$ and $E[\widetilde{s}_n^*] = \pi_n\xi_n$. Now we estimate a lower bound of the socially beneficial investment from the punishment contest:

$$\sum_{i=1}^{n} E[\tilde{s}_i^*] = \left[\sum_{i=1}^{n} \pi_i\xi_i + (n-1)\xi_n - \sum_{i=1}^{n-1} \xi_i\right] \tag{20}$$

The probabilities of not being punished are equal to $n-1$, $\sum_{i=1}^n \pi_i = n-1$ and $\pi_i \in [0, 1]$ for all i. Therefore, $\sum_{i=1}^n \pi_i\xi_i$ is at least $\sum_{i=2}^n \xi_i$ and it follows that

$$\begin{aligned}\sum_{i=1}^{n} E[\widetilde{s}_i^*] &\geq \xi_n - \xi_1 + (n-1)\xi_n \\ &= n\xi_n - \xi_1\end{aligned} \tag{21}$$

We obtain a sufficient condition for punishment of the worst performers to create more socially responsible behavior than rewarding the best performers by setting the lower bound for punishing the worst performers to be greater than the upper bound for rewarding the best performers.

$$n\xi_n - \xi_1 \geq \gamma_2. \tag{22}$$

□

References

Adam, A. M., & Shavit, T. (2008). How can a ratings-based method for assessing corporate social responsibility (CSR) provide an incentive to firms excluded from socially responsible investment indices to invest in CSR? *Journal of Business Ethics, 82*(4), 899–905.

Aguilera, R. V., Rupp, D. E., Williams, C. A., & Ganapathi, J. (2007). Putting the S back in corporate social responsibility: A multilevel theory of social change in organizations. *Academy of Management Review, 32*(3), 836–863.

Anas, A., Rashid, H. M. A., & Annuar, H. A. (2015). The effect of award on CSR disclosures in annual reports of Malaysian PLCs. *Social Responsibility Journal, 11*(4), 831–852.

Arenas, D., Lozano, J. M., & Albareda, L. (2009). The role of NGOs in CSR: Mutual perceptions among stakeholders. *Journal of Business Ethics, 88*(1), 175–197.

Be´nabou,R., &Tirole,J. (2006). Incentives and prosocial behavior. *American Economic Review, 96*(5), 1652–1678.

Bansal, P. (2005). Evolving sustainably: A longitudinal study of corporate sustainable development. *Strategic Management Journal, 26*(3), 197–218.

Bansal, P., & Roth, K. (2000). Why companies go green: A model of ecological responsiveness. *Academy of Management Journal, 43*(4), 717–736.

Barney, J. B. (1991). Firm resources and sustained competitive advantage. *Journal of Management, 17,* 99–120.

Baron, D. P. (2001). Private politics, corporate social responsibility, and integrated strategy. *Journal of Economics & Management Strategy, 10*(1), 7–45.

Baron, D. P. (2009). A positive theory of moral management, social pressure, and corporate social performance. *Journal of Economics & Management Strategy, 18*(1), 7–43.

Baron, D. P., Harjoto, M. A., & Jo, H. (2011). The economics and politics of corporate social performance. *Business & Politics, 13*(2), 1–46.

Bartley, T., & Child, C. (2011). Movements, markets and fields: The effects of anti- sweatshop campaigns on US firms, 1993–2000. *Social Forces, 90*(2), 425–451.

Baye, M. R., Kovenock, D., & de Vries, C. G. (1996). The all-pay auction with complete information. *Economic Theory, 8,* 291–305.

Bhattacharyya, S. S., Sahay, A., Arora, A. P., & Chaturvedi, A. (2008). A toolkit for designing firm level strategic corporate social responsibility (CSR) initiatives. *Social Responsibility Journal, 4*(3), 265–282.

Binder, S., & Neumayer, E. (2005). Environmental pressure group strength and air pollution: An empirical analysis. *Ecological Economics, 55*(4), 527–538.

Chatterji, A. K., & Toffel, M. W. (2010). How firms respond to being rated. *Strategic Management Journal, 31,* 917–945.

