

# Sustainable Organizations\*

Thomas Geelen<sup>†</sup>      Jakub Hajda<sup>‡</sup>      Jan Starmans<sup>§</sup>

October 12, 2023

## Abstract

We develop a theory of stakeholder governance to study how pro-social stakeholders shape organizations. Conflicts of interest arising from diverging pro-social preferences can lead to shifts in control rights, which significantly affect organizational sustainability. We provide conditions under which more pro-social stakeholders benefit or harm an organization's sustainability, showing that they do not necessarily make it more sustainable. The key insight of our analysis is that a top-down approach to addressing sustainability concerns consistently improves organizational sustainability, while a bottom-up approach can harm it. Our results apply to CEO authority and retention, board composition, and shareholder proposals and engagement.

**Keywords:** Sustainability, ESG, stakeholders, delegation of authority, control rights, corporate governance.

**JEL Classifications:** D23, G30, L20, Q56.

---

\*We would like to thank Patrick Bolton, Martin Boyer, Alvin Chen, Alex Edmans, Peter Feldhütter, Simon Gervais, Mariassunta Giannetti, Deeksha Gupta, Martin Hellwig, Marcin Kacperczyk, Tomislav Ladika, Doron Levit, Nadya Malenko, Ulrike Malmendier, Vincent Maurin, Salvatore Miglietta (discussant), Erwan Morellec, Marcus Opp, Lasse Heje Pedersen, Enrico Perotti, Lars Persson (discussant), Georgii Riabov, Magdalena Rola-Janicka (discussant), Bogdan Stacescu (discussant), Rune Stenbacka (discussant), Vladimir Vladimirov, Paul Voss, Paul Whelan, Begum Ipek Yavuz (discussant), and Josef Zechner, seminar participants at Copenhagen Business School, HEC Montréal, Humboldt University, Johns Hopkins University, McGill University, the Stockholm School of Economics, the University of Amsterdam, and the University of Bonn, and conference participants at the 2023 Conference in Sustainable Finance at the University of Luxembourg, the 2023 Erasmus Corporate Governance Conference, the Finance Theory Group Meetings at the University of Michigan and at the University of Wisconsin-Madison, the 2023 Journal of Corporate Finance Special Issue Conference, the 2023 Nordic Initiative for Corporate Economics Conference, the 2023 Rotman Junior Finance Conference, and the 2022 Young Scholars Nordic Finance Workshop for helpful comments. Support from the Danish Finance Institute and the Jan Wallanders and Tom Hedelius Foundation is gratefully acknowledged.

<sup>†</sup>Copenhagen Business School and Danish Finance Institute. E-mail: tag.fi@cbs.dk.

<sup>‡</sup>HEC Montréal. E-mail: jakub.hajda@hec.ca.

<sup>§</sup>Stockholm School of Economics. E-mail: jan.starmans@hhs.se.

Stakeholders such as employees, managers, and investors are demonstrating a growing commitment to addressing environmental, social, and governance (ESG) issues within their organizations. This shift in stakeholder preferences can be exemplified by the dramatic increase in sustainable investing over the last decade (e.g., [US SIF, 2020](#)) and raises important questions concerning the functioning of organizations. First, how do pro-social stakeholders affect organizations? Second, do stakeholders impact organizations' sustainability differently depending on their role within the organization? Finally, in which cases can pro-social preferences lead to conflicts of interest within organizations, and how do these conflicts of interest impact how organizations are run? Understanding the conditions under which growing ESG concerns lead to more sustainable organizations is important for policymakers trying to promote organizational sustainability due to climate change concerns, for stakeholders trying to advance sustainability in their organizations, and for empirical researchers studying the link between stakeholders' ESG concerns and organizational sustainability.

To address these questions, we develop a theory of stakeholder governance in which an owner and a manager are involved in implementing a project. The owner is the controlling stakeholder who also relies on the manager in running the organization. The key insight of our paper is that while strengthening the pro-social preferences of the owner always enhances the organization's sustainability, a more pro-social manager can make an organization less sustainable. In essence, we show that a top-down approach to addressing sustainability concerns consistently improves the organization's sustainability, whereas a bottom-up approach can harm it. More specifically, we demonstrate that strengthening the pro-social preferences of a stakeholder can exacerbate conflicts of interest about sustainability policies, which can prompt the owner to withdraw the control rights from the manager. This loss of control rights and influence negatively affects the organization's sustainability if the control rights shift from a more pro-social manager to a less pro-social owner. However, the organization's sustainability benefits if the control rights shift from a less pro-social manager to a more pro-social owner. Consequently, having more pro-social stakeholders may not always improve organizations' sustainability.

To investigate how the diverging pro-social preferences of stakeholders affect organizations, we extend the delegation of authority model of [Aghion and Tirole \(1997\)](#) by incorporating the provision of a public good by the organization and by endowing its stakeholders

with pro-social preferences. In the model, the organization comprises two stakeholders—an owner and a manager—who are involved in implementing a project. Each stakeholder trades off monetary and social payoffs when deciding which project to implement. Importantly, each stakeholder may have pro-social preferences in addition to monetary incentives. Thus, rather than focusing on the production of private benefits as the existing literature does, our model features the joint provision of public and private goods. Importantly, unlike private benefits, public goods jointly affect the utility of all stakeholders with pro-social preferences. The fact that stakeholders can differ in terms of their pro-social preferences can lead to conflicts of interest about the projects they wish to implement. For example, a car rental company may have to decide whether to purchase more expensive—but less polluting—electric cars instead of gasoline cars. When selecting the type of car to purchase, the company’s owner and manager may disagree about whether to focus on profitability or environmental concerns.

The stakeholder holding the control rights has the authority to choose the project. However, a stakeholder must acquire information to implement a project and thus exercise control rights. This means that a stakeholder without formal control rights can exert effective control over the project choice by being better informed than the stakeholder who holds the control rights. For example, the owner of a car rental company who holds the control rights can decide which type of car to purchase, but choosing a car requires carefully investigating the pros and cons of different models. Suppose the manager running the day-to-day operations understands these pros and cons better than the owner. In that case, the owner should follow the manager’s advice regarding the type of car to purchase, thus granting the manager effective control over this decision.

The owner faces a trade-off when deciding whether to delegate the control rights to the manager. On the one hand, the conflict of interest about the project choice makes delegation costly. On the other hand, delegating the control rights also has benefits because it shifts the burden of costly information acquisition to the manager. Consequently, if the owner’s and the manager’s pro-social preferences differ substantially, then the cost of delegating the control rights outweighs its benefits, and the owner refrains from delegating. Conversely, if the pro-social preferences of the two stakeholders are aligned enough so that the conflict of interest between them is less severe, then the owner delegates the control rights to the manager. In particular, this view challenges the notion that a manager’s engagement in sustainability

only results in private benefits at the expense of the owners (Friedman, 1970). Instead, our framework suggests that this engagement can benefit the owners when sustainability drives effort.

The first key result of our paper is that an organization's sustainability can decrease when the manager becomes more pro-social. That is, a bottom-up approach to promoting organizational sustainability may be detrimental to the organization's sustainability. There are two cases in which this happens. The first case occurs when the manager becomes significantly more pro-social than the owner. When this happens, the owner withdraws the control rights from the more pro-social manager because the conflict of interest becomes too severe. The second case occurs when an increase in the manager's pro-social preferences results in a better alignment with the owner's preferences. In this situation, the owner delegates the control rights to the less pro-social manager to save on information acquisition costs. In both cases, the manager becomes more pro-social, and the control rights shift from the more pro-social stakeholder to the less pro-social one, harming the organization's sustainability. However, in the first case, the shift in the control rights is due to the conflict of interest becoming more severe, whereas in the second case, it results from the conflict of interest becoming less severe. Yet, in both cases, the organization's sustainability declines.

The second key result of our paper is that a shift in the control rights always benefits an organization's sustainability when it is caused by the owner becoming more pro-social. That is, a top-down approach to promoting organizational sustainability always improves the organization's sustainability. There are again two cases. First, if the owner is less pro-social than the manager, then the owner becoming more pro-social can cause the owner to delegate the control rights to the more pro-social manager. Second, if the owner is more pro-social than the manager, then this change can result in the owner withdrawing the control rights from the less pro-social manager. In both cases, the control rights shift from the less pro-social to the more pro-social stakeholder, which benefits the organization's sustainability.

We also demonstrate that changes in pro-social preferences play an important role when the control rights remain unchanged. There are two forces at play. First, a stakeholder with stronger pro-social preferences prefers a project with a higher social but a lower monetary payoff. Second, changes in pro-social preferences impact stakeholders' incentives to become informed, which can strengthen or weaken the influence of more pro-social stakeholders. For

example, we show that making the manager more pro-social relative to the owner benefits the organization's sustainability when control rights remain unchanged since the effects on both the project choice and effort make the organization more sustainable. Specifically, the manager's preferred project becomes more pro-social. Moreover, a more pro-social manager has a greater incentive to acquire information due to a higher utility level, which results in a shift in effective control from the less pro-social owner to the more pro-social manager, even though control rights remain unchanged.

In sum, we provide a theory of stakeholder governance that sheds light on how pro-social stakeholders affect organizations. Our framework allows us to identify cases in which strengthening a stakeholder's pro-social preferences benefits or harms an organization's sustainability. In particular, our results show that although a more pro-social owner always benefits an organization's sustainability, it is not always true that a more pro-social manager does. In other words, stronger pro-social preferences enhance organizational sustainability when implemented from the top down but may yield unintended negative consequences when operating from the bottom up. Moreover, we demonstrate that even minor changes in pro-social preferences can have a significant effect on an organization's sustainability because they can alter the allocation of the control rights.

We extend our model to examine the role of managers' ESG-linked compensation and owners' hiring decisions in shaping an organization's sustainability. We show that while social compensation can mitigate the conflict of interest among the stakeholders by incentivizing the manager to choose projects with a higher social payoff, it may not necessarily contribute to the overall sustainability of the organization. This negative effect arises when an increase in the manager's social compensation leads to a shift in the control rights, which can harm the organization's sustainability. Finally, we show that when choosing among managers with varying levels of pro-social preferences, the owner is inclined to hire a manager who is more but not overly pro-social. The rationale is that a more pro-social manager has greater incentives to exert effort in acquiring information that is beneficial to the owner. However, if the manager were too pro-social, the conflict of interest would become too severe, which would negatively impact the owner. Our results thus imply a theory of matching based on pro-social preferences that could be explored in future research.

Although we primarily focus on owners and managers as our main application, our find-

ings have implications for different types of relationships between stakeholders, including managers and employees, entrepreneurs and investors, as well as company boards of directors and CEOs. The central implication of our model that making stakeholders more pro-social can benefit or harm an organization’s sustainability is in line with the mixed empirical evidence regarding the impact of investors on the sustainability footprint of firms and other organizations (e.g., [Di Giuli and Kostovetsky, 2014](#); [Chen et al., 2020](#); [Heath et al., 2021](#); [Huang et al., 2021](#); [Kim et al., 2022](#)). In general, our results provide a theoretical underpinning for the increasing prevalence of stakeholder activism and engagement—stakeholders trying to obtain effective control—about sustainability issues (e.g., [Eccles and Klimenko, 2019](#)) and help interpret existing results from the empirical literature on the topic. Section IV contains a detailed discussion of the model’s implications for CEO authority and retention, board composition and dynamics, shareholder proposals, and shareholder engagement and activism, as well as suggestions for future research.

Our paper provides a theory of stakeholder society (e.g., [Tirole, 2001](#); [Allen et al., 2015](#); [Magill et al., 2015](#)) when stakeholders have pro-social preferences and contributes to several strands of the literature. First, our analysis contributes to the literature in organizational economics that studies control rights and stakeholders’ incentives to acquire information to exercise these rights.<sup>1</sup> In particular, we extend the framework of [Aghion and Tirole \(1997\)](#) by introducing project choice along two dimensions—social and monetary payoffs—and by endowing stakeholders with pro-social preferences. Our approach allows us to study how pro-social stakeholders affect organizations and, in particular, to contrast the implications of adopting a bottom-up or top-down approach to promoting sustainability in organizations. Notably, the results in our paper cannot be obtained by simply reinterpreting the private benefits for the principal and agent in [Aghion and Tirole \(1997\)](#) as capturing the owner’s and manager’s pro-social preferences, which we formally demonstrate in Appendix A. The reason is that our framework features the production of a public good, which jointly affects the utility of the owner *and* the manager. Therefore, varying the pro-social preferences of one stakeholder affects the utility of the other by affecting the provision of the public good.

---

<sup>1</sup>A related literature studies how control rights affect the communication between agents in organizations (e.g., [Dessein, 2002](#); [Harris and Raviv, 2005](#); [Alonso et al., 2008](#); [Grenadier et al., 2016](#)). See [Bolton and Dewatripont \(2013\)](#) for a survey of the literature on authority in organizations. Delegation of authority models have been used to study questions in, amongst others, corporate finance, corporate governance, and corporate culture (e.g., [Burkart et al., 1997](#); [Stein, 2002](#); [Van den Steen, 2010a,b](#); [Chen, 2022](#)).

In addition, the strength and direction of the effect on the other stakeholder depends on how much the other stakeholder values the public good (i.e., the other stakeholder’s pro-social preferences). Studying how stakeholders’ pro-social preferences affect organizations thus requires microfounding these preferences and allowing for project choice involving a public good.

Second, our paper contributes to the theoretical literature on corporate governance by studying the impact of pro-social stakeholders.<sup>2</sup> This topic is particularly important in light of increasing empirical evidence highlighting the importance of sustainability concerns in the context of corporate governance (e.g., [Di Giuli and Kostovetsky, 2014](#); [Dimson et al., 2015](#); [McCahery et al., 2016](#); [Hoepner et al., 2018](#); [Dyck et al., 2019](#); [Krueger et al., 2020](#); [Dasgupta et al., 2021](#)). In contrast to the small but growing theoretical literature on the impact of pro-social stakeholders on corporate governance ([Matsusaka and Shu, 2021](#); [Gollier and Pouget, 2022](#); [Levit et al., 2022](#); [Jin and Noe, 2023](#); [Malenko and Malenko, 2023](#)), we study a delegation problem when the organization produces a public good and stakeholders have pro-social preferences. We demonstrate that control rights and effective control are crucial dimensions for understanding the conditions under which more pro-social stakeholders benefit or harm organizational sustainability.

Finally, we contribute to the growing theoretical literature on the impact of pro-social stakeholders on organizations, which has primarily focused on investors and firms (e.g., [Heinkel et al., 2001](#); [Chowdhry et al., 2019](#); [Morgan and Tumlinson, 2019](#); [Landier and Lovo, 2020](#); [Green and Roth, 2021](#); [Roth, 2021](#); [Broccardo et al., 2022](#); [Gupta et al., 2022](#); [Hart and Zingales, 2022](#); [Oehmke and Opp, 2022](#)).<sup>3</sup> In contrast to this literature, we focus on how control rights shape stakeholder influence on organizational sustainability. In addition, our insights can be applied to a broader range of stakeholders and organizations. As stakeholders increasingly demand that organizations address ESG issues, it is important to understand how the interactions between stakeholders influence the sustainability of organizations.

---

<sup>2</sup>See [Malenko \(2022\)](#) for a survey on the literature on corporate governance.

<sup>3</sup>See [Gillan et al. \(2021\)](#) for a survey. A related literature studies the asset pricing implications of socially responsible investors (e.g., [Pástor et al., 2021](#); [Pedersen et al., 2021](#); [Pedersen and Feldhütter, 2022](#)).

# I Model

We consider an organization composed of two risk-neutral stakeholders: an owner ( $j = O$ ) and a manager ( $j = M$ ). The owner is the controlling stakeholder of the organization. There exists a set of projects that differ in terms of their social and monetary payoffs, and the organization can implement one of these projects. There are three dates without time discounting. At time zero, the owner decides whether to delegate the control rights to the manager. At time one, the owner and the manager decide how much effort to exert to become informed about the projects' payoffs. At time two, a project may be implemented if at least one stakeholder is informed. We describe the model in more detail below.

The organization's output is a pair  $(\pi, s)$ , where  $\pi$  is the monetary payoff and  $s$  is the social payoff. We refer to a payoff pair  $(\pi, s)$  as a project. The organization has one unit of initial resources that can be employed to produce the monetary and social payoffs. Let  $\iota \in [0, 1]$  denote the investment in the social payoff, then  $1 - \iota$  is the investment in the monetary payoff. There exists a production technology that generates a monetary payoff of  $\pi = \sqrt{1 - \iota}$  and a social payoff of  $s = \sqrt{\iota}$ . Consequently, the relevant set of projects is given by

$$\mathcal{P} = \{(\sqrt{1 - \iota}, \sqrt{\iota}) \mid \iota \in [0, 1]\}.$$

There is no savings technology; thus, the organization generates a zero monetary payoff and a zero social payoff if the initial resources are not employed.

The two stakeholders face an informational friction because without acquiring additional information, there is a chance of generating highly negative payoffs. As a result, if neither stakeholder is informed, then no project is implemented.<sup>4</sup> Both stakeholders can exert costly effort to become informed. Specifically, a stakeholder  $j \in \{O, M\}$  chooses a probability  $q_j \in [0, 1]$  with which the stakeholder becomes informed at time two. Stakeholder  $j$ 's private cost of effort is  $\frac{\phi_j}{2} q_j^2$ . If informed, stakeholder  $j$  can choose a project  $(\pi, s)$  from the set  $\mathcal{P}$ . The stakeholders make their effort choices simultaneously at time one, and the outcomes of the owner's and the manager's effort choices at time two are independent. An alternative interpretation of the informational friction is that there exists a search cost to identify

---

<sup>4</sup>What matters in our framework is that the owner is better off with the manager's preferred project than with a random project and vice versa for the manager. The existence of projects with highly negative payoffs is a sufficient condition for this assumption to hold.



relevant projects.

It is well documented that some stakeholders have pro-social preferences and that they may differ in terms of these preferences (e.g., [List, 2009](#); [Riedl and Smeets, 2017](#); [Bonnefon et al., 2019](#); [Bauer et al., 2021](#); [Humphrey et al., 2021](#); [Baker et al., 2022](#); [Heeb et al., 2022](#)).<sup>5</sup> Thus, we assume that stakeholder  $j$ 's utility from implementing a project  $(\pi, s)$  is

$$u_j(\pi, s) = \beta_j \pi + \gamma_j s,$$

where  $\gamma_j \geq 0$  captures the stakeholder's pro-social preferences and  $\beta_j > 0$  represents the monetary incentives. To ensure an interior equilibrium in the stakeholders' effort choices, we assume that effort is sufficiently costly:  $\phi_j > \max_{(\pi, s) \in \mathcal{P}} u_j(\pi, s)$ ,  $j \in \{O, M\}$ .

Finally, at time zero, the owner decides on the delegation of the control rights over the organization's project choice:  $d \in \{O, M\}$ . The owner either retains the control rights,  $d = O$ , or delegates them to the manager,  $d = M$ . The stakeholder holding the control rights has the authority to choose the project at time two but can also delegate the organization's project choice to the other stakeholder ex post. We refer to the stakeholder whose preferred project is implemented as the stakeholder holding effective control.<sup>6,7</sup>

Two remarks regarding the model are in order. First, as [Aghion and Tirole \(1997\)](#), we adopt an incomplete contracting approach (e.g., [Grossman and Hart, 1986](#)), assuming that project choice cannot be described and contracted on. In [Appendix B](#), we demonstrate that introducing optimal linear contracting does not resolve the conflict of interest that arises due to pro-social preferences in our setting. Second, note that while the social payoff is non-negative, this assumption does not imply that the organization generates positive externalities. What matters in the model is the difference between the social payoffs when implementing a project and when no project is implemented. Without loss of generality, we normalize the social payoff when no project is implemented to zero. We could instead assume

<sup>5</sup>For indirect evidence, see, for example, [Hong and Kacperczyk \(2009\)](#), [Hong and Kostovetsky \(2012\)](#), and [Gibson et al. \(2021\)](#).

<sup>6</sup>If the owner retains the control rights at time zero, then this is equivalent to postponing the delegation of the control rights until after the information acquisition stage.

<sup>7</sup>We assume that the owner cannot renege on the delegation decision at time two. This assumption clearly holds if the control rights are contractually agreed on. If this is not the case, the assumption can be justified if the owner incurs a reputational cost larger than  $u_O(\pi_O, s_O) - u_O(\pi_M, s_M)$  when reneging on the promise to delegate the control rights. [Baker et al. \(1999\)](#) micro-found this reputational cost in a repeated delegation of authority model.



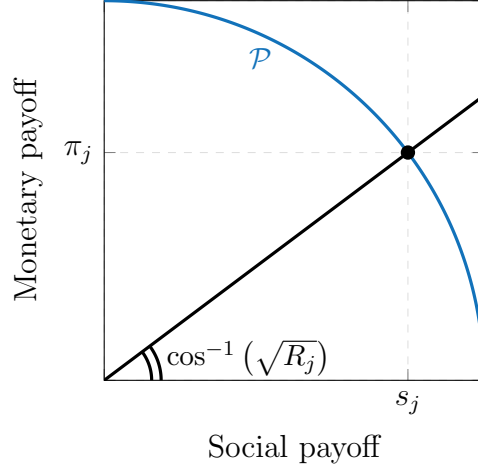


Figure 2: **Set of projects and stakeholder’s preferred project.** The figure plots the set of projects  $\mathcal{P}$  and stakeholder  $j$ ’s preferred project  $(\pi_j, s_j)$  for a given level of the stakeholder’s relative pro-social preferences  $R_j$ .

preferences is more willing to accept a lower wage or bonus if the organization generates a higher social payoff. This implication is consistent with the findings of [Krueger et al. \(2022\)](#), who show that employees in more sustainable sectors earn lower wages. The authors attribute this wage gap to employees’ preferences for environmental sustainability. Similarly, an owner with stronger pro-social preferences is willing to accept a lower financial return in exchange for a higher social return, consistent with the evidence in, for example, [Riedl and Smeets \(2017\)](#), [Bonneton et al. \(2019\)](#), [Bauer et al. \(2021\)](#), [Humphrey et al. \(2021\)](#), [Baker et al. \(2022\)](#), [Heeb et al. \(2022\)](#), [Giglio et al. \(2023\)](#).

When the owner and the manager have the same relative pro-social preferences,  $R_O = R_M$ , then their preferred projects are the same:  $(\pi_O, s_O) = (\pi_M, s_M)$ . In particular, without pro-social preferences,  $\gamma_O = \gamma_M = 0$ , their preferred projects coincide. Increasing the wedge between the stakeholders’ relative pro-social preferences,  $|R_O - R_M|$ , makes their preferred projects differ more, making the conflict of interest regarding project choice more severe. As shown below, this conflict of interest is crucial in determining who holds the control rights and, therefore, in shaping the organization’s sustainability.

## B Effort

Given the owner's and manager's preferred projects, we can determine the stakeholders' expected utilities at time one, which in turn determine their effort choices. Figure 3 summarizes the project choice at time two, depending on the allocation of the control rights and the information of the owner and the manager. For example, assume that the owner holds the control rights, depicted in Figure 3a. In this case, when informed, the owner implements the preferred project  $(\pi_O, s_O)$ . When the owner is uninformed while the manager is informed, then the owner follows the manager's recommendation and implements the manager's preferred project  $(\pi_M, s_M)$ . This means that the manager has effective control.<sup>8</sup> If neither stakeholder is informed, then no project is implemented due to the risk of generating highly negative payoffs. The other case in which the manager holds the control rights is similar, the only difference being that the manager's preferred project is implemented if both stakeholders are informed.

Stakeholder  $j$ 's expected utility at time one is given by

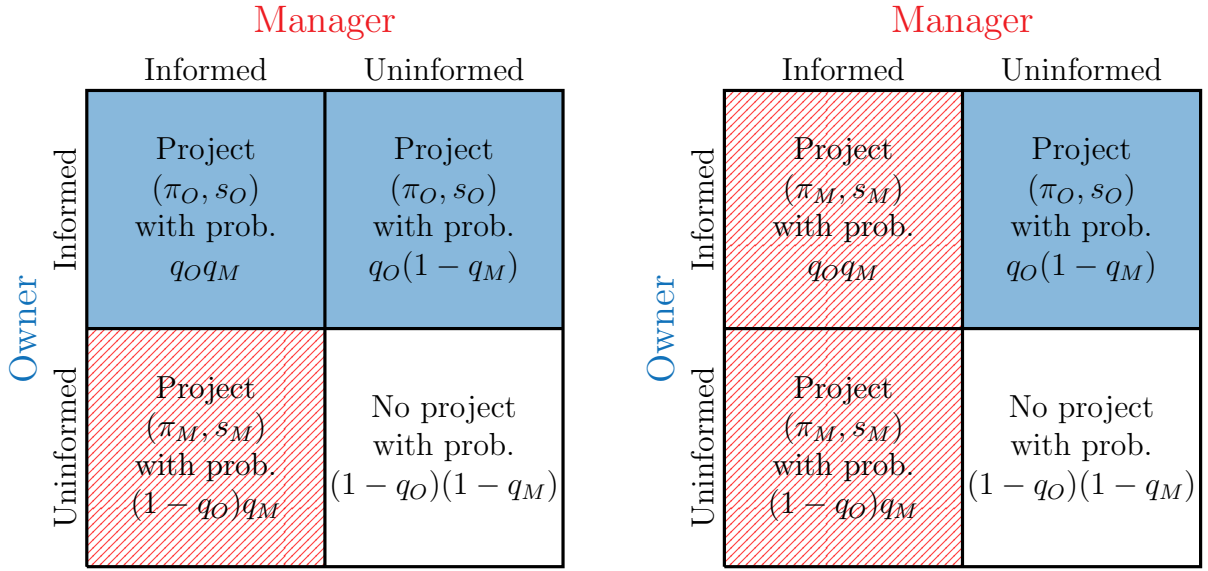
$$U_j(q_O, q_M, d) = \begin{cases} q_O u_j(\pi_O, s_O) + (1 - q_O) q_M u_j(\pi_M, s_M) - \frac{\phi_j}{2} q_j^2, & \text{if } d = O, \\ (1 - q_M) q_O u_j(\pi_O, s_O) + q_M u_j(\pi_M, s_M) - \frac{\phi_j}{2} q_j^2, & \text{if } d = M. \end{cases}$$

The owner has effective control with probability  $q_O$  when holding the control rights,  $d = O$ . In contrast, the manager has effective control with probability  $(1 - q_O)q_M$ , which occurs when the manager is informed but the owner is not. On the other hand, if the owner delegates the control rights to the manager,  $d = M$ , the probability of having effective control for the owner decreases to  $(1 - q_M)q_O$ , while the probability for the manager increases to  $q_M$ .

The stakeholders choose their effort levels simultaneously. As we show in Lemma 4 in Appendix E, the equilibrium effort choices  $(q_O(d), q_M(d))$  conditional on the delegation decision  $d$ , are determined by the two first-order conditions of the owner's and manager's expected utilities and satisfy  $(q_O(d), q_M(d)) \in (0, 1)^2$ . This implies a positive probability that each stakeholder determines the organization's project choice and thus has effective control. As a result, both the owner's and the manager's relative pro-social preferences affect the

---

<sup>8</sup>The owner follows the manager's recommendation because  $u_O(\pi_M, s_M) > 0$ . The stakeholders' relative pro-social preferences satisfy  $R_j < 1$  because  $\beta_j > 0$ . As a result, we have  $\pi_j > 0$  while  $s_j \geq 0$  and therefore  $u_{j'}(\pi_j, s_j) \geq \beta_{j'} \pi_j > 0$ , where  $j$  and  $j'$  denote the two stakeholders.



(a) Owner holds control rights:  $d = O$ .

(b) Manager holds control rights:  $d = M$ .

Figure 3: **Control rights and effective control.** This figure summarizes which project is undertaken (if any) at time two and the probability of the different cases as a function of the allocation of the control rights and the owner's and manager's effort. The background color indicates who holds effective control, where blue-filled (red-shaded) indicates that the owner (manager) has effective control.

organization's expected social payoff.

The best response function of stakeholder  $j$  is informative about the stakeholder's effort choice and is given by

$$B_j(q_{j'}, d) = \frac{(1 - q_{j'})u_j(\pi_j, s_j) + \mathbb{I}_{\{d=j\}}q_{j'}\Delta u_j}{\phi_j},$$

where  $j'$  denotes the other stakeholder and  $\Delta u_j = u_j(\pi_j, s_j) - u_j(\pi_{j'}, s_{j'}) \geq 0$ . The best response function reveals three important properties of the stakeholders' incentives to exert effort. First, the stakeholders' effort choices are strategic substitutes because  $\frac{\partial B_j(q_{j'}, d)}{\partial q_{j'}} < 0$ . Second, a stakeholder's monetary incentives and pro-social preferences affect the incentives to exert effort through the utility  $u_j(\pi, s)$  received from implementing the preferred project  $(\pi_j, s_j)$  and from the implementation of the other stakeholder's preferred project  $(\pi_{j'}, s_{j'})$ . Finally, holding the control rights increases the incentives to exert effort. This is because the stakeholder with the control rights can implement the preferred project when both stake-

holders are informed.

We can now translate the effort choices into the allocation of effective control. Specifically, we define the owner's allocation of effective control as the probability that the owner determines the organization's project choice, conditional on a project being implemented, denoted by  $e_O(d)$ .<sup>9</sup> The manager's allocation of effective control is  $e_M(d) = 1 - e_O(d)$ .

As Lemma 1 shows, in the case where the owner has no pro-social preferences,  $\gamma_O = 0$ , an increase in the manager's pro-social preferences  $\gamma_M$  unambiguously increases the manager's effort and effective control.

**Lemma 1** (Pro-social Preferences, Effort, and Effective Control). *When the owner has no pro-social preferences,  $\gamma_O = 0$ , then an increase in the manager's pro-social preferences  $\gamma_M$  leads to an increase in the manager's effort,  $\frac{\partial q_M(d)}{\partial \gamma_M} \geq 0$ , and to a decrease in the owner's effort,  $\frac{\partial q_O(d)}{\partial \gamma_M} \leq 0$ . This substitution in effort between the stakeholders translates into a shift in effective control from the owner to the manager, that is,  $\frac{\partial e_M(d)}{\partial \gamma_M} \geq 0$  and  $\frac{\partial e_O(d)}{\partial \gamma_M} \leq 0$ .*

An increase in the manager's pro-social preferences affects the manager's effort choice in two ways. First, becoming more pro-social increases the manager's level of utility and incentives to exert effort. Second, higher pro-social preferences of the manager render the conflict of interest with the owner more severe, which can further increase the manager's incentives to exert effort. Even though a more severe conflict of interest can also increase the owner's incentives to exert effort, the direct effect on the manager's utility always dominates the indirect effect on the owner's utility.

One of the key forces through which an increase in a stakeholder's pro-social preferences affects the effort level is by increasing the stakeholder's utility. Intuitively, strengthening the pro-social preferences of a stakeholder means that the stakeholder cares more about the social payoff. In particular, we argue that absolute shifts in pro-social preferences rather than relative shifts from monetary incentives to pro-social preferences are the right way to examine changes in stakeholders' pro-social preferences. For example, if a manager receives a salary and becomes more environmentally conscious, it does not mean that the manager

---

<sup>9</sup>We focus on the probability conditional on a project being implemented and not on the unconditional probability because, in practice, projects that are not undertaken cannot be observed in the data. Both stakeholders are uninformed with probability  $(1 - q_O(d))(1 - q_M(d))$ . As a result, the probability of a project being implemented is given by  $q_O(d) - q_O(d)q_M(d) + q_M(d)$  and we get  $e_O(O) = \frac{q_O(O)}{q_O(O) - q_O(O)q_M(O) + q_M(O)}$  and  $e_O(M) = \frac{(1 - q_M(M))q_O(M)}{q_O(M) - q_O(M)q_M(M) + q_M(M)}$ .

cares less about monetary income, but only that the environmental impact of the company becomes more important to the manager.<sup>10</sup>

Due to the symmetry of the model, we obtain a symmetric result when the manager has no pro-social preferences,  $\gamma_M = 0$ , in that an increase in the owner’s pro-social preferences  $\gamma_O$  leads to a substitution in effort and effective control from the manager to the owner. In Appendix C, we study how changes in stakeholders’ monetary incentives and pro-social preferences impact the equilibrium effort levels and effective control when both stakeholders have pro-social preferences.<sup>11</sup>

Given that changes in the stakeholders’ pro-social preferences alter the project choices and the allocation of effective control, it is crucial to understand how those changes jointly affect the organization’s sustainability. To this end, we examine the expected social payoff conditional on the delegation decision and a project being implemented,

$$\mathbb{E}[\tilde{s}|\tilde{\pi} > 0, d] = e_O(d)s_O + e_M(d)s_M,$$

where  $\tilde{\pi}$  and  $\tilde{s}$  are the random monetary and social payoffs and where  $\mathbb{E}$  denotes the expectation at time zero. We refer to the organization’s expected social payoff conditional on a project being implemented,  $\mathbb{E}[\tilde{s}|\tilde{\pi} > 0]$ , as the *organization’s sustainability*. Intuitively, the higher the expected social payoff, the higher the organization would score on sustainability KPIs and the more sustainable it would be deemed. Note that this result does not imply that more sustainable organizations are always desirable from a welfare perspective because higher social payoffs come at the cost of lower monetary payoffs.<sup>12</sup>

The effect of changing the manager’s pro-social preferences on the organization’s sustain-

---

<sup>10</sup>Our results on how shifts in control rights driven by changes in pro-social preferences affect an organization’s sustainability (Proposition 6 and 9) do not depend on whether we consider absolute changes in pro-social preferences (an increase in  $\gamma_M$ ) or relative changes (an increase in  $\gamma_M$  while  $\gamma_M + \beta_M = 1$ ).

<sup>11</sup>The key difference compared to the case considered here is that an increase in the manager’s pro-social preferences can also make the conflict of interest with the owner less severe, decreasing the manager’s incentives to exert effort.

<sup>12</sup>Note that if a social planner has the objective function  $\pi + s$ , then the social planner would want to implement the project  $(\pi, s) = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$  corresponding to the relative pro-social preferences  $R = \frac{1}{2}$ . Clearly, we can have  $R_j < \frac{1}{2}$  and  $R_j > \frac{1}{2}$  such that the social planner would prefer a more or a less pro-social project compared to stakeholder  $j$ .

ability, conditional on the delegation decision, is

$$\frac{\partial \mathbb{E} [\tilde{s} | \tilde{\pi} > 0, d]}{\partial \gamma_M} = \underbrace{\frac{\partial e_M(d)}{\partial \gamma_M} (s_M - s_O)}_{\Delta \text{ Effective Control}} + \underbrace{e_M(d) \frac{\partial s_M}{\partial \gamma_M}}_{\Delta \text{ Project Choice}}.$$

There are two effects at play. First, making the manager more pro-social changes effective control in the organization. This effect positively influences on the organization's sustainability when effective control shifts from a less pro-social stakeholder to a more pro-social one. The second effect is that making the manager more pro-social tilts the manager's preferred project towards one with a higher social payoff, positively influencing the organization's sustainability.

**Proposition 1** (Pro-social Preferences and Organization's Sustainability). *When the owner has no pro-social preferences,  $\gamma_O = 0$ , then an increase in the manager's pro-social preferences  $\gamma_M$  leads to a shift in effective control to the manager,  $\frac{\partial e_M(d)}{\partial \gamma_M} \geq 0$ , who chooses a more pro-social project,  $\frac{\partial s_M}{\partial \gamma_M} \geq 0$ . As a result, the organization's sustainability, conditional on the delegation decision, is increasing in the manager's pro-social preferences:*

$$\frac{\partial \mathbb{E} [\tilde{s} | \tilde{\pi} > 0, d]}{\partial \gamma_M} \geq 0.$$

When taking the delegation decision as given, Proposition 1 demonstrates that as the manager becomes more pro-social relative to the owner, the organization becomes more sustainable. In this case, both the change in effective control and the change in the project choice increase the organization's sustainability. Note that a symmetric result holds when the manager has no pro-social preferences,  $\gamma_M = 0$ , and the owner's pro-social preferences  $\gamma_O$  become stronger. That is, the organization's sustainability increases.