Christensen, L. J., Peirce, E., Hartman, L. P., Hoffman, W. M., & Carrier, J. (2007). Ethics, CSR, and sustainability education in the Financial Times top 50 global business schools: Baseline data and future research directions. *Journal of Business Ethics, 73*(4), 347–368.

Clark, A., Kofford, A., Christensen, L.J., & Barney, J. (2020). Can firms with a stakeholder-oriented identity gain a competitive advantage over firms with a shareholder-oriented identity? *Working Paper*.

Clark, D. J., & Riis, C. (1998). Competition over more than one prize. *American Economic Review, 88,* 276–289.

Cording, M., Harrison, J. S., Hoskisson, R. E., & Jonsen, K. (2014). Walking the talk: A multistakeholder exploration of organizational authenticity, employee productivity, and post-merger performance. *Academy of Management Perspectives, 28*(1), 38–56.

Cuypers, I. R., Koh, P. S., & Wang, H. (2016). Sincerity in corporate philanthropy, stakeholder perceptions and firm value. *Organization Science, 27*(1), 173–188.

David, D. P., Bloom, M., & Hillman, A. J. (2007). Investor activism, managerial responsiveness, and corporate social performance. *Strategic Management Journal, 28,* 91–100.

Davidson, W. N., III., Worrell, D. L., & El-Jelly, A. (1995). Influencing managers to change unpopular corporate behavior through boycotts and divestitures: A stock market test. *Business & Society, 34*(2), 171–196.

Denis, D. J., & Serrano, J. M. (1996). Active investors and management turnover following unsuccessful control contests. *Journal of Financial Economics*, *40*(2), 239–266.

den Hond, F., & de Bakker, F. (2007). Ideologically motivated activism: How activist groups influence corporate social change activities. *Academy of Management Review, 32*(3), 901–924.

Ditlev-Simonsen, C. D., & Wenstøp, F. (2013). How stakeholders view stakeholders as CSR motivators. *Social Responsibility Journal, 9*(1), 137–147.

Elsbach, K. D., & Kramer, R. M. (1996). Members' responses to organizational identity threats: Encountering and countering the Business Week rankings. *Administrative Science Quarterly, 41*(3), 442–476.

Espeland, W. N., & Sauder, M. (2007). Rankings and reactivity: How public measures recreate social worlds. *American Journal of Sociology, 113*(1), 1–40.

Flammer, C. (2015). Does product market competition foster corporate social responsibility? Evidence from trade liberalization. *Strategic Management Journal, 36*(10), 1469–1485.

Flammer, C., Hong, B., & Minor, D. (2017). Corporate governance and the rise of integrating CSR criteria in executive compensation. In *Academy of Management Proceedings* (Vol. 2017, No. 1, p. 11206). Briarcliff Manor, NY: Academy of Management.

Flammer, C., Hong, B., & Minor, D. (2019). Corporate governance and the rise of integrating corporate social responsibility criteria in executive compensation: Effectiveness and implications for firm outcomes. *Strategic Management Journal, 40*(7), 1097–1122.

Freeman, R. E. (1984). *Strategic management: A stakeholder perspective*. Boston: Pitman.

Friedman, M. (1962). *Capitalism and freedom*. Chicago, IL: University of Chicago Press.

Friedman, M. (1970). The social responsibility of business is to increase its profits. *New York Times Magazine*, 13 September, pp. 122–126.

Fung, A., Graham, M., & Weil, D. (2007). *Full disclosure: The perils and promise of transparency*. Cambridge: Cambridge University Press.

Gillan, S. L., & Starks, L. T. (1998). A survey of shareholder activism: Motivation and empirical evidence. *Contemporary Finance Digest, 2,* 10–34.

Glac, K. (2014). The influence of shareholders on corporate social responsibility. *Economics, Management, and Financial Markets*, *9*(3), 34–79.

Godfrey, P. C. (2005). The relationship between corporate philanthropy and shareholder wealth: A risk management perspective. *Academy of Management Review, 30*(4), 777–798.

Gonzalez-Diaz, J. (2012). First-price winner-takes-all contests. *Optimization, 61*(7), 779–804.

Hart, S. L. (1995). A natural-resource-based view of the firm. *Academy of Management Review, 20*(4), 986–1014.