While Proposition 1 shows that more pro-social stakeholders can make an organization more sustainable, this result takes the delegation decision as given. However, the owner takes this decision and it can change as stakeholders become more pro-social. We study this last crucial step in the following section.



## C Delegation of Control Rights

The owner decides whether to delegate the control rights to the manager, taking into account the future actions of both stakeholders. That is, the owner solves

$$\max_{d \in \{O, M\}} U_O(q_O(d), q_M(d), d),$$

where  $q_O(d)$  and  $q_M(d)$  are the stakeholders' effort choices.

We first study the impact of the delegation decision on the effort choices of the stakeholders.

**Lemma 2** (Control Rights and Effort). *Allocating the control rights to a stakeholder increases the stakeholder's effort and reduces the other stakeholder's effort, that is,  $q_O(O) \geq q_O(M)$  and  $q_M(M) \geq q_M(O)$ .*

As Lemma 2 shows, allocating the control rights to a stakeholder increases the stakeholder's effort and reduces the other stakeholder's effort. This substitution in effort between the stakeholders translates into a shift in effective control, that is,  $e_O(O) \geq e_O(M)$  and  $e_M(M) \geq e_M(O)$ . The reason for the increase is that holding the control rights increases the likelihood of having effective control, which results in stronger incentives to exert effort.

To study the owner's delegation decision, we define the wedge in the owner's utility gained from retaining rather than delegating the control rights as

$$\Delta U_O = U_O(q_O(O), q_M(O), O) - U_O(q_O(M), q_M(M), M).$$

In particular, the owner delegates the control rights,  $d = M$ , if  $\Delta U_O < 0$  and retains the control rights,  $d = O$ , if  $\Delta U_O > 0$ .

**Proposition 2** (Irrelevance of Control Rights). *When the stakeholders have the same relative pro-social preferences,  $R_O = R_M$ , then the delegation decision  $d$  does not affect their effort choices,  $(q_O(O), q_M(O)) = (q_O(M), q_M(M))$ , and expected payoffs. As a result,  $\Delta U_O = 0$ .*

Proposition 2 shows that the owner's expected utility does not depend on the delegation decision if there is no conflict of interest between the stakeholders. The result highlights that both the pro-social preferences and the wedge in stakeholders' relative pro-social preferences are necessary to make the delegation decision relevant.

To understand how pro-social preferences affect the delegation decision, note that the wedge  $\Delta U_O$  can be rewritten as

$$\Delta U_O = \underbrace{\mathbb{P}(\tilde{\pi} > 0 | d = O) \mathbb{E} [u_O(\tilde{\pi}, \tilde{s}) | \tilde{\pi} > 0, d = O] - \frac{\phi_O}{2} q_O^2(O)}_{\text{Expected utility when } d = O} - \underbrace{\left( \mathbb{P}(\tilde{\pi} > 0 | d = M) \mathbb{E} [u_O(\tilde{\pi}, \tilde{s}) | \tilde{\pi} > 0, d = M] - \frac{\phi_O}{2} q_O^2(M) \right)}_{\text{Expected utility when } d = M}. \quad (1)$$

In equation (1), we express the owner's expected utility as the probability that a project is implemented times the expected utility conditional on a project being implemented, minus the effort cost. The equation shows that the delegation decision affects the expected utility in three ways. The first effect of delegating the control rights to the manager is that it impacts the probability of a project being undertaken, which we refer to as the *project implementation effect*. Thus, if

$$\mathbb{P}(\tilde{\pi} > 0 | d = M) > \mathbb{P}(\tilde{\pi} > 0 | d = O),$$

then the probability of a project being implemented is higher when the manager holds the control rights. This, in turn, is beneficial to the owner. In other words, delegating the control rights to the manager may reduce the risk of implementing no project.<sup>13</sup>

The second effect of delegating the control rights to the manager is an *effort cost effect*. When delegating the control rights to the manager, the owner exerts less effort, and the manager exerts more effort, which lowers the owner's effort cost by

$$\frac{\phi_O}{2} (q_O^2(O) - q_O^2(M)) \geq 0,$$

and always gives the owner a reason to delegate control rights.

The third effect of delegating the control rights to the manager is that it alters the likelihood that each stakeholder's preferred project is undertaken. Delegating the control rights to the manager increases the manager's effort and reduces the owner's effort, strengthening

---

<sup>13</sup>In Appendix F, we show that as  $\gamma_M$  gets sufficiently large, then the project implementation effect is positive in that delegation increases the probability of a project being implemented:  $\mathbb{P}(\tilde{\pi} > 0 | d = M) > \mathbb{P}(\tilde{\pi} > 0 | d = O)$ .

the manager's effective control. As a consequence, we have

$$\mathbb{E}[u_O(\tilde{\pi}, \tilde{s}) | \tilde{\pi} > 0, d = O] \geq \mathbb{E}[u_O(\tilde{\pi}, \tilde{s}) | \tilde{\pi} > 0, d = M].$$

Intuitively, delegating the control rights to the manager means that the manager's preferred project is relatively more likely to be implemented, which reduces the owner's expected utility. This effect, referred to as the *project selection effect*, discourages the owner from delegating the control rights to the manager.

**Proposition 3** (Relative Pro-social Preferences and Control Rights). *Taking as given the owner's monetary incentives  $\beta_O$  and pro-social preferences  $\gamma_O$ , if the wedge in relative pro-social preferences,  $|R_O - R_M|$ , is positive but sufficiently small, then the owner delegates the control rights to the manager:  $d = M$ . If the wedge is relatively large, then the owner retains the control rights:  $d = O$ .*

Intuitively, if the conflict of interest between the owner and the manager is minor, then the project selection effect is small. That is, the shift in effective control to the manager resulting from delegating control rights only leads to a small loss in expected utility for the owner. In this case, the value the owner obtains from the manager's increased effort and the lower effort cost dominates the project selection effect. Therefore, the owner delegates the control rights to the manager. In contrast, if the conflict of interest is severe, then the project selection effect dominates, and the owner retains the control rights.

We can fully characterize the delegation decision when one of the stakeholders has no pro-social preferences. That is, when  $\gamma_O = 0$  and  $\phi_M$  is sufficiently large, we obtain the result in Proposition 4.<sup>14</sup>

**Proposition 4** (Owner without Pro-social Preferences and Control Rights). *When the owner has no pro-social preferences,  $\gamma_O = 0$ , and  $\phi_M > \hat{\phi}_M$ , where  $\hat{\phi}_M$  is defined in the appendix, then there exists a threshold  $\hat{\gamma}_M > 0$  such that the owner delegates the control rights when  $\gamma_M \in (0, \hat{\gamma}_M)$  and retains the control rights when  $\gamma_M > \hat{\gamma}_M$ .*

Recall from Proposition 1 that the organization's sustainability, conditional on the delegation decision, is increasing in  $\gamma_M$  when  $\gamma_O = 0$ . The crucial insight from Proposition

---

<sup>14</sup>The condition  $\phi_M > \hat{\phi}_M$  in Proposition 4 ensures that if  $\gamma_M$  gets sufficiently large, then the owner wants to retain the control rights.

4 is that increasing the manager’s pro-social preferences can lead to a withdrawal of the control rights from the manager. The following result shows that this negatively affects the organization’s sustainability.

**Proposition 5** (Owner without Pro-social Preferences and Organization’s Sustainability).

When the owner has no pro-social preferences,  $\gamma_O = 0$ , and  $\phi_M > \hat{\phi}_M$ , then an increase in the manager’s pro-social preferences  $\gamma_M$  increases the organization’s sustainability for all  $\gamma_M \neq \hat{\gamma}_M$ , where  $\hat{\phi}_M$  and  $\hat{\gamma}_M$  are the thresholds from Proposition 4, that is,

$$\forall \gamma_M \neq \hat{\gamma}_M, \quad \frac{\partial \mathbb{E}[\tilde{s} | \tilde{\pi} > 0]}{\partial \gamma_M} \geq 0.$$

At  $\hat{\gamma}_M$ , the organization’s sustainability decreases discontinuously, that is,

$$\lim_{\gamma_M \uparrow \hat{\gamma}_M} \mathbb{E}[\tilde{s} | \tilde{\pi} > 0] > \lim_{\gamma_M \downarrow \hat{\gamma}_M} \mathbb{E}[\tilde{s} | \tilde{\pi} > 0].$$

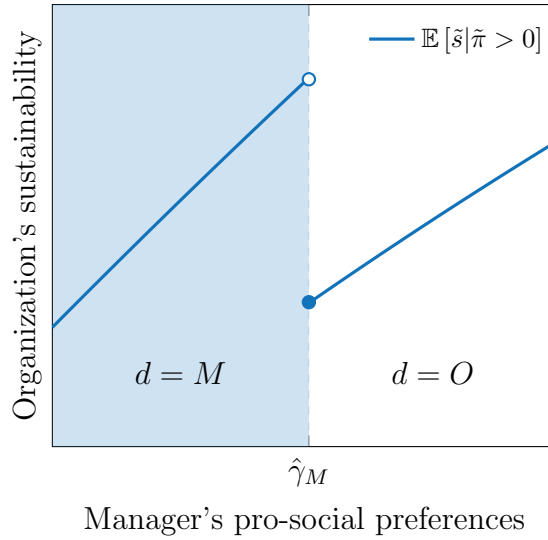


Figure 4: **Manager’s pro-social preferences and organization’s sustainability.** The figure plots the organization’s sustainability,  $\mathbb{E}[\tilde{s} | \tilde{\pi} > 0]$ , as a function of the manager’s pro-social preferences  $\gamma_M$  when  $\gamma_O = 0$ . If  $\gamma_M < \hat{\gamma}_M$ , the owner delegates the control rights, and the owner retains the control rights if  $\gamma_M > \hat{\gamma}_M$ .

Figure 4 shows the total effect of increasing the manager’s pro-social preferences from Proposition 5, taking into account the endogenous delegation decision by the owner. The figure highlights the downward jump in the organization’s sustainability as the owner with-

draws the control rights.<sup>15</sup> For example, if the CEO becomes more socially responsible, the CEO may lose the control rights and therefore have less influence on the firm, which can ultimately harm the firm’s sustainability.<sup>16</sup>

In many real-life applications, control rights are discrete. For example, the control of a firm can change around the majority-voting threshold. However, our result in Proposition 5, which demonstrates that an increase in the manager’s pro-social preferences can reduce the organization’s sustainability due to a withdrawal of the control rights, does not rely on the discrete nature of changing the control rights. In Appendix D, we extend our framework to a continuum of projects and control rights. By doing so, we demonstrate that it is not crucial that the control rights are discrete. Rather, it is important that strengthening the manager’s pro-social preferences significantly reduces the manager’s control rights, which negatively impacts the organization’s sustainability.

As we show in Proposition 6, even when the owner has pro-social preferences, a change in the control rights that results from an increase in the manager’s pro-social preferences harms the organization’s sustainability.

**Proposition 6** (Manager’s Pro-Social Preferences, Changes in Control Rights, and Organization’s Sustainability). *For a given set of preference parameters  $\{\beta_O, \gamma_O, \beta_M\}$ , assume that the delegation set is convex:  $\{\gamma_M | d = M\} = [\underline{\gamma}_M, \bar{\gamma}_M]$ . Further, assume that the  $\gamma_M$  that implies no conflict of interest,  $\{\gamma_M | R_O = R_M\}$ , is part of the feasible parameter space.<sup>17</sup> If the allocation of the control rights changes at  $\gamma'_M \in \{\underline{\gamma}_M, \bar{\gamma}_M\}$ , then the organization’s sustainability decreases discontinuously:*

$$\lim_{\gamma_M \uparrow \gamma'_M} \mathbb{E}[\tilde{s} | \tilde{\pi} > 0] > \lim_{\gamma_M \downarrow \gamma'_M} \mathbb{E}[\tilde{s} | \tilde{\pi} > 0].$$

The control rights can change at the boundary of the delegation set  $\{\underline{\gamma}_M, \bar{\gamma}_M\}$ . If they

---

<sup>15</sup>Note that the result is not driven by the fact that we consider the organization’s sustainability as the conditional expectation  $\mathbb{E}[\tilde{s} | \tilde{\pi} > 0]$ . For the unconditional expectation  $\mathbb{E}[\tilde{s}]$ , we get that  $\mathbb{E}[\tilde{s} | d = M] = q_M(M)s_M$  and  $\mathbb{E}[\tilde{s} | d = O] = q_M(O)(1 - q_O(O))s_M$ . Lemmas 1 and 2 imply that each of these expectations is non-decreasing in  $\gamma_M$  and at  $\hat{\gamma}_M$ , where the control rights shift, we have that  $\mathbb{E}[\tilde{s} | d = M] > \mathbb{E}[\tilde{s} | d = O]$ .

<sup>16</sup>Given the symmetry in our model, we can interchange the monetary and social preferences and the payoffs in all formal results to study changes in stakeholders’ monetary incentives. Note that in the model we assume that  $\beta_O > 0$  and  $\beta_M > 0$  so for these results we would assume that  $\gamma_O > 0$  and  $\gamma_M > 0$ . For example, for the result in Proposition 5, this would imply that as the manager’s monetary incentives increase, the organization’s expected profitability conditional on a project being undertaken jumps down at  $\hat{\beta}_M$ .

<sup>17</sup>In Proposition 6, we assume that if the owner is indifferent between delegating or retaining the control rights,  $\Delta U_O = 0$ , then the owner delegates the control rights.

change at  $\underline{\gamma}_M$ , then the manager is less pro-social than the owner.<sup>18</sup> At this threshold, the owner starts delegating the control rights to the manager because the conflict of interest becomes weaker when the manager's pro-social preferences increase. At  $\bar{\gamma}_M$ , the manager is more pro-social than the owner. At this threshold, the owner withdraws the control rights from the manager because the conflict of interest becomes more severe. In both cases, the control rights shift from the more pro-social stakeholder to the less pro-social one, harming the organization's sustainability. This highlights one of our key implications that more pro-social stakeholders can harm an organization's sustainability. In other words, stronger pro-social preferences may yield unintended consequences when operating from the bottom up. Importantly, in the first case, the shift in the control rights is due to a less severe conflict of interest, whereas in the second case, it results from a more severe conflict of interest. Yet, in both cases, the organization's sustainability declines.

We now analyze the case in which the owner becomes more pro-social. As we show below, changes in the control rights lead to an upward jump in the organization's sustainability, in sharp contrast to the case in which the manager becomes more pro-social. In Proposition 7, we first study the owner's delegation decision when the manager has no pro-social preferences.

**Proposition 7** (Manager without Pro-social Preferences and Control Rights). *When the manager has no pro-social preferences,  $\gamma_M = 0$ , and  $\phi_O > \hat{\phi}_O$ , where  $\hat{\phi}_O$  is defined in the appendix, then there exists a threshold  $\hat{\gamma}_O > 0$  such that the owner delegates the control rights when  $\gamma_O \in (0, \hat{\gamma}_O)$  and retains the control rights when  $\gamma_O > \hat{\gamma}_O$ .*

Proposition 7 shows that if the conflict of interest between the stakeholders is minor, then the owner delegates the control rights to the manager. If the conflict of interest is severe, then the cost of delegating the control rights is too high and the owner retains the control rights.

**Proposition 8** (Manager without Pro-social Preferences and Organization's Sustainability). *When the manager has no pro-social preferences,  $\gamma_M = 0$ , and  $\phi_O > \hat{\phi}_O$ , then an increase in the owner's pro-social preferences  $\gamma_O$  increases the organization's sustainability for all*

---

<sup>18</sup>The control rights only change at  $\underline{\gamma}_M$  if it is positive. Otherwise, the owner always delegates the control rights if  $\gamma_M$  is close to or equal to zero. Furthermore, we can show that if  $\underline{\gamma}_M > 0$ , then  $\underline{\gamma}_M < \{\gamma_M | R_O = R_M\}$  and therefore the manager is less pro-social than the owner at  $\underline{\gamma}_M$ .

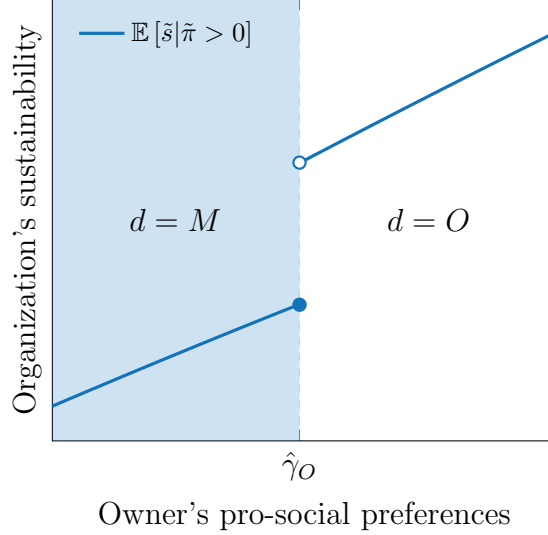


Figure 5: **Owner's pro-social preferences and organization's sustainability.** The figure plots the organization's sustainability,  $\mathbb{E}[\tilde{s}|\tilde{\pi} > 0]$ , as a function of the owner's pro-social preferences  $\gamma_O$  when  $\gamma_M = 0$ . If  $\gamma_O < \hat{\gamma}_O$ , the owner delegates the control rights, and the owner retains the control rights if  $\gamma_O > \hat{\gamma}_O$ .

$\gamma_O \neq \hat{\gamma}_O$ , where  $\hat{\phi}_O$  and  $\hat{\gamma}_O$  are the thresholds from Proposition 7, that is,

$$\forall \gamma_O \neq \hat{\gamma}_O, \quad \frac{\partial \mathbb{E}[\tilde{s}|\tilde{\pi} > 0]}{\partial \gamma_O} \geq 0.$$

At  $\hat{\gamma}_O$ , the organization's sustainability increases discontinuously, that is,

$$\lim_{\gamma_O \uparrow \hat{\gamma}_O} \mathbb{E}[\tilde{s}|\tilde{\pi} > 0] < \lim_{\gamma_O \downarrow \hat{\gamma}_O} \mathbb{E}[\tilde{s}|\tilde{\pi} > 0].$$

Proposition 8 shows that if the manager has no pro-social preferences and the owner becomes more pro-social, then the organization's sustainability always increases, even when taking into account the change in the delegation decision. The reason is that the owner withdraws the control rights from the less pro-social manager at the threshold  $\hat{\gamma}_O$ , which benefits the organization's sustainability. Figure 5 shows the overall effect of increasing the owner's pro-social preferences from Proposition 8 when taking into account the owner's endogenous delegation decision.

More generally, as Proposition 9 shows, even when the manager has pro-social preferences, a change in the control rights resulting from an increase in the owner's pro-social preferences benefits the organization's sustainability.

**Proposition 9** (Owner’s Pro-Social Preferences, Changes in Control Rights, and Organization’s Sustainability). *For a given set of preference parameters  $\{\beta_O, \beta_M, \gamma_M\}$ , assume that the delegation set is convex:  $\{\gamma_O | d = M\} = [\underline{\gamma}_O, \bar{\gamma}_O]$ . Further, assume that the  $\gamma_O$  that implies no conflict of interest,  $\{\gamma_O | R_O = R_M\}$ , is part of the feasible parameter space.<sup>19</sup> If the allocation of the control rights changes at  $\gamma'_O \in \{\underline{\gamma}_O, \bar{\gamma}_O\}$ , then the organization’s sustainability increases discontinuously:*

$$\lim_{\gamma_O \uparrow \gamma'_O} \mathbb{E}[\tilde{s} | \tilde{\pi} > 0] < \lim_{\gamma_O \downarrow \gamma'_O} \mathbb{E}[\tilde{s} | \tilde{\pi} > 0].$$

Similar to Proposition 6, the control rights can change at the boundary of the delegation set  $\{\underline{\gamma}_O, \bar{\gamma}_O\}$ . If they change at  $\underline{\gamma}_O$ , then the control rights shift from the less pro-social owner to the more pro-social manager, while at  $\bar{\gamma}_O$ , they shift from the less pro-social manager to the more pro-social owner. In both cases, the shift in the control rights benefits the organization’s sustainability.

Our results thus imply that while more pro-social owners always benefit an organization’s sustainability, more pro-social managers may not. In other words, stronger pro-social preferences enhance organizational sustainability when implemented from the *top down*, but may yield unintended negative consequences when operating from the *bottom up*. In addition, our analysis highlights that even minor changes in pro-social preferences can have significant negative effects on the sustainability of an organization due to significant changes in the allocation of control rights.

### III Extensions

In this section, we study several extensions of our model. Specifically, we focus on the manager’s social compensation (e.g., ESG-linked compensation) and the owner’s hiring decision.

---

<sup>19</sup>In Proposition 9, we assume that if the owner is indifferent between delegating or retaining the control rights,  $\Delta U_O = 0$ , then the owner delegates the control rights.



## A Social Compensation

In this section, we introduce the notion of social compensation for the manager into our model. While an optimal contracting approach is beyond the scope of this paper, we study how an exogenous and linear social compensation contract affects the organization’s outcomes.<sup>20</sup> The key insight of this section is that while social compensation can reduce the conflict of interest between the stakeholders, it can also hurt the organization’s sustainability through shifts in the control rights.

Even without explicitly introducing social compensation, the analysis in Section II.B already implies that paying for social performance may be misguided. That is because an organization’s sustainability reflects not only the owner’s choices but also the manager’s. Intuitively, if an organization’s social payoff reflects the preferences and choices of multiple stakeholders, it is unclear how individual stakeholders should be rewarded for their individual choices. This broad insight poses a challenge for designing incentive schemes for managers or employees based on ESG KPIs, which have become more prevalent recently. For example, as many as 57% of the S&P 500 firms currently evaluate their managers’ performance based on ESG metrics (e.g., Ikram et al., 2019; Semler Brossy, 2021; Cohen et al., 2022; Rajan et al., 2022).

To study the additional effects of social compensation, we extend the baseline model from Section I by providing the manager with an additional monetary compensation contract that is linear in the organization’s social payoff  $s$ :  $\alpha s$ . In this setting, the manager’s utility from a project  $(\pi, s)$  is

$$u_M(\pi, s) + \alpha s = \beta_M \pi + (\gamma_M + \alpha)s.$$

Thus, social compensation changes the manager’s effective pro-social preferences from  $\gamma_M$  to  $\gamma_M + \alpha$ . We assume that  $\gamma_M + \alpha \geq 0$ . In particular,  $\alpha = 0$  corresponds to our baseline model.

The manager’s social compensation leads to a change in the manager’s effective relative pro-social preferences. That is, if  $\alpha \neq 0$ , then

$$R_M^\alpha = \frac{(\gamma_M + \alpha)^2}{\beta_M^2 + (\gamma_M + \alpha)^2} \neq \frac{\gamma_M^2}{\beta_M^2 + \gamma_M^2} = R_M,$$

---

<sup>20</sup>See Appendix B for a discussion of optimal contracting in our framework as a tool to reduce conflicts of interest ex post.

and the manager’s preferred project becomes  $(\pi_M, s_M) = (\sqrt{1 - \iota_M}, \sqrt{\iota_M})$ , where  $\iota_M = R_M^\alpha$ . Consequently, the first implication of introducing social compensation is that if the owner can flexibly adjust the manager’s social compensation, then it allows the owner to eliminate any conflict of interest between the two stakeholders.

**Corollary 1** (Social Compensation and Project Choice). *There exists a compensation contract  $\alpha$  for the manager such that the effective relative pro-social preferences of the owner and manager are the same, that is,  $R_O = R_M^\alpha$ , and therefore the stakeholders’ preferred projects are the same, that is,  $(\pi_O, s_O) = (\pi_M, s_M)$ .*

While the social compensation we consider can reduce the conflict of interest in our model, there exist many constraints to using social compensation in reality. For example, if the manager is protected by limited liability, then the social compensation  $\alpha$  needs to be nonnegative. In this case, social compensation can only increase the manager’s effective pro-social preferences but cannot reduce them. In addition, when  $\alpha$  is positive, then compensating the manager is costly for the organization. Thus, even though social compensation may eliminate the conflict of interest, it may not be optimal to do so once the cost of compensation is considered. Social payoffs may also be hard to measure or have multiple dimensions, making compensation based on these measures potentially problematic. For example, social preferences may concern issues such as the environment, social causes, or governance, which are by themselves multi-dimensional.<sup>21</sup>

Social compensation also impacts the manager’s effort. Given that the manager’s pro-social preferences effectively become  $\gamma_M + \alpha$ , the comparative statics with respect to  $\alpha$  in the extended model are the same as those for  $\gamma_M$  in the baseline model, which we discuss in Appendix C. We can also use the results from Propositions 3 and 4 to study the effects of social compensation on the delegation decision and the organization’s sustainability. When the wedge in effective relative pro-social preferences  $|R_O - R_M^\alpha|$  is small, then the owner delegates the control rights to the manager, while when the wedge is large, then the owner retains the control rights. This implies that introducing social compensation can align stakeholders’ preferences and incentivize the owner to delegate the control rights to the manager.

---

<sup>21</sup>We can show in our model that if there are multiple social payoffs and if there exists heterogeneity in preferences across stakeholders regarding the different dimensions of these payoffs, then social compensation based on a single rating measuring the overall social performance of the organization is generally insufficient to eliminate the conflict of interest between stakeholders.

As the following result shows, an increase in the manager’s social compensation can also hurt the organization’s sustainability due to the shift in the control rights.

**Proposition 10** (Social Compensation, Control Rights, and Organization’s Sustainability).

*Assume that the owner has strong relative pro-social preferences, that is,  $\gamma_O > 0$  and  $\beta_O$  is sufficiently small. Then there exists a threshold  $\hat{\alpha}$  such that, at this threshold, the organization’s sustainability decreases discontinuously, that is,*

$$\lim_{\alpha \uparrow \hat{\alpha}} \mathbb{E}[\tilde{s} | \tilde{\pi} > 0] > \lim_{\alpha \downarrow \hat{\alpha}} \mathbb{E}[\tilde{s} | \tilde{\pi} > 0].$$

Figure 6 illustrates that the owner does not delegate the control rights to the manager if the conflict of interest is severe. When social compensation increases, then the owner starts delegating the control rights to the manager at some threshold. However, as the manager is still relatively less pro-social than the owner, this shift in the control rights reduces the organization’s sustainability.

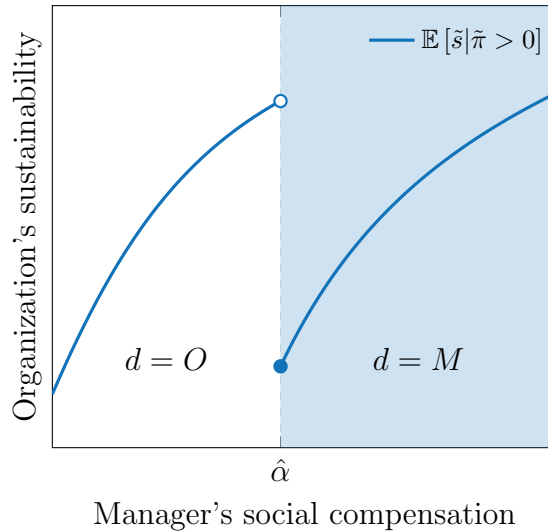


Figure 6: **Effect of manager’s social compensation on organization’s sustainability.** The figure plots the organization’s sustainability,  $\mathbb{E}[\tilde{s} | \tilde{\pi} > 0]$ , as a function of the manager’s social compensation  $\alpha$ . If  $\alpha < \hat{\alpha}$ , the owner retains the control rights, and the owner delegates the control rights if  $\alpha > \hat{\alpha}$ .

## B Hiring

This section discusses the implications of hiring a manager from a set of managers. Given that differences in relative pro-social preferences generate a conflict of interest between the owner and the manager, a crucial question is whether the owner has an incentive to hire a manager with different relative pro-social preferences. There exists growing empirical evidence documenting that ESG considerations play an important role in labor markets (e.g., [Cen et al., 2022](#); [Krueger et al., 2022](#); [Yao, 2022](#)). Our analysis suggests that these considerations are important for owners hiring managers.

One may expect the owner to hire a manager with similar pro-social preferences to avoid any conflict of interest. However, as [Proposition 11](#) shows, this may not always be the case.

**Proposition 11** (Owner without Pro-social Preferences and Hiring). *When the owner has no pro-social preferences,  $\gamma_O = 0$ , and  $\phi_M > \hat{\phi}_M$ , where  $\hat{\phi}_M$  is defined in the appendix, then there exists a threshold  $\tilde{\gamma}_M$  such that hiring a more pro-social manager improves the owner's expected utility if and only if  $\gamma_M \in (0, \tilde{\gamma}_M)$ , that is,<sup>22</sup>*

$$\frac{\partial \max_{d \in \{O, M\}} U_O(q_O(d), q_M(d), d)}{\partial \gamma_M} > 0 \quad \Leftrightarrow \quad \gamma_M \in (0, \tilde{\gamma}_M).$$

Furthermore, for  $\gamma_M \in (0, \tilde{\gamma}_M)$ , the organization's sustainability also increases, that is, for  $\gamma_M \in (0, \tilde{\gamma}_M)$ , we have  $\frac{\partial \mathbb{E}[\tilde{\pi} > 0]}{\partial \gamma_M} \geq 0$ .

Keeping the monetary incentives fixed, the higher the manager's pro-social preferences, the higher the manager's utility  $u_M$ , independent of project choice. This, in turn, increases the manager's incentives to exert effort. Intuitively, more pro-social managers are intrinsically more motivated to exert effort because they are more concerned about the organization's sustainability. When the manager is not overly pro-social, this effort effect dominates any reduction in the owner's utility due to the diverging project choice of the manager. As a result, the owner prefers to hire a pro-social manager but not an overly pro-social manager, and this hiring decision increases the organization's sustainability. Our model thus highlights that managerial hiring can benefit the sustainability of an organization. Moreover, it implies a theory of matching in labor markets based on pro-social preferences. This extension is

---

<sup>22</sup>The function  $\max_{d \in \{P, A\}} U_O(q_O(d), q_M(d), d)$  is continuously differentiable except at  $\hat{\gamma}_M$  from [Proposition 5](#), where it is only continuous.

beyond the scope of this paper but would be interesting to explore in future research.

## IV Empirical Implications

In this section, we discuss the empirical implications of our main results regarding control rights, effective control, and the sustainability of organizations. The key difference between the two stakeholders in our model lies in the controlling stakeholder’s authority to delegate control rights to the non-controlling stakeholder. While we primarily focus on owners and managers as the controlling and non-controlling stakeholders, our findings extend to various stakeholder relationships, such as managers and employees, entrepreneurs and investors, and company boards of directors and CEOs. More broadly, our results provide a theoretical underpinning for the increasing prevalence of stakeholder activism and engagement in addressing ESG issues.<sup>23</sup>

### A Control Rights

Our model demonstrates that control rights can change due to the increasing concerns of stakeholders regarding ESG issues and that the change in control rights can impact an organization’s sustainability policy. For example, our results imply that when the conflict of interest between the board and the CEO becomes more severe, the board of directors may withdraw control rights from the CEO. The board may do so by limiting the CEO’s authority, for instance, by modifying the CEO’s contract or changing corporate bylaws. In particular, dismissing a CEO can be interpreted as the most severe withdrawal of all control rights. The implication regarding shifts in control rights is consistent with the findings of [Huang et al. \(2020\)](#), who show that disagreement between investors and management is an important driver of CEO turnover.

Anecdotal evidence suggests that disagreements about pro-social policies, which correspond to the project choice in our model, can also induce turnover. For example, in 2021, Danone’s CEO Emmanuel Faber was removed from his position after an attempt to trans-

---

<sup>23</sup>See, for example, “[The investor revolution](#)” Harvard Business Review, May-June 2019, “[Let employees take the lead on ESG](#),” Wall Street Journal, June 31, 2021, “[Employees demand that we become more sustainable](#),” Forbes, October 31, 2021, and “[A catalyst for greening the financial system](#),” ECB Blog, July 8, 2022.

form Danone into a company that not only focuses on profits but also on environmental sustainability.<sup>24</sup> An example of a manager being potentially less pro-social than his employer is HSBC Asset Management’s former head of responsible investing, Stuart Kirk, who was suspended after giving a controversial speech entitled “Why investors need not worry about climate risk.” He left the bank shortly after.<sup>25</sup> A recent literature has examined the impact of CEOs’ pro-social preferences on corporate decisions, such as layoffs of workers (e.g., [Guenzel et al., 2023](#)). Our analysis in Section III further suggests that CEOs’ pro-social preferences can also influence CEO hiring and retention, which presents an interesting avenue for future research.

Our results related to control rights can also help interpret changes in board composition and dynamics. One measure of control rights is the number of board seats aligned with shareholders (e.g., [Cotter et al., 1997](#); [Nguyen and Nielsen, 2010](#)). If conflicts of interest between shareholders and management arise, shareholders may initiate a proxy fight to obtain more control over the board. The hedge fund Engine No. 1’s engagement with ExxonMobil’s management is an example of how pro-social shareholders may attempt to gain control rights. Engine No. 1’s demands resulted in the election of three new directors to ExxonMobil’s board, which challenged the company’s existing strategy and pushed for a transition towards renewable energy.<sup>26</sup> While prior research has studied the role of pro-social shareholders in corporate governance (e.g., [Di Giuli and Kostovetsky, 2014](#); [Dimson et al., 2015](#); [McCahery et al., 2016](#); [Hoepner et al., 2018](#); [Dyck et al., 2019](#); [Krueger et al., 2020](#); [Dasgupta et al., 2021](#)), the effect of pro-social stakeholders on board composition in terms of directors’ pro-social preferences presents another interesting avenue for future research given the important role of boards in determining corporate policies.

Another important channel through which stakeholders exercise their control rights is through shareholder proposals. In line with the mechanism in our model, [Kim et al. \(2019\)](#) document that firms act pro-socially at the request of their stakeholders. Consistent with this notion, the authors document that local institutional investors exert significant influence on the environmental policies of firms via shareholder proposals. Similarly, [Chen et al. \(2020\)](#) show that institutional shareholders use the same channel to generate improvements

---

<sup>24</sup>See “[A top CEO was ousted after making his company more environmentally conscious. Now he’s speaking out,](#)” Time Magazine, 21 November 2021.

<sup>25</sup>See “[HSCB banker quits over climate change furore,](#)” Financial Times, 7 July 2022.

<sup>26</sup>See “[Signs of change at ExxonMobil a year after hedge fund proxy fight,](#)” Financial Times, 24 May 2022.

in social impact outcomes, while [He et al. \(2023\)](#) provide supporting evidence based on mutual fund votes. Finally, [Huang et al. \(2021\)](#) provide causal evidence for this channel by documenting that a higher interest of institutional investors regarding ESG issues translates into more pro-social voting patterns. However, our analysis also suggests that exercising control rights through shareholder voting can negatively affect firms' sustainability efforts and outcomes if it is driven by shareholders becoming less pro-social. While the existing literature seems to suggest that the average effect of shareholder voting on firms' sustainability is positive, it would be interesting to explore the cross-sectional heterogeneity in the effect and to understand the determinants of when shareholder voting makes firms more or less sustainable.