Hill, R. P., Ainscough, T., Shank, T., & Manullang, D. (2007). Corporate social responsibility and socially responsible investing: A global perspective. *Journal of Business Ethics, 70*(2), 165–174.

Jawahar, I. M., & McLaughlin, G. L. (2001). Toward a descriptive stakeholder theory: An organizational life cycle approach. *Academy of Management Review, 26*(3), 397–414.

Jin, G., & Leslie, P. (2003). The effect of information on product quality: Evidence from restaurant hygiene grade cards. *Quarterly Journal of Economics, 118*(2), 409–451.

Johnson, R. A., & Greening, D. W. (1999). The effects of corporate governance and institutional ownership types on corporate social performance. *Academy of Management Journal, 42*(5), 564–576.

Jones Christensen, L., Mackey, A., & Whetten, D. A. (2014). Taking responsibility for corporate social responsibility: The role of leaders in creating, implementing, sustaining, or avoiding socially responsible firm behaviors. *Academy of Management Perspectives, 28*(2), 164–178.

Kacperczyk, A. (2009). With greater power comes greater responsibility? Takeover protection and corporate attention to stakeholders. *Strategic Management Journal, 30*(3), 261–285.

Kang, J. (2016). Labor market evaluation versus legacy conservation: What factors determine retiring CEOs' decisions about long-term investment? *Strategic Management Journal, 37*(2), 389–405.

King, B. G. (2008). A political mediation model of corporate response to social movement activism. *Administrative Science Quarterly, 53*(3), 395–421.

King, B. G. (2011). The tactical disruptiveness of social movements: Sources of market and mediated disruption in corporate boycotts. *Social Problems, 58*(4), 491–517.

Kurtz, L. (2008). Socially responsible investment and shareholder activism. In A. Crane, D. Matten, A. McWilliams, J. Moon, & D. S. Siegel (Eds.), *The Oxford handbook of corporate social responsibility*. Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199211593.003.0011

Lenox, M. J., & Eesley, C. E. (2009). Private environmental activism and the selection and response of firm targets. *Journal of Economics and Management Strategy, 18*(1), 45–73.

Lepore, J. J., Mackey, A., & Mackey, T. B. (2012). Punishment versus reward in all-pay contests with perfect information. *Economics Bulletin*, *32*(4), 3090–3097.

Lewis, B. W., & Carlos, W. C. (2019). The risk of being ranked: Investor response to marginal inclusion on the 100 Best Corporate Citizens list. *Strategic Management Journal*. https://doi.org/10.1002/smj.3083

Mackey, A., Mackey, T. B., & Barney, J. B. (2007). Corporate social responsibility and firm performance: Investor preferences and corporate strategies. *Academy of Management Review, 32,* 817–835.

Madsen, P. M., & Rodgers, Z. J. (2015). Looking good by doing good: The antecedents and consequences of stakeholder attention to