## **B Effective Control**

While formal control rights in the form of CEO contracts or board representation have important implications for sustainability outcomes, informal forms of control or soft power, which correspond to effective control in our model, also matter in ESG-related shareholder engagement. Consistent with our model's notion of effective control, [Dimson et al. \(2015, 2021\)](#) document that shareholders exert informal influence on firms by persuading firms to address environmental and social issues. Similarly, [McCahery et al. \(2016\)](#) show that institutional investors exert informal influence on firms. [Hoepner et al. \(2018\)](#) show that shareholder engagement on ESG issues can benefit shareholders by reducing firms' downside risks. Institutional investors play a particularly important role in shareholder engagement on ESG issues through exercising effective control. For example, [Di Giuli and Kostovetsky \(2014\)](#) document that political convictions of firms' stakeholders determine firms' corporate social responsibility (CSR) spending and rating. [Azar et al. \(2021\)](#) provide evidence consistent with the idea that engagement by investors such as BlackRock, Vanguard, and State Street Global Advisors reduces firms' carbon emissions. This type of engagement is often informal, such as private meetings with management. One of the main channels through which this type of engagement affects corporate ESG policies is by diffusing new ESG knowledge among companies and investors ([UN PRI, 2018](#)). This diffusion of ESG knowledge is consistent with the informational friction present in our framework.

## C The Sustainability of Organizations

One of the key implications of our paper is that more pro-social stakeholders can benefit or harm the sustainability of organizations, driven by changes in control rights and effective control. An example of the negative impact of a more pro-social non-controlling stakeholder (relative to the controlling stakeholder) is the recent controversy surrounding BlackRock's ESG strategy. Several American states have withdrawn investment mandates from BlackRock over the concern that its increasingly important ESG strategy will have a negative impact on investor returns. Consequently, while BlackRock's strategy became more pro-social, its ability to influence the efforts of its portfolio companies to promote sustainability has been reduced.<sup>27</sup>

We also show that the change in control rights resulting from stakeholders becoming more pro-social can, in certain cases, increase an organization's sustainability. Importantly, our results imply that while more pro-social non-controlling stakeholders, such as managers, may not always benefit an organization's sustainability, more pro-social controlling stakeholders, such as firm owners, always do. An example of the positive effect of a principal becoming more pro-social is the case of the hedge fund Engine No. 1 discussed earlier. Their engagement led to pro-social shareholders obtaining more control of ExxonMobile, allowing them to push the company to adopt a more sustainable strategy.

The ambiguous effect of more pro-social stakeholders on the sustainability of organizations is also in line with the mixed empirical evidence regarding the impact of investors on the sustainability footprint of firms and other organizations. [Kim et al. \(2022\)](#) document the potentially negative effect by showing that issuing ESG-linked loans can lead to a deterioration in ESG scores. [Heath et al. \(2021\)](#) show that there is no effect by demonstrating that socially responsible investment funds do not improve the behavior of their portfolio companies in terms of environmental and social outcomes. In contrast, other studies document a positive effect of stakeholders on firms' sustainability (e.g., [Di Giuli and Kostovetsky, 2014](#); [Chen et al., 2020](#); [Huang et al., 2021](#); [Gantchev et al., 2022](#)). Our analysis highlights that when assessing the impact of pro-social stakeholders, it is crucial to determine if it is associated with a change in control rights and whether the change in control rights is driven by more

---

<sup>27</sup>See "[Florida to pull \\$2bn from BlackRock in spreading ESG backlash,](#)" Financial Times, December 1, 2022.



pro-social controlling or non-controlling stakeholders, as this informs whether the impact on a firm's sustainability is positive or negative. This conceptual framework can guide future research exploring the effect of pro-social stakeholders on organizations' sustainability.

## V Conclusion

We develop a theory of stakeholder governance to study how stakeholders with pro-social preferences influence an organization's sustainability. Our analysis highlights that while strengthening the pro-social preferences of the owner always enhances the organization's sustainability, a more pro-social manager can make an organization less sustainable. In essence, we show that a top-down approach to addressing sustainability concerns consistently improves the organization's sustainability, whereas a bottom-up approach can falter. In addition to analyzing how pro-social stakeholders impact control rights, effective control, and organizations' sustainability, we also study how social compensation and hiring affect organizations. We find that social compensation can negatively impact an organization's sustainability and that organizations prefer to hire more but not overly pro-social managers.

Our analysis can be applied to different types of relationships between stakeholders such as managers and employees, entrepreneurs and investors, and company boards and CEOs. In general, our model provides a theoretical underpinning for the increasing prevalence of ESG-related stakeholder activism and engagement.

## References

- Aghion, P. and Tirole, J. (1997). Formal and real authority in organizations. *Journal of Political Economy*, 105(1):1–29.
- Allen, F., Carletti, E., and Marquez, R. (2015). Stakeholder governance, competition, and firm value. *Review of Finance*, 19(3):1315–1346.
- Alonso, R., Dessein, W., and Matouschek, N. (2008). When does coordination require centralization? *American Economic Review*, 98(1):145–79.
- Azar, J., Duro, M., Kadach, I., and Ormazabal, G. (2021). The big three and corporate carbon emissions around the world. *Journal of Financial Economics*, 142(2):674–696.
- Baker, G., Gibbons, R., and Murphy, K. J. (1999). Informal authority in organizations. *Journal of Law, Economics, and Organization*, 15(1):56–73.
- Baker, M., Egan, M. L., and Sarkar, S. K. (2022). How do investors value ESG? *Working Paper*.
- Bauer, R., Ruof, T., and Smeets, P. (2021). Get real! individuals prefer more sustainable investments. *Review of Financial Studies*, 34(8):3976–4043.
- Bolton, P. and Dewatripont, M. (2013). Authority in organizations. *Handbook of Organizational Economics*, pages 342–372.
- Bonnefon, J.-F., Landier, A., Sastry, P., and Thesmar, D. (2019). Do investors care about corporate externalities? Experimental evidence. *Working Paper, HEC Paris*.
- Broccardo, E., Hart, O., and Zingales, L. (2022). Exit vs. voice. *Forthcoming, Journal of Political Economy*.
- Burkart, M., Gromb, D., and Panunzi, F. (1997). Large shareholders, monitoring, and the value of the firm. *Quarterly Journal of Economics*, 112(3):693–728.
- Cen, X., Qiu, Y., and Yue Wang, T. (2022). Corporate social responsibility and employee retention. *Working Paper, Texas A&M University*.
- Chen, J. (2022). Optimal managerial authority. *Working Paper, Boston University*.
- Chen, T., Dong, H., and Lin, C. (2020). Institutional shareholders and corporate social responsibility. *Journal of Financial Economics*, 135(2):483–504.
- Chowdhry, B., Davies, S. W., and Waters, B. (2019). Investing for impact. *Review of Financial Studies*, 32(3):864–904.
- Cohen, S., Kadach, I., Ormazabal, G., and Reichelstein, S. (2022). Executive compensation tied to ESG performance: International evidence. *Working Paper, Stanford University*.
- Cotter, J. F., Shivdasani, A., and Zenner, M. (1997). Do independent directors enhance target shareholder wealth during tender offers? *Journal of Financial Economics*, 43(2):195–218.

- Dasgupta, A., Fos, V., and Sautner, Z. (2021). Institutional investors and corporate governance. *Foundations and Trends in Finance*, 12(4):276–394.
- Dessein, W. (2002). Authority and communication in organizations. *Review of Economic Studies*, 69(4):811–838.
- Di Giuli, A. and Kostovetsky, L. (2014). Are red or blue companies more likely to go green? Politics and corporate social responsibility. *Journal of Financial Economics*, 111(1):158–180.
- Dimson, E., Karakaş, O., and Li, X. (2015). Active ownership. *Review of Financial Studies*, 28(12):3225–3268.
- Dimson, E., Karakaş, O., and Li, X. (2021). Coordinated engagements. *Working Paper, University of Cambridge*.
- Dyck, A., Lins, K. V., Roth, L., and Wagner, H. F. (2019). Do institutional investors drive corporate social responsibility? International evidence. *Journal of Financial Economics*, 131(3):693–714.
- Eccles, R. G. and Klimenko, S. (2019). The investor revolution. *Harvard Business Review*.
- Friedman, M. (1970). The social responsibility of business is to increase its profits. *New York Times Magazine*.
- Gantchev, N., Giannetti, M., and Li, R. (2022). Does money talk? Divestitures and corporate environmental and social policies. *Review of Finance*, 26(6):1469–1508.
- Gibson, R., Krueger, P., and Mitali, S. (2021). The sustainability footprint of institutional investors: ESG driven price pressure and performance. *Working Paper, University of Geneva*.
- Giglio, S., Maggiori, M., Stroebel, J., Tan, Z., Utkus, S., and Xu, X. (2023). Four facts about ESG beliefs and investor portfolios. *Working Paper*.
- Gillan, S. L., Koch, A., and Starks, L. T. (2021). Firms and social responsibility: A review of ESG and CSR research in corporate finance. *Journal of Corporate Finance*, 66:101889.
- Gollier, C. and Pouget, S. (2022). Investment strategies and corporate behaviour with socially responsible investors: A theory of active ownership. *Economica*, 89(356):997–1023.
- Green, D. and Roth, B. (2021). The allocation of socially responsible capital. *Working Paper, Harvard Business School*.
- Grenadier, S. R., Malenko, A., and Malenko, N. (2016). Timing decisions in organizations: Communication and authority in a dynamic environment. *American Economic Review*, 106(9):2552–81.
- Grossman, S. J. and Hart, O. D. (1986). The costs and benefits of ownership: A theory of vertical and lateral integration. *Journal of Political Economy*, 94(4):691–719.
- Guenzel, M., Hamilton, C., and Malmendier, U. (2023). CEO social preferences and layoffs. *Working Paper, University of Pennsylvania*.

- Gupta, D., Kopytov, A., and Starmans, J. (2022). The pace of change: Socially responsible investing in private markets. *Working Paper, Stockholm School of Economics*.
- Harris, M. and Raviv, A. (2005). Allocation of decision-making authority. *Review of Finance*, 9(3):353–383.
- Hart, O. and Zingales, L. (2022). The new corporate governance. *Working Paper, University of Chicago*.
- He, Y., Kahraman, B., and Lowry, M. (2023). ES risks and shareholder voice. *Forthcoming, Review of Financial Studies*.
- Heath, D., Macciocchi, D., Michaely, R., and Ringgenberg, M. C. (2021). Does socially responsible investing change firm behavior? *Forthcoming, Review of Finance*.
- Heeb, F., Kölbel, J. F., Paetzold, F., and Zeisberger, S. (2022). Do investors care about impact? *Forthcoming, Review of Financial Studies*.
- Heinkel, R., Kraus, A., and Zechner, J. (2001). The effect of green investment on corporate behavior. *Journal of Financial and Quantitative Analysis*, 36(4):431–449.
- Hoepner, A. G., Oikonomou, I., Sautner, Z., Starks, L. T., and Zhou, X. (2018). ESG shareholder engagement and downside risk. *Working Paper, University of Texas at Austin*.
- Hong, H. and Kacperczyk, M. (2009). The price of sin: The effects of social norms on markets. *Journal of Financial Economics*, 93(1):15–36.
- Hong, H. and Kostovetsky, L. (2012). Red and blue investing: Values and finance. *Journal of Financial Economics*, 103(1):1–19.
- Huang, S., Maharjan, J., and Thakor, A. V. (2020). Disagreement-induced CEO turnover. *Journal of Financial Intermediation*, 43:100819.
- Huang, W., Karolyi, G. A., and Kwan, A. (2021). Paying attention to ESG matters: Evidence from big data analytics. *Working Paper, Hong Kong University Business School*.
- Humphrey, J., Kogan, S., Sagi, J., and Starks, L. (2021). The asymmetry in responsible investing preferences. *Working Paper*.
- Ikram, A., Li, Z. F., and Minor, D. (2019). CSR-contingent executive compensation contracts. *Forthcoming, Journal of Banking and Finance*.
- Jin, D. and Noe, T. H. (2023). Making sure your vote does not count: ESG activism, pseudo-green shareholders, and insincere proxy voting. *Working Paper, University of Oxford*.
- Kim, I., Wan, H., Wang, B., and Yang, T. (2019). Institutional investors and corporate environmental, social, and governance policies: Evidence from toxics release data. *Management Science*, 65(10):4901–4926.
- Kim, S., Kumar, N., Lee, J., and Oh, J. (2022). ESG lending. *Working Paper, University of Florida*.

- Krueger, P., Metzger, D., and Wu, J. (2022). The sustainability wage gap. *Working Paper, University of Geneva*.
- Krueger, P., Sautner, Z., and Starks, L. T. (2020). The importance of climate risks for institutional investors. *Review of Financial Studies*, 33(3):1067–1111.
- Landier, A. and Lovo, S. (2020). ESG investing: How to optimize impact? *Working Paper, HEC Paris*.
- Levit, D., Malenko, N., and Maug, E. G. (2022). Trading and shareholder democracy. *Forthcoming, Journal of Finance*.
- List, J. A. (2009). Social preferences: Some thoughts from the field. *Annual Review of Economics*, 1(1):21–1.
- Magill, M., Quinzii, M., and Rochet, J.-C. (2015). A theory of the stakeholder corporation. *Econometrica*, 83(5):1685–1725.
- Malenko, A. and Malenko, N. (2023). Voting choice. *Working Paper, Boston College*.
- Malenko, N. (2022). Information flows, organizational structure, and corporate governance. *Working Paper, Boston College*.
- Matsusaka, J. G. and Shu, C. (2021). A theory of proxy advice when investors have social goals. *Working Paper, University of Southern California*.
- McCahery, J. A., Sautner, Z., and Starks, L. T. (2016). Behind the scenes: The corporate governance preferences of institutional investors. *Journal of Finance*, 71(6):2905–2932.
- Morgan, J. and Tumlinson, J. (2019). Corporate provision of public goods. *Management Science*, 65(10):4489–4504.
- Nguyen, B. D. and Nielsen, K. M. (2010). The value of independent directors: Evidence from sudden deaths. *Journal of Financial Economics*, 98(3):550–567.
- Oehmke, M. and Opp, M. M. (2022). A theory of socially responsible investment. *Working Paper, Stockholm School of Economics*.
- Pástor, L., Stambaugh, R. F., and Taylor, L. A. (2021). Sustainable investing in equilibrium. *Journal of Financial Economics*, 142(2):550–571.
- Pedersen, L. H. and Feldhütter, P. (2022). Is capital structure irrelevant with ESG investors? *Working Paper, Copenhagen Business School*.
- Pedersen, L. H., Fitzgibbons, S., and Pomorski, L. (2021). Responsible investing: The ESG-efficient frontier. *Journal of Financial Economics*, 142(2):572–597.
- Rajan, R. G., Ramella, P., and Zingales, L. (2022). What purpose do corporations purport? Evidence from letters to shareholders. *Working Paper, University of Chicago*.
- Riedl, A. and Smeets, P. (2017). Why do investors hold socially responsible mutual funds? *Journal of Finance*, 72(6):2505–2550.

- Roth, B. (2021). Impact investing: A theory of financing social enterprises. *Working Paper, Harvard Business School*.
- Semler Brossy (2021). ESG+Incentives 2021 report.
- Stein, J. C. (2002). Information production and capital allocation: Decentralized versus hierarchical firms. *Journal of Finance*, 57(5):1891–1921.
- Tirole, J. (2001). Corporate governance. *Econometrica*, 69(1):1–35.
- UN PRI (2018). How ESG engagement creates value for investors and companies. *United Nations Principles for Responsible Investing*.
- US SIF (2020). Report on US Sustainable and Impact Investing Trends 2020.
- Van den Steen, E. (2010a). Culture clash: The costs and benefits of homogeneity. *Management Science*, 56(10):1718–1738.
- Van den Steen, E. (2010b). On the origin of shared beliefs (and corporate culture). *The RAND Journal of Economics*, 41(4):617–648.
- Yao, T. (2022). ESG education and job choice. *Working Paper, HEC Paris*.

# Appendix

Appendix A provides a formal discussion of the relation to [Aghion and Tirole \(1997\)](#). Appendix B shows that optimal contracting cannot resolve the conflict of interest between the stakeholders. Appendix C discusses the impact of monetary incentives and pro-social preferences on effort and effective control. Appendix D shows that our results do not depend on the discrete nature of the allocation of control rights. Appendix E contains the proofs. Appendix F discusses the project implementation effect.

## A Relation to [Aghion and Tirole \(1997\)](#)

In this appendix, we formally demonstrate the difference between our framework and [Aghion and Tirole \(1997\)](#). Specifically, we show why microfounding stakeholders’ pro-social preferences and allowing for a project choice with monetary and social payoffs is necessary to study how shifts in pro-social preferences affect organizational sustainability.

Table A.1 summarizes the payoff structure in [Aghion and Tirole \(1997\)](#). Note that we use the terms “owner” and “manager” instead of “principal” and “agent”, respectively. In [Aghion and Tirole \(1997\)](#), the parameters  $B \geq 0$  and  $b \geq 0$  describe the level of the owner’s and manager’s utility when their preferred projects are implemented, respectively, while the parameters  $\alpha \in [0, 1]$  and  $\beta \in [0, 1]$  capture the degree of the conflict of interests between the owner and the manager.

	Owner’s preferred project	Manager’s preferred project
Owner’s utility	$B$	$\alpha B$
Manager’s utility	$\beta b$	$b$

Table A.1: **Payoff structure in [Aghion and Tirole \(1997\)](#)**. Note that we use the terms “owner” and “manager” instead of “principal” and “agent,” respectively.

Table A.2 summarizes the payoff structure in our framework, taking into account the endogenous preferred projects of the owner and the manager  $(\pi_O, s_O)$  and  $(\pi_M, s_M)$ , respectively.

	Owner’s preferred project	Manager’s preferred project
Owner’s utility	$u_O(\pi_O, s_O) = \sqrt{\beta_O^2 + \gamma_O^2}$	$u_O(\pi_M, s_M) = \frac{\beta_O \beta_M + \gamma_O \gamma_M}{\sqrt{\beta_M^2 + \gamma_M^2}}$
Manager’s utility	$u_M(\pi_O, s_O) = \frac{\beta_O \beta_M + \gamma_O \gamma_M}{\sqrt{\beta_O^2 + \gamma_O^2}}$	$u_M(\pi_M, s_M) = \sqrt{\beta_M^2 + \gamma_M^2}$

Table A.2: **Payoff structure in our framework.**

The primary advantage of our framework is that it allows us to study the effect of shifts in control rights driven by changes in pro-social preferences on an organization’s sustainability. [Aghion and Tirole \(1997\)](#) allows us to study when control rights shift as any preference

parameter  $b$ ,  $B$ ,  $\alpha$ , or  $\beta$  changes. As Table A.1 and A.2 make clear, there is no straightforward way to map Aghion and Tirole (1997)'s preference parameters into our framework.

The fact that Aghion and Tirole (1997) microfound neither the project choice nor the preferences across the different project dimensions mean that many economic concepts related to sustainability cannot be defined in their framework. For example, in Aghion and Tirole (1997), there is no notion of when one stakeholder is more pro-social than the other stakeholder. In our setting, the owner is more pro-social than the manager in the sense that they choose a project with a higher social payoff if and only if

$$R_O = \frac{\gamma_O^2}{\gamma_O^2 + \beta_O^2} > \frac{\gamma_M^2}{\gamma_M^2 + \beta_M^2} = R_M.$$

Importantly, the case  $R_O > R_M$  can be consistent with both  $u_O(\pi_O, s_O) > u_M(\pi_M, s_M)$  and  $u_O(\pi_O, s_O) < u_M(\pi_M, s_M)$  such that one cannot simply map the ranking of stakeholders in terms of their relative pro-social preferences to a ranking of  $B$  and  $b$  in Aghion and Tirole (1997). That is, one cannot interpret the stakeholder with the higher or lower utility as the more pro-social stakeholder. Microfounding project choice with a monetary payoff  $\pi$  and a social payoff  $s$  and the stakeholders' preferences over these two dimensions therefore allows us to address the question of whether shifts in control rights are driven by stakeholders becoming more or less pro-social and how this impacts the organization's sustainability.

In addition, the microfoundations in our model introduce new economic forces, which are absent in Aghion and Tirole (1997) but are crucial for the mechanism in our paper. To highlight the rich interactions that arise when stakeholders' pro-social preferences change, consider the manager becoming more pro-social. The first effect is that the manager's preferred project,  $(\pi_M, s_M)$ , changes. Specifically, the manager chooses a project with a higher social and lower monetary payoff. That is,

$$\frac{\partial \pi_M}{\partial \gamma_M} < 0 \quad \text{and} \quad \frac{\partial s_M}{\partial \gamma_M} > 0.$$

As a result, the manager's utility increases when the manager's preferred project is implemented. That is,

$$\frac{\partial u_M(\pi_M, s_M)}{\partial \gamma_M} = \beta_M \frac{\partial \pi_M}{\partial \gamma_M} + \gamma_M \frac{\partial s_M}{\partial \gamma_M} + s_M > 0,$$

and the manager's utility when the owner's preferred project is implemented (weakly) increases. That is,

$$\frac{\partial u_M(\pi_O, s_O)}{\partial \gamma_M} = s_O \geq 0.$$

In particular, the extent to which the manager's utility increases depends on  $s_O$ , which is in turn determined by the owner's pro-social preferences  $\gamma_O$ .

Crucially, the manager's increasing pro-social preferences not only impact the manager's utility but also influence the owner's utility due to the public good nature of the social payoff.



Specifically, when the manager has effective control, the owner's utility changes. That is,

$$\frac{\partial u_O(\pi_M, s_M)}{\partial \gamma_M} = \beta_O \frac{\partial \pi_M}{\partial \gamma_M} + \gamma_O \frac{\partial s_M}{\partial \gamma_M}.$$

Importantly, whether the owner's utility increases when the manager becomes more pro-social depends on whether the other owner is more or less pro-social than the manager. That is, if  $R_O > R_M$ , then  $\frac{\partial u_O(\pi_M, s_M)}{\partial \gamma_M} > 0$ , and if  $R_O < R_M$ , then  $\frac{\partial u_O(\pi_M, s_M)}{\partial \gamma_M} < 0$ . The reason is that making one stakeholder more pro-social can reduce or amplify the conflict of interest between the two stakeholders.

In summary, if the manager becomes more pro-social, then the manager's utility levels  $u_M(\pi_M, s_M)$  and  $u_M(\pi_O, s_O)$  (weakly) increase, where the slope of the latter depends on the owner's pro-social preferences. In addition, the owner's utility  $u_O(\pi_M, s_M)$  changes, with the sign depending on whether the owner is more pro-social or less pro-social than the manager in terms of their relative pro-social preferences  $R_M$  and  $R_O$ . Comparing Tables A.1 and A.2, it is clear that there is no way in which one can map the comparative statics with respect to pro-social preferences to the parameters in Aghion and Tirole (1997).

As discussed in Section II, the owner's and manager's utilities from Table A.2 determine their effort choices. In turn, the conflict of interest and the equilibrium effort levels determine the allocation of the control rights and effective control in our model. Thus, studying how pro-social preferences affect organizational sustainability requires microfounding social payoffs and pro-social preferences. The question can, therefore, not be studied in the context of Aghion and Tirole (1997).

## B Contracting on Project Choice

In this appendix, we show that it is too costly for the owner to fully resolve the conflict of interest with the manager using linear contracts ex post. Consequently, introducing optimal contracting ex post does not resolve the conflict of interest that arises due to pro-social preferences in our model. Note that focusing on optimal contracting ex post is the most flexible form of contracting as it can condition on the allocation of effective control. Specifically, if the owner has effective control, then the owner has no reason to offer a contract to the manager. Since ex-post contracting can consider this contingency, it is more flexible and, therefore, superior to ex-ante contracting, assuming that the contract cannot condition on the allocation of effective control.

To study ex-post optimal contracting, we assume that the owner can offer a contract  $\alpha = (\alpha_\pi, \alpha_s) \geq 0$  to the manager at time 2 when the manager has effective control. The owner can use this contract to induce the manager to choose a project closer to the owner's preferred project, reducing the conflict of interest. The contract  $\alpha$  yields the manager a payoff in utility terms of

$$\alpha_\pi \pi + \alpha_s s,$$

and therefore the manager's utility of undertaking a project  $(\pi, s)$  becomes

$$u_M(\pi, s) + \alpha_\pi \pi + \alpha_s s = (\beta_M + \alpha_\pi) \pi + (\gamma_M + \alpha_s) s.$$

If the owner offers the contract  $\alpha$  to the manager, the manager chooses the project

$$(\tilde{\pi}_M, \tilde{s}_M) = \left( \sqrt{\frac{(\gamma_M + \alpha_s)^2}{(\gamma_M + \alpha_s)^2 + (\beta_M + \alpha_\pi)^2}}, \sqrt{\frac{(\beta_M + \alpha_\pi)^2}{(\gamma_M + \alpha_s)^2 + (\beta_M + \alpha_\pi)^2}} \right).$$

Given the manager's preferred project  $(\tilde{\pi}_M, \tilde{s}_M)$  under the contract  $\alpha$ , the owner's utility minus the cost of the manager's contract becomes

$$u_O(\tilde{\pi}_M, \tilde{s}_M) - (\alpha_\pi \tilde{\pi}_M + \alpha_s \tilde{s}_M) = (\beta_O - \alpha_\pi) \tilde{\pi}_M + (\gamma_O - \alpha_s) \tilde{s}_M.$$

We first show that it can never be optimal to have both  $\alpha_\pi > 0$  and  $\alpha_s > 0$ . The reason is that both dimensions of the contract are costly but have the opposite effect on project choice. While a higher  $\alpha_\pi$  induces the manager to pick a project with a higher monetary and a lower social payoff, a higher  $\alpha_s$  achieves the opposite.

**Lemma 3.** *The owner never selects a contract with  $\alpha_\pi > 0$  and  $\alpha_s > 0$ .*

*Proof of Lemma 3.* Assume that there exists an optimal contract with  $\alpha_\pi > 0$  and  $\alpha_s > 0$ . If  $\tilde{\pi}_M > \pi_M$ , then the owner can induce the manager to implement the same project at a lower cost for the owner by setting  $\alpha_s = 0$  and using a lower  $\alpha_\pi$ . A symmetric argument applies when  $\tilde{\pi}_M < \pi_M$ . In this case, the owner can induce the manager to implement the same project by setting  $\alpha_\pi = 0$  and by lowering  $\alpha_s$ . Therefore, an optimal contract satisfies either  $\alpha_\pi = 0$  or  $\alpha_s = 0$ .  $\square$

We next show that the optimal contract never fully resolves the conflict of interest between the owner and the manager. That is, the owner does not choose a contract that induces the manager to implement the owner's preferred project  $(\pi_O, s_O)$ .

**Proposition 12.** *When  $R_O \neq R_M$ , then the contract the owner offers to the manager does not fully resolve the conflict of interest.*

*Proof of Proposition 12.* Consider the case in which  $R_O < R_M$ . The case  $R_O > R_M$  follows a symmetric argument. When  $R_O > R_M$ , the owner would never offer the manager a contract with  $\alpha_s > 0$  as this would only increase the conflict of interest between the two. Therefore, the owner offers the manager a contract  $(\alpha_\pi, 0)$ .

Without taking into account the cost of compensation, the optimal contract the owner offers to the manager solves

$$\frac{\partial u_O(\tilde{\pi}_M, \tilde{s}_M)}{\partial \alpha_\pi} = 0.$$

This benchmark contract would induce the manager to implement the owner's preferred project  $(\pi_O, s_O)$  and, therefore, fully resolve the conflict of interest.

Taking the cost of compensation into account, the optimal contract the owner offers to the manager solves

$$\frac{\partial u_O(\tilde{\pi}_M, \tilde{s}_M)}{\partial \alpha_\pi} - \tilde{\pi}_M - \alpha_\pi \frac{\partial \tilde{\pi}_M}{\partial \alpha_\pi} = 0.$$

Since  $\frac{\partial \tilde{\pi}_M}{\partial \alpha_\pi} > 0$ , the solutions of the two first-order conditions do not coincide.<sup>28</sup> Therefore, the owner selects a compensation contract  $(\alpha_\pi, 0)$  that yields a project choice by the manager  $(\tilde{\pi}_M, \tilde{s}_M)$  that is different from the owner's preferred project  $(\pi_O, s_O)$ . As a result, contracting does not fully resolve the conflict of interest.  $\square$

## C Effort and Effective Control

This appendix studies how monetary incentives and pro-social preferences affect stakeholders' incentives to exert effort and their allocation of effective control.

As the following result shows, changes in the stakeholders' incentive and preference parameters lead to a substitution between their effort levels in equilibrium.

**Proposition 13** (Monetary Incentives, Pro-social Preferences, and Effort). *A change in the owner's or manager's monetary incentives or pro-social preferences leads to a substitution of effort between the owner and the manager. That is, for  $\theta \in \{\beta_O, \gamma_O, \beta_M, \gamma_M\}$ ,*

$$\frac{\partial q_O(d)}{\partial \theta} \frac{\partial q_M(d)}{\partial \theta} \leq 0.$$

The substitution of effort between the owner and manager, as demonstrated in Proposition 13, implies a substitution of effective control between the two stakeholders. This substitution has the same direction as the substitution of effort.

**Corollary 2** (Monetary Incentives, Pro-social Preferences, and Effective Control). *A change in the owner's or manager's monetary incentives or pro-social preferences leads to a substitution of effective control between the owner and the manager in line with their changes in effort. That is, for  $\theta \in \{\beta_O, \gamma_O, \beta_M, \gamma_M\}$ ,*

$$\frac{\partial e_O(d)}{\partial \theta} \frac{\partial e_M(d)}{\partial \theta} \leq 0 \quad \text{and} \quad \frac{\partial e_O(d)}{\partial \theta} \frac{\partial q_O(d)}{\partial \theta} \geq 0 \quad \text{and} \quad \frac{\partial e_M(d)}{\partial \theta} \frac{\partial q_M(d)}{\partial \theta} \geq 0.$$

Proposition 13 and Corollary 2 highlight a key force that arises in our model. In addition to changing the preferred project, altering a stakeholder's monetary incentives or pro-social preferences also leads to a substitution of effort from one stakeholder to the other and thus to a substitution of effective control. In particular, making a stakeholder more pro-social not only shifts the stakeholder's preferred project towards the social payoff, but also changes the allocation of effective control—the extent to which the stakeholder can actually influence the project the organization eventually implements.

To understand the effect a more pro-social stakeholder has on the organization's payoffs, we need to analyze how altering pro-social preferences affects effective control. To this end, we first study the effect of changing the manager's pro-social preferences  $\gamma_M$ . We then discuss the comparative statics with respect to the owner's pro-social preferences  $\gamma_O$  as well as the stakeholders' monetary incentives  $\beta_O$  and  $\beta_M$ .<sup>29</sup>

<sup>28</sup>Note that  $\frac{\partial u_O(\tilde{\pi}_M, \tilde{s}_M)}{\partial \alpha_\pi} > 0$  at  $\alpha_\pi = 0$ , such that an optimal contract satisfies  $\alpha_\pi > 0$ .

<sup>29</sup>Note that the comparative statics with respect to the owner's and manager's effort costs can be directly

**Proposition 14** (Pro-social Preferences and Effort when Owner Holds Control Rights). *When the owner holds the control rights,  $d = O$ , then the owner exerts less effort and the manager exerts more effort when the manager’s pro-social preferences  $\gamma_M$  increase, that is,*

$$\frac{\partial q_O(O)}{\partial \gamma_M} \leq 0 \quad \text{and} \quad \frac{\partial q_M(O)}{\partial \gamma_M} \geq 0.$$

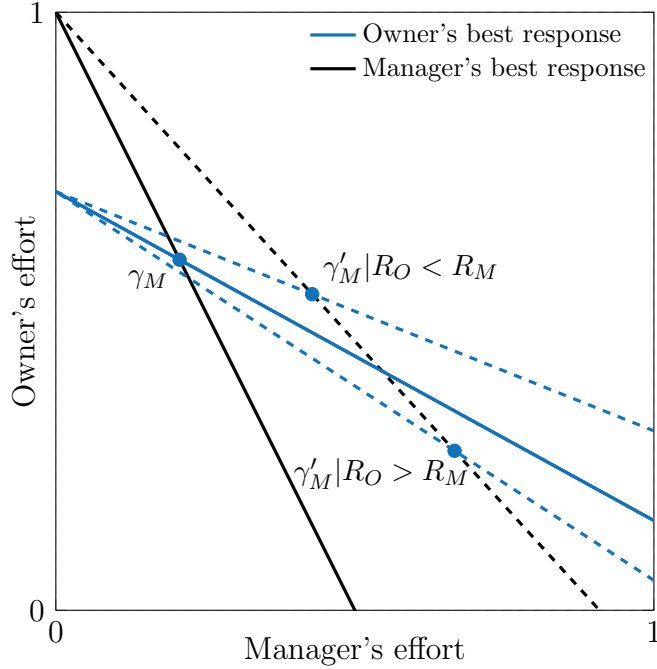


Figure A.1: **Manager’s pro-social preferences and equilibrium effort when the owner holds the control rights.** The figure plots the owner’s best response function  $B_O(q_M, O)$  and the manager’s best response function  $B_M(q_O, O)$ . The solid lines are the best response functions for some initial level of the manager’s pro-social preferences  $\gamma_M$  and the dashed lines for a marginally higher level of the manager’s pro-social preferences  $\gamma'_M > \gamma_M$ . The figure distinguishes between two cases of the owner’s best response function, one in which the initial  $\gamma_M$  satisfies  $R_O > R_M$  and one in which it satisfies  $R_O < R_M$ .

Strengthening the manager’s pro-social preferences unambiguously increases the manager’s incentives to exert effort. Figure A.1 illustrates this by showing that the manager’s best response function  $B_M(q_M, O)$  shifts outwards as  $\gamma_M$  increases. This happens because the manager’s utility when having effective control,  $u_M(\pi_M, s_M)$ , increases. Intuitively, because the manager cares more about the organization’s social payoff, the utility when having effective control increases, which in turn increases the incentives to exert effort.

Importantly, the owner’s best response function and therefore incentives to exert effort

---

determined from their impact on the best response functions. A higher effort cost  $\phi_O$  of the owner decreases the incentives to exert effort and therefore lowers the owner’s best response function, while the best response function of the manager remains unaffected. Therefore, the owner’s effort decreases while the manager’s effort increases. A similar argument can be made when increasing the manager’s effort cost  $\phi_M$ .

are also affected by a change in the manager's pro-social preferences, that is,

$$\frac{\partial B_O(q_M, O)}{\partial \gamma_M} = -q_M \frac{\frac{\partial u_O(\pi_M, s_M)}{\partial \gamma_M}}{\phi_O} < 0 \quad \Leftrightarrow \quad R_O > R_M.$$

This happens because the manager's project choice, and therefore the owner's utility, changes when the manager has effective control. If the owner has stronger relative pro-social preferences than the manager, that is, if  $R_O > R_M$ , then the owner's utility when the manager's preferred project  $(\pi_M, s_M)$  is implemented increases in response to a higher  $\gamma_M$  as it brings the manager's preferred project closer to the owner's. Thus, losing effective control to the manager becomes less costly, which in turn reduces the owner's incentives to exert effort. Put differently, the manager's effective control provides a better hedge for the owner in this case. In contrast, if an increase in the manager's pro-social preferences aggravates the conflict of interest between the stakeholders, that is, if  $R_O < R_M$ , the owner's incentives to exert effort increase.

The two cases are illustrated in Figure A.1. It turns out that in our model, the direct effect on the manager's utility always dominates the indirect effect on the owner's utility. As such, a higher  $\gamma_M$  increases the manager's effort and decreases the owner's effort in equilibrium. For example, even if the manager of a firm controls the firm's decision making, employees becoming more pro-social causes them to gain more effective control and therefore more influence on the firm's outcomes.