corporate disaster relief. *Strategic Management Journal, 36,* 776–794.
Makov, T., & Newman, G. E. (2016). Economic gains stimulate negative evaluations of corporate sustainability initiatives. *Nature Climate Change, 6*(9), 844–846.
Martins, L. L. (2005). A model of the effects of reputational rankings on organizational change. *Organization Science, 16*(6), 701–720.
Mathiasen, C. (1994). *The SEC and social policy shareholder resolutions in the 1990s*. Washington, DC: Investor Responsibility Research Center.
McWilliams, A., & Siegel, D. (2001). Corporate social responsibility: A theory of the firm perspective. *Academy of Management Review, 26,* 117–127.
Newman, G. E., & Cain, D. M. (2014). Tainted altruism: When doing some good is evaluated as worse than doing no good at all. *Psychological Science, 25*(3), 648–655.
O'Rourke, A. (2003). A new politics of engagement: Shareholder activism for corporate social responsibility. *Business Strategy and the Environment, 12*(4), 227–239.
Otteson, J. R. (2019). *Honorable business: A framework for business in a just and humane society*. New York: Oxford University Press.
Parthiban, D. P., Kochhar, R., & Levitas, E. (1998). The Effect of institutional investors on the level and mix of CEO compensation. *Academy of Management Journal*, *41*(2), 200–208. https://doi.org/10.5465/257102.
Paul, K., & Lydenberg, S. D. (1992). Applications of corporate social monitoring systems; types, dimensions, and goals. *Journal of Business Ethics, 11*(1), 1–10.
Peteraf, M. A., & Barney, J. B. (2003). Unraveling the resource-based tangle. *Managerial Decision Economics, 24,* 309–323.
Pinder, C. C. (2014). *Work motivation in organizational behavior*. London: Psychology Press.
Pruitt, S. W., & Friedman, M. (1986). Determining the effectiveness of consumer boycotts: A stock price analysis of their impact on corporate targets. *Journal of Consumer Policy, 9*(4), 375–387.
Pruitt, S. W., Wei, K. J., & White, R. E. (1988). The impact of union-sponsored boycotts on the stock prices of target firms. *Journal of Labor Research, 9*(3), 285–289.
Rehbein, K., Waddock, S., & Graves, S. B. (2004). Understanding shareholder activism: Which corporations are targeted? *Business and Society, 43,* 239–268.
Reid, E. M., & Toffel, M. W. (2009). Responding to public and private politics: Corporate disclosure of climate change strategies. *Strategic Management Journal, 30*(11), 1157–1178.
Rondinelli, D. A., & London, T. (2003). How corporations and environmental groups cooperate: Assessing cross-sector alliances and collaborations. *Academy of Management Perspectives, 17*(1), 61–76.
Roth, A. E., & Wilson, R. B. (2019). How market design emerged from game theory: A mutual interview. *Journal of Economic Perspectives, 33*(3), 118–143.
Sadowski, M., Whitaker, K., & Buckingham, F. (2010). Rate the raters: Phase one. London: SustainAbility. https://www.ioes.ucla.edu/wp-content/uploads/2013-Delmas-Etzion-Nairn-AMP.pdf
Scalet, S., & Kelly, T. F. (2010). CSR rating agencies: What is their global impact? *Journal of Business Ethics, 94,* 69–88.
Shrivastava, P. (1995). The role of corporations in achieving ecological sustainability. *Academy of Management Review, 20*(4), 936–960.
Siegel, R. (2009). All-pay contests. *Econometrica, 77,* 71–92.
US SIF. (2018). Report on US sustainable, responsible and impact investing trends, October 2019. Downloaded from: https://www.ussif.org/.
Turban, D. B., & Greening, D. W. (1997). Corporate social performance and organizational attractiveness to prospective employees. *Academy of Management Journal, 40*(3), 658–672.
Ujah, N. U., & Okafor, C. E. (2019). The effect of a golden parachute on managed earnings. *Managerial Finance*, *45*, 925–949.
Vasi, I. B., & King, B. G. (2012). Social movements, risk perceptions, and economic outcomes: The effect of primary and secondary stakeholder activism on firms' perceived environmental risk and financial performance. *American Sociological Review, 77*(4), 573–596.
Vroom, V. H., & Jago, A. G. (1978). On the validity of the Vroom-Yetton model. *Journal of Applied Psychology, 63*(2), 151–160.
Waddock, S. (2004). Parallel universes: Companies, academics, and the progress of corporate citizenship. *Business and Society Review, 109,* 5–42.
Walsh, J. P., & Kosnik, R. D. (1993). Corporate raiders and their disciplinary role in the market for corporate control. *Academy of Management Journal*, *36*(4), 671–700. https://doi.org/10.2307/256755
Wang, H., & Qian, C. (2011). Corporate philanthropy and corporate financial performance: The roles of stakeholder response and political access. *Academy of Management Journal, 54*(6), 1159–1181.
Wood, D. J., & Jones, R. E. (1995). Stake holder mismatching: A theoretical problem in empirical research on corporate social performance. *The International Journal of Organizational Analysis*. https://doi.org/10.1108/eb028831
Wowak, A. J., Gomez-Mejia, L. R., & Steinbach, A. L. (2017). Inducements and motives at the top: A holistic perspective on the drivers of executive behavior. *Academy of Management Annals, 11*(2), 669–702.
Zedelus, C. M., Veling, H., Custers, R., Bijleveld, E., Chiew, K. S., & Aarts, H. (2014). A new perspective on human reward research: How consciously and unconsciously perceived reward information influences performance. *Cognitive Affective Behavioral Neuroscience, 14*(2), 493–508.