**Proposition 15** (Pro-social Preferences and Effort when Manager Holds Control Rights). *When the manager holds the control rights,  $d = M$ , then there exists a threshold  $\tilde{\gamma}_M$  such that when  $\gamma_M < \tilde{\gamma}_M$ , an increase in the manager's pro-social preferences increases the owner's effort and decreases the manager's effort, that is,*

$$\frac{\partial q_O(M)}{\partial \gamma_M} \geq 0 \quad \text{and} \quad \frac{\partial q_M(M)}{\partial \gamma_M} \leq 0,$$

and vice versa when  $\gamma_M > \tilde{\gamma}_M$ , that is,

$$\frac{\partial q_O(M)}{\partial \gamma_M} \leq 0 \quad \text{and} \quad \frac{\partial q_M(M)}{\partial \gamma_M} \geq 0.$$

Proposition 15 highlights that the effect of strengthening a stakeholder's pro-social preferences on the allocation of effective control critically depends on whether that stakeholder holds the control rights. As is clear from the owner's best response function,  $B_O(q_M, M)$ , the incentives to exert effort do not directly depend on the manager's preferred project  $(\pi_M, s_M)$  and therefore on  $\gamma_M$ . Intuitively, because the manager holds the control rights, the owner cannot directly reduce the probability that the manager has effective control and therefore the manager's preferred project does not directly affect the owner's effort incentives.

In contrast, the manager's best response function, and therefore incentives to exert effort,

depends on pro-social preferences  $\gamma_M$ . In particular, we have

$$\frac{\partial B_M(q_O, M)}{\partial \gamma_M} = (1 - q_O) \frac{\frac{\partial u_M(\pi_M, s_M)}{\partial \gamma_M}}{\phi_M} + q_O \frac{\frac{\partial \Delta u_M}{\partial \gamma_M}}{\phi_M} = \underbrace{(1 - q_O) \frac{s_M}{\phi_M}}_{\text{Direct Utility Effect}} + \underbrace{q_O \frac{(s_M - s_O)}{\phi_M}}_{\text{Hedging Effect}}.$$

The first term captures the direct effect on the manager's utility, which is always positive because increasing the manager's pro-social preferences results in a higher utility level. The second—hedging—effect arises as the manager can delegate project choice to the owner if failing to generate information. If  $\gamma_M$  is sufficiently low, such that  $R_M < R_O \Leftrightarrow s_M < s_O$ , the hedging effect is negative because the conflict of interest between the stakeholders becomes less severe as  $\gamma_M$  increases, which makes the hedge more valuable and therefore lowers the manager's incentives to exert effort. On the other hand, if  $\gamma_M$  is sufficiently high, such that  $R_M > R_O \Leftrightarrow s_M > s_O$ , the hedging effect is positive. Taken together, when  $\gamma_M$  is low, the hedging effect dominates and therefore the manager's effort decreases and the owner's increases as the manager becomes more pro-social. At  $\gamma_M = \tilde{\gamma}_M$ , the direct utility effect starts to dominate.<sup>30</sup> A further increase in  $\gamma_M$  thus leads to the manager exerting more effort and to the owner exerting less effort.

The comparative statics with respect to the owner's pro-social preferences  $\gamma_O$  as well as the stakeholders' monetary incentives  $\beta_O$  and  $\beta_M$  follow from the results obtained above due to the symmetry of our model. Specifically, we can relabel the social payoff as the monetary payoff and therefore the comparative statics with respect to  $\beta_M$  and  $\gamma_M$  are qualitatively identical. Furthermore, we can interchange the role of the owner and the manager conditional on the delegation decision. As a result, the comparative statics with respect to  $\gamma_M$  and  $\beta_M$  when  $d = M$  ( $d = O$ ) are qualitatively identical to those with respect to  $\gamma_O$  and  $\beta_O$  when  $d = O$  ( $d = M$ ).

## D Continuous Delegation of Control Rights

In this section, we show that our result in Proposition 5 that an increase in the manager's pro-social preferences can reduce the organization's sustainability due to a withdrawal of the control rights does not rely on the fact that the change in the control rights is discrete. As we show below in a simple extension of our model, what we need is that strengthening the manager's pro-social preferences leads to a significant reduction in the manager's control rights.

In the baseline model, the organization has a single task, namely it needs to decide which project to undertake, if any at all. Assume now that the organization has  $N > 1$  tasks indexed by  $i \in \{1, \dots, N\}$  instead of a single one. Each task  $i$  consists of a project choice similar to the one in the baseline model. The payoffs from task  $i$  are  $\frac{1}{N}$  times the payoffs of a project from the baseline model and zero if no project is undertaken, that is, the payoffs

<sup>30</sup>The condition  $(1 - q_O(M))s_M + q_O(M)(s_M - s_O) = 0$  implicitly characterizes  $\tilde{\gamma}_M$  when it is positive. Observe that at  $R_O = R_M$ , the hedging effect turns from negative to positive. This in combination with the fact that the direct utility effect always gives the manager stronger incentives to exert effort implies that the threshold  $\tilde{\gamma}_M$  must be below the level at which  $R_O = R_M$ . Thus, the threshold  $\tilde{\gamma}_M$  satisfies  $R_M < R_O$ .

for task  $i$  from the owner's and manager's preferred projects are given by  $\frac{1}{N}(\pi_O, s_O)$  and  $\frac{1}{N}(\pi_M, s_M)$ , respectively, and the effort cost to learn about the project payoffs for task  $i$  is  $\frac{1}{N}$  times the effort cost in the baseline model, that is,  $\frac{1}{N} \frac{\phi_j}{2} q_j^2$ , where  $j \in \{O, M\}$ . In addition, the owner receives an extra utility  $\frac{\epsilon_i}{N}$  from retaining the control rights for task  $i$ , where the random variables  $\epsilon_i$ ,  $i \in \{1, \dots, N\}$ , are independently drawn from a uniform distribution with support  $[-\sigma, \sigma]$  with  $\sigma \geq 0$ . This setup implies that for each task  $i$ , the owner delegates the control rights to the manager when  $\Delta U_O^i = \frac{1}{N}(\Delta U_O + \epsilon_i) < 0$  and retains the control rights when  $\Delta U_O^i > 0$ .

As a result, when  $N \rightarrow \infty$ , because of the law of large numbers, the fraction of tasks for which the owner delegates control rights to the manager is one when  $\Delta U_O < -\sigma$ ,  $\frac{\sigma - \Delta U_O}{2\sigma}$  when  $\Delta U_O \in [-\sigma, \sigma]$ , and zero when  $\Delta U_O > \sigma$ . If  $\sigma = 0$ , then we are back to our baseline model. For  $\sigma > 0$ , the fraction of the control rights delegated to the manager changes continuously as  $\gamma_M$  changes.

For  $N$  finite, the organization's sustainability is the sum over the tasks  $i \in \{1, \dots, N\}$  of the expected social payoff conditional on a project being undertaken for that task. Observe that when  $\sigma > 0$  and  $N \rightarrow \infty$ , the organization's sustainability is continuous in  $\gamma_M$ . Furthermore, when  $\sigma$  gets sufficiently small, there exist two thresholds  $\gamma'_M$  and  $\gamma''_M$  satisfying  $\gamma'_M < \hat{\gamma}_M < \gamma''_M$ , where  $\hat{\gamma}_M$  is defined in Proposition 5, such that:

*i)* For  $\gamma'_M$ , the owner delegates authority for all tasks while for  $\gamma''_M$  the owner retains authority for all tasks. Thus, for  $\gamma'_M$  and  $\gamma''_M$ , the organization's sustainability is the same as in the baseline model.

*ii)* For  $\gamma'_M$  and  $\gamma''_M$ , the organization's sustainability, which is the same as in the baseline model, satisfies

$$\mathbb{E}[\tilde{s} | \tilde{\pi} > 0, \gamma'_M] > \mathbb{E}[\tilde{s} | \tilde{\pi} > 0, \gamma''_M].$$

Continuity of the organization's sustainability in  $\gamma_M$  then implies that there exists a  $\gamma_M \in [\gamma'_M, \gamma''_M]$  such that  $\frac{\partial \mathbb{E}[\tilde{s} | \tilde{\pi} > 0]}{\partial \gamma_M} < 0$ .

## E Proofs

We organize the proofs into three sections. The first section contains the proofs for the baseline model related to the effort results taking the delegation of the control rights as given (Section II.B and Appendix C). The second section contains the proofs for the baseline model related to the delegation of the control rights (Section II.C). The third section contains the proofs for the model extensions (Section III).

### I Proofs for Section II.B and Appendix C

**Lemma 4** (Equilibrium Effort Choices). *Given the delegation decision  $d$ , there exists a unique Nash equilibrium in effort choices  $(q_O(d), q_M(d)) \in (0, 1)^2$  at time one, which is the solution to the first-order conditions of the owner's expected utility and the manager's expected utility with respect to their effort levels.*



*Proof of Lemma 4.* First, we want to show that the lower bounds for  $\phi_O$  and  $\phi_M$  ensure that  $q_O < 1$  and  $q_M < 1$ . If  $q_O = q_M = 1$  then the manager without the control rights would be better off setting the effort to zero. Therefore, either  $q_O < 1$  or  $q_M < 1$ . Assume, without loss of generality, that  $q_O < 1$ , then the lower bound for  $\phi_M$  implies that  $q_M < 1$ . Therefore,  $q_O < 1$  and  $q_M < 1$ . Second, from the first-order conditions it directly follows that  $q_M > 0$  and  $q_O > 0$ . As a consequence, the two first-order conditions define the optimal effort levels.

Finally, the first-order conditions define a system of two linear equations with two unknowns  $(q_O, q_M)$ . Direct calculations allow us to show that this system has a unique solution  $(q_O(d), q_M(d))$ .  $\square$

*Proof of Proposition 13.* Given  $d = O$ , we need to sign the product of four different pairs of derivatives. We start by signing two after which the other two follow from symmetry within the model.

1. Observe that

$$\begin{aligned}\frac{\partial q_O(O)}{\partial \gamma_M} &= \frac{\gamma_O \phi_M \left( \sqrt{\beta_O^2 + \gamma_O^2} - \phi_O \right)}{(\beta_M \beta_O + \gamma_M \gamma_O - \phi_M \phi_O)^2} \leq 0, \\ \frac{\partial q_M(O)}{\partial \gamma_M} &= \frac{\left( \sqrt{\beta_O^2 + \gamma_O^2} - \phi_O \right) (\beta_M (\beta_O \gamma_M - \beta_M \gamma_O) - \gamma_M \phi_M \phi_O)}{\sqrt{\beta_M^2 + \gamma_M^2} (\beta_M \beta_O + \gamma_M \gamma_O - \phi_M \phi_O)^2} \geq 0.\end{aligned}$$

The first inequality follows from

$$\begin{aligned}(\beta_M \beta_O + \gamma_M \gamma_O) &= u_O(\pi_O, s_O) u_M(\pi_O, s_O) \\ &\leq u_O(\pi_O, s_O) u_M(\pi_M, s_M) \\ &= \sqrt{\beta_O^2 + \gamma_O^2} \sqrt{\beta_M^2 + \gamma_M^2} \\ &< \phi_O \phi_M,\end{aligned}\tag{A.1}$$

and the fact that

$$\sqrt{\beta_O^2 + \gamma_O^2} = u_O(\pi_O, s_O) < \phi_O.\tag{A.2}$$

While the second inequality follows from equation (A.1), equation (A.2), and the fact that

$$\begin{aligned}\beta_M (\beta_O \gamma_M - \beta_M \gamma_O) - \gamma_M \phi_M \phi_O &\leq \beta_M (\beta_O \gamma_M - \beta_M \gamma_O) - \gamma_M (\beta_M \beta_O + \gamma_M \gamma_O) \\ &= -(\beta_M^2 + \gamma_M^2) \gamma_O \leq 0.\end{aligned}$$

2. Observe that

$$\begin{aligned}\frac{\partial q_O(O)}{\partial \gamma_O} \frac{\partial q_M(O)}{\partial \gamma_O} &= \\ &- \frac{\phi_M \sqrt{\beta_M^2 + \gamma_M^2} \left( -\beta_M \beta_O \gamma_O - \gamma_M \phi_O \sqrt{\beta_O^2 + \gamma_O^2} + \beta_O^2 \gamma_M + \gamma_O \phi_M \phi_O \right)^2}{(\beta_O^2 + \gamma_O^2) (\beta_M \beta_O + \gamma_M \gamma_O - \phi_M \phi_O)^4} \leq 0.\end{aligned}$$



The denominator is positive because of equation (A.1), which proves the inequality.

The comparative statics with respect stakeholders' monetary incentives  $\beta_M$  and  $\beta_O$  follow from the results obtained above due to the symmetry in our model. Specifically, we can relabel the social payoff as the monetary payoff and therefore the comparative statics with respect to  $\beta_M$  ( $\beta_O$ ) and  $\gamma_M$  ( $\gamma_O$ ) are qualitatively identical.

Furthermore, we can interchange the role of the owner and the manager conditional on the delegation decision to obtain the results when  $d = M$ .  $\square$

*Proof of Corollary 2.* The first result follows directly from that fact that  $e_O(d) = 1 - e_M(d)$  and therefore

$$\frac{\partial e_O(d)}{\partial \theta} \frac{\partial e_M(d)}{\partial \theta} \leq 0.$$

Observe that

$$e_O^{-1}(O) = \frac{q_O(O) + (1 - q_O(O))q_M(O)}{q_O(O)} = 1 + \frac{1 - q_O(O)}{q_O(O)}q_M(O).$$

If  $q_O(O)$  increases then  $q_M(O)$  decreases and therefore  $\frac{1 - q_O(O)}{q_O(O)}q_M(O)$  decreases and  $e_O(O)$  increases. Therefore,

$$\frac{\partial e_O(O)}{\partial \theta} \frac{\partial q_O(O)}{\partial \theta} \geq 0.$$

Similar arguments show that

$$\frac{\partial e_O(M)}{\partial \theta} \frac{\partial q_O(M)}{\partial \theta} \geq 0.$$

The final result follows from the symmetry in the model and can be obtained by interchanging the owner and the manager in the steps above.  $\square$

*Proof of Proposition 14.* The result follows directly from the derivations in the proof of Proposition 13.  $\square$

*Proof of Proposition 15.* Given that the manager holds the control rights,  $d = M$ , the owner's best response function  $B_O(q_M, M)$  does not change when varying  $\gamma_M$ . The manager's best response function is given by

$$B_M(q_O, M) = \frac{(1 - q_O)u_M(\pi_M, s_M) + q_O\Delta u_M}{\phi_M}.$$

From the envelope theorem it then follows that

$$\frac{\partial B_M(q_O, M)}{\partial \gamma_M} = \frac{s_M - q_O s_O}{\phi_M}.$$

Therefore, if  $s_M - q_O s_O > 0$ , then  $\frac{\partial q_M(M)}{\partial \gamma_M} > 0$  and, as implied by Proposition 13,  $\frac{\partial q_O(M)}{\partial \gamma_M} \leq 0$ . Similarly, if  $s_M - q_O s_O < 0$ , then  $\frac{\partial q_M(M)}{\partial \gamma_M} < 0$  and, as implied by Proposition 13,  $\frac{\partial q_O(M)}{\partial \gamma_M} \geq 0$ .

Finally, if  $s_M - q_O s_O = 0$ , then  $\frac{\partial B_M(q_O, M)}{\partial \gamma_M} = \frac{\partial B_O(q_M, M)}{\partial \gamma_M} = 0$ , and therefore  $\frac{\partial q_M(M)}{\partial \gamma_M} = \frac{\partial q_O(M)}{\partial \gamma_M} = 0$ .

Assume that  $s_M - q_O s_O \geq 0$ , then  $s_M$  is increasing in  $\gamma_M$  while  $q_O$  is weakly decreasing in  $\gamma_M$  and  $s_O$  remains unchanged. Therefore, if there exists a  $\gamma_M$  such that  $s_M - q_O s_O \geq 0$ , then  $s_M - q_O s_O > 0$  for any  $\gamma'_M > \gamma_M$ . Furthermore, we know that when  $R_M > R_O$ , then  $s_M - q_O s_O > s_M - s_O > 0$ . Taken together, this proves that there exists a unique  $\tilde{\gamma}_M$  such that  $s_M - q_O s_O = 0$ . For  $\gamma'_M > \tilde{\gamma}_M$ , we have  $s_M - q_O s_O > 0$ , and for  $\gamma'_M < \tilde{\gamma}_M$ , we have  $s_M - q_O s_O < 0$ , which proves the result.  $\square$

*Proof of Lemma 1.* When  $d = O$ , the result follows directly from Proposition 14 and Corollary 2.

When  $d = M$  and  $\gamma_O = 0$ , then  $\tilde{\gamma}_M = 0$  because

$$s_M - q_O s_O = s_M \geq 0.$$

The result then follows from Proposition 15 and Corollary 2.  $\square$

*Proof of Proposition 1.* The result follows from Lemma 1 and the fact that

$$\mathbb{E}[\tilde{s} | \tilde{\pi} > 0, d] = e_M(d) s_M.$$

$\square$

## II Proofs for Section II.C

*Proof of Lemma 2.* Direct calculations imply that

$$q_O(O) - q_O(M) = q_M(M) - q_M(O) = \frac{-\sqrt{(\beta_M^2 + \gamma_M^2)(\beta_O^2 + \gamma_O^2)} + \beta_M \beta_O + \gamma_M \gamma_O}{\beta_M \beta_O + \gamma_M \gamma_O - \phi_M \phi_O} \geq 0.$$

The inequality follows from equation (A.1), which shows that the numerator and denominator are nonpositive and negative, respectively.  $\square$

*Proof of Proposition 2.* Since  $R_O = R_M$ , it follows that  $(\pi_O, s_O) = (\pi_M, s_M)$  and therefore  $\Delta u_O = \Delta u_M = 0$ . The first-order conditions that determine the effort levels are therefore independent of  $d$  and as a result

$$(q_O(O), q_M(O)) = (q_O(M), q_M(M)).$$

These observations then imply that

$$U_O(q_O(O), q_M(O), O) = U_O(q_O(M), q_M(M), M),$$

which completes the proof.  $\square$

*Proof of Proposition 3.* We first prove the result for when the wedge is small after which we prove the result for when the wedge is large.

Consider a set of parameters  $(\beta'_O, \gamma'_O, \beta'_M, \gamma'_M)$ . Define  $(\beta'_O, \gamma'_O, \beta'_M, \gamma''_M)$ , where  $\gamma''_M = \beta'_M \gamma'_O / \beta'_O \geq 0$ . For this set of parameters, the relative pro-social preferences are the same and from Proposition 2 it then follows that  $\Delta U_O(\beta'_O, \gamma'_O, \beta'_M, \gamma''_M) = 0$ . As a consequence,

$$\Delta U_O(\beta'_O, \gamma'_O, \beta'_M, \gamma'_M) = \int_{\gamma''_M}^{\gamma'_M} \frac{\partial \Delta U_O(\beta'_O, \gamma'_O, \beta'_M, \gamma_M)}{\partial \gamma_M} d\gamma_M.$$

Observe that

$$\begin{aligned} \left. \frac{\partial \Delta U_O(\beta'_O, \gamma'_O, \beta'_M, \gamma_M)}{\partial \gamma_M} \right|_{\gamma_M = \gamma''_M} &= 0, \\ \left. \frac{\partial^2 \Delta U_O(\beta'_O, \gamma'_O, \beta'_M, \gamma_M)}{\partial \gamma_M^2} \right|_{\gamma_M = \gamma''_M} &= \frac{(\beta'_O)^4 \left( (\beta'_O)^2 + (\gamma'_O)^2 - \sqrt{(\beta'_O)^2 + (\gamma'_O)^2} \phi_O \right)}{\sqrt{(\beta'_O)^2 + (\gamma'_O)^2} \left( \beta'_M \left( (\beta'_O)^2 + (\gamma'_O)^2 \right) - \beta'_O \phi_M \phi_O \right)^2} < 0, \end{aligned}$$

because of equation (A.2) and the fact that when the relative pro-social preferences are aligned (i.e.,  $R_O = R_M$ ), then

$$u_O(\pi_O, s_O) u_M(\pi_O, s_O) = \frac{\beta_M}{\beta_O} (\beta_O^2 + \gamma_O^2) < \phi_M \phi_O.$$

Note that

$$|R'_O - R'_M| = \left| \frac{\left( \frac{\gamma'_O}{\beta'_O} \right)^2}{\left( \frac{\gamma'_O}{\beta'_O} \right)^2 + 1} - \frac{\left( \frac{\gamma'_M}{\beta'_M} \right)^2}{\left( \frac{\gamma'_M}{\beta'_M} \right)^2 + 1} \right| = \left| \frac{(r'_O)^2}{(r'_O)^2 + 1} - \frac{(r'_M)^2}{(r'_M)^2 + 1} \right|,$$

where  $r'_O := \frac{\gamma'_O}{\beta'_O}$  and  $r'_M := \frac{\gamma'_M}{\beta'_M}$ , and

$$|\gamma''_M - \gamma'_M| = \left| \frac{\beta'_M}{\beta'_O} \gamma'_O - \gamma'_M \right| = \beta'_M |r'_O - r'_M| \leq \phi_M |r'_O - r'_M|,$$

because  $\beta'_M \leq \sqrt{(\beta'_M)^2 + (\gamma'_M)^2} < \phi_M$ . As a result, for given parameters  $\gamma'_O$  and  $\beta'_O$ , as  $|R'_O - R'_M| \rightarrow 0$ , we have that  $|r'_O - r'_M| \rightarrow 0$  because the function  $f(r) = \frac{r^2}{r^2+1}$  is continuous and strictly increasing for  $r > 0$ . Thus, if  $|R'_O - R'_M|$  is sufficiently small, then  $\gamma'_M$  is sufficiently close to  $\gamma''_M$ .

Continuity of  $\Delta U_O$  and its first- and second-order derivative with respect to  $\gamma_M$  then implies that for  $\gamma'_M > \gamma''_M$  sufficiently close to  $\gamma''_M$ ,

$$\Delta U_O(\beta'_O, \gamma'_O, \beta'_M, \gamma'_M) = \int_{\gamma''_M}^{\gamma'_M} \frac{\partial \Delta U_O(\beta'_O, \gamma'_O, \beta'_M, \gamma_M)}{\partial \gamma_M} d\gamma_M < 0,$$

while when  $\gamma'_M < \gamma''_M$  sufficiently close to  $\gamma''_M$ ,

$$\Delta U_O(\beta'_O, \gamma'_O, \beta'_M, \gamma'_M) = - \int_{\gamma'_M}^{\gamma''_M} \frac{\partial \Delta U_O(\beta'_O, \gamma'_O, \beta'_M, \gamma_M)}{\partial \gamma_M} d\gamma_M < 0,$$

which proves the result when  $|R_O - R_M| > 0$  is sufficiently small.

For  $|R_O - R_M| \rightarrow 1$ , we have

$$\begin{aligned} \lim_{|R_O - R_M| \rightarrow 1} U_O(q_O(O), q_M(O), O) &= \tilde{q}_O(O) \tilde{u}_O(\pi_O, s_O) - \frac{\phi_O}{2} (\tilde{q}_O(O))^2 \\ &= \max_q \left\{ q \tilde{u}_O(\pi_O, s_O) - \frac{\phi_O}{2} q^2 \right\} \\ &> \max_q \left\{ (1 - \tilde{q}_M(M)) q \tilde{u}_O(\pi_O, s_O) - \frac{\phi_O}{2} q^2 \right\} \\ &= (1 - \tilde{q}_M(M)) \tilde{q}_O(M) \tilde{u}_O(\pi_O, s_O) - \frac{\phi_O}{2} (\tilde{q}_O(M))^2 \\ &= \lim_{|R_O - R_M| \rightarrow 1} U_O(q_O(M), q_M(M), M), \end{aligned}$$

where the variables and functions with a tilde are the limits of their respective variables and functions as  $|R_O - R_M| \rightarrow 1$ . The first and last equalities follow from the product rule for limits using the fact that  $\tilde{u}_O(\pi_M, s_M) = 0$ . The second and third equalities follow from the fact that the limits of the first-order conditions that define  $\tilde{q}_O(d)$  solve these optimization problems. While the inequality follows from the fact that  $\tilde{q}_M(M) > 0$ .

Continuity of the owner's expected utility in the model parameters, taking as given the delegation decision, then proves the result for the case when the wedge in relative pro-social preferences is large.  $\square$

*Proof of Proposition 4.* Observe that  $\Delta U_O$  has the same sign as

$$\hat{\Delta} U_O = \Delta U_O \left( 2\sqrt{\beta_M^2 + \gamma_M^2} (\beta_M \beta_O - \phi_M \phi_O)^2 \right).$$

Furthermore, when  $\gamma_O = 0$ , then

$$\frac{\partial \hat{\Delta} U_O}{\partial \gamma_M} = \frac{\beta_O^2 \gamma_M \left( 2\beta_M^2 (\beta_O - \phi_O) + 4\phi_M \phi_O \sqrt{\beta_M^2 + \gamma_M^2} - 4\beta_M \phi_M \phi_O - 3\gamma_M^2 \phi_O \right)}{\sqrt{\beta_M^2 + \gamma_M^2}},$$

which is zero for at most two  $\gamma_M > 0$ . To show this, replace  $\gamma_M^2$  by  $C^2 - \beta_M^2$  and notice that the term in the brackets in the numerator yields a quadratic equation in  $C$ .

At  $\gamma_M = 0$ , we get that

$$\hat{\Delta} U_O = 0 \quad \text{and} \quad \frac{\partial \hat{\Delta} U_O}{\partial \gamma_M} = 0 \quad \text{and} \quad \frac{\partial^2 \hat{\Delta} U_O}{\partial \gamma_M^2} = 2\beta_M \beta_O^2 (\beta_O - \phi_O) < 0.$$

Therefore, in a neighbourhood above  $\gamma_M = 0$  we have that  $\frac{\partial \hat{\Delta}U_O}{\partial \gamma_M} < 0$  and  $\hat{\Delta}U_O < 0$ .

Furthermore, for  $\gamma_M \rightarrow \sqrt{\phi_M^2 - \beta_M^2}$ , we have that

$$\lim_{\gamma_M \rightarrow \sqrt{\phi_M^2 - \beta_M^2}} \hat{\Delta}U_O = -\beta_O^2(\beta_M - \phi_M)(2\beta_M^2\beta_O - 3\beta_M\phi_M\phi_O + \phi_M^2\phi_O),$$

which is positive if  $8\beta_O > 9\phi_O$ , in which case  $\hat{\phi}_M = 0$ , or if

$$\phi_M > \hat{\phi}_M = \frac{1}{2}\beta_M \left( 3 + \sqrt{9 - \frac{8\beta_O}{\phi_O}} \right).$$

We thus have that: *i*) as  $\gamma_M \rightarrow 0$ ,  $\hat{\Delta}U_O$  is negative and decreasing in  $\gamma_M$ , *ii*)  $\frac{\partial \hat{\Delta}U_O}{\partial \gamma_M} = 0$  has at most two solutions for  $\gamma_M > 0$ , and *iii*) if  $\phi_M > \hat{\phi}_M$ , then  $\lim_{\gamma_M \rightarrow \sqrt{\phi_M^2 - \beta_M^2}} \hat{\Delta}U_O > 0$ . Therefore,  $\hat{\Delta}U_O$  crosses zero once and we define  $\hat{\gamma}_M \in \left(0, \sqrt{\phi_M^2 - \beta_M^2}\right)$  as this point of crossing.

For  $\gamma_M \in (0, \hat{\gamma}_M)$ , the owner delegates the control rights because  $\hat{\Delta}U_O < 0$  and therefore  $\Delta U_O < 0$ . For  $\gamma_M > \hat{\gamma}_M$ , the owner retains the control rights because  $\hat{\Delta}U_O > 0$  and therefore  $\Delta U_O > 0$ .  $\square$

*Proof of Proposition 5.* First, from Proposition 4 it follows that the manager follows a threshold delegation strategy where below  $\hat{\gamma}_M$ , the owner delegates the control rights while above it the owner retains the control rights. From Proposition 1 it then follows that, in each of these two regions, the organization's sustainability is weakly increasing. Furthermore, from the proof of Lemma 2 it follows that  $q_M(O) < q_M(M)$  at  $\hat{\gamma}_M > 0$  because  $R_O = 0 < R_M$ . Therefore, the organization's sustainability jumps downwards at  $\hat{\gamma}_M$ .  $\square$

*Proof of Proposition 6.* Recall that we assume that the delegation set is convex, that is,  $\{\gamma_M | d = M\} = [\underline{\gamma}_M, \bar{\gamma}_M]$ , where  $0 \leq \underline{\gamma}_M \leq \bar{\gamma}_M$ . Denote  $\gamma_M^* = \{\gamma_M | R_O = R_M\}$ . From Proposition 2 and the fact that if indifferent, the owner delegates the control rights to the manager, it follows that  $\gamma_M^* \in [\underline{\gamma}_M, \bar{\gamma}_M]$ . We first show that  $\gamma_M^* < \bar{\gamma}_M$ . Assume that  $\gamma_M^* = \bar{\gamma}_M$ . We know that there exists an  $\epsilon > 0$  such that any  $\gamma_M \in (\bar{\gamma}_M, \bar{\gamma}_M + \epsilon)$  is a feasible parameter value in that  $\gamma_M \geq 0$  and  $\phi_M > \max_{(\pi, s) \in \mathcal{P}} u_M(\pi, s)$ . Therefore,  $\gamma_M^* = \bar{\gamma}_M$  contradicts Proposition 3, which shows that in a neighborhood around  $\gamma_M^*$ , the owner delegates the control rights. As consequence, we must have  $\gamma_M^* < \bar{\gamma}_M$ .

In the following, we distinguish between two cases. First, assume that  $\gamma_M^* = \underline{\gamma}_M$ . From the previous argument we know that  $\underline{\gamma}_M = \gamma_M^* < \bar{\gamma}_M$ . If  $\gamma_M^* = \underline{\gamma}_M > 0$ , then any  $\gamma_M \in [0, \underline{\gamma}_M)$  is feasible. Therefore,  $\gamma_M^* = \underline{\gamma}_M > 0$  contradicts Proposition 3, which shows that in a neighborhood around  $\gamma_M^*$ , the owner delegates the control rights. As a result, when  $\gamma_M^* = \underline{\gamma}_M$ , then  $\underline{\gamma}_M = 0$ , and the only shift in the control rights takes place at  $\bar{\gamma}_M$ . At  $\bar{\gamma}_M$ ,  $R_M > R_O$  and the control rights are withdrawn from the manager. According to the proof of Lemma 2, this shift in the control rights leads to a reduction in the manager's effort,  $q_M(O) < q_M(M)$ , and increase in the owner's effort,  $q_O(O) > q_O(M)$ . These effort changes

result in a drop in the manager's effective control,  $e_M(O) < e_M(M)$ , and therefore in a drop in the organization's sustainability.

Second, assume that  $\gamma_M^* \in (\underline{\gamma}_M, \bar{\gamma}_M)$ . The same arguments as in the first case show that at  $\bar{\gamma}_M$ , the shift in the control rights leads to a decrease in the organization's sustainability. Similar arguments show that if  $\underline{\gamma}_M > 0$ , then at  $\underline{\gamma}_M$ , the shift in the control rights leads to a decrease in the organization's sustainability.  $\square$

*Proof of Proposition 7.* Observe  $\Delta U_O$  has the same sign as

$$\hat{\Delta}U_O = \Delta U_O \frac{2(\beta_M \beta_O - \phi_M \phi_O)^2}{\beta_M}.$$

Furthermore, when  $\gamma_M = 0$ , then

$$\frac{\partial \hat{\Delta}U_O}{\partial \gamma_O} = 2\gamma_O \left( \frac{\beta_M \beta_O^2}{\sqrt{\beta_O^2 + \gamma_O^2}} - \beta_M \phi_O - \frac{2\beta_O \phi_M \phi_O}{\sqrt{\beta_O^2 + \gamma_O^2}} + 2\phi_M \phi_O \right),$$

which is zero for at most one  $\gamma_O > 0$  because the function in brackets is monotonic in  $\gamma_O$ .

At  $\gamma_O = 0$ , we get that

$$\hat{\Delta}U_O = 0 \quad \text{and} \quad \frac{\partial \hat{\Delta}U_O}{\partial \gamma_O} = 0 \quad \text{and} \quad \frac{\partial^2 \hat{\Delta}U_O}{\partial \gamma_O^2} = 2\beta_M(\beta_O - \phi_O) < 0.$$

Therefore, in a neighbourhood above  $\gamma_O = 0$ , we have that  $\frac{\partial \hat{\Delta}U_O}{\partial \gamma_O} < 0$  and  $\hat{\Delta}U_O < 0$ .

Furthermore, as  $\gamma_O \rightarrow \sqrt{\phi_O^2 - \beta_O^2}$ , we have that

$$\lim_{\gamma_O \rightarrow \sqrt{\phi_O^2 - \beta_O^2}} \hat{\Delta}U_O = -(\beta_O - \phi_O)^2(-2\phi_M \phi_O + \beta_M(2\beta_O + \phi_O)),$$

which is positive if  $\phi_O > \hat{\phi}_O = -((2\beta_M \beta_O)/(\beta_M - 2\phi_M))$ .

We thus have that: *i*) as  $\gamma_O \rightarrow 0$ ,  $\hat{\Delta}U_O$  is negative and decreasing in  $\gamma_O$ , *ii*)  $\frac{\partial \hat{\Delta}U_O}{\partial \gamma_O} = 0$  has at most one solution for  $\gamma_O > 0$ , and *iii*) if  $\phi_O > \hat{\phi}_O$ , then  $\lim_{\gamma_O \rightarrow \sqrt{\phi_O^2 - \beta_O^2}} \hat{\Delta}U_O > 0$ . Therefore,  $\hat{\Delta}U_O$  crosses zero once and we define  $\hat{\gamma}_O \in (0, \sqrt{\phi_O^2 - \beta_O^2})$  as this point of crossing.

For  $\gamma_O \in (0, \hat{\gamma}_O)$ , the owner delegates the control rights because  $\hat{\Delta}U_O < 0$  and therefore  $\Delta U_O < 0$ . For  $\gamma_O > \hat{\gamma}_O$ , the owner retains the control rights because  $\hat{\Delta}U_O > 0$  and therefore  $\Delta U_O > 0$ .  $\square$

*Proof of Proposition 8.* First, from Proposition 7 it follows that the manager follows a threshold delegation strategy where below  $\hat{\gamma}_O$ , the owner delegates the control rights, while above it the owner retains the control rights. From the symmetry in the model and Proposition 14 it follows that  $e_O(M)$  is increasing in  $\gamma_O$ . Furthermore, from the symmetry in the model and Proposition 15 it follows that when  $d = O$  and  $\gamma_M = 0$ , then  $\tilde{\gamma}_O = 0$  and therefore  $e_O(O)$  is increasing in  $\gamma_O$ . As a result, in each of these two regions, the organization's sustainability

is increasing. Furthermore, from the proof of Lemma 2 it follows that  $q_O(M) < q_O(O)$  at  $\hat{\gamma}_O > 0$  because  $R_M = 0 < R_O$ . Therefore, the organization's sustainability jumps upwards at  $\hat{\gamma}_O$ .  $\square$

*Proof of Proposition 9.* The proof works analogous to the proof of Proposition 6.  $\square$

### III Proofs for Section III

*Proof of Corollary 1.* We need to find an  $\alpha \geq -\gamma_M$  such that

$$R_M^\alpha = \frac{(\gamma_M + \alpha)^2}{\beta_M^2 + (\gamma_M + \alpha)^2} = R_O, \quad (\text{A.3})$$

where  $R_O \in [0, 1)$ . Observe that  $R_M^\alpha$  is strictly increasing in  $\alpha$ , for  $\alpha = -\gamma_M$  we have that  $R_M^\alpha = 0$ , and  $\lim_{\alpha \rightarrow \infty} R_M^\alpha = 1$ . Therefore, there exists a unique  $\alpha$  such that equation (A.3) is satisfied, meaning that the manager's effective relative pro-social preferences are the same as the owner's relative pro-social preferences and therefore the stakeholders' preferred projects are the same.  $\square$

*Proof of Proposition 10.* From Corollary 1, we know that there exists an  $\tilde{\alpha} \in [-\gamma_M, \infty)$  such that  $R_M^\alpha = R_O$ . Given that changing  $\alpha$  in the extended model is equivalent to changing the manager's pro-social preferences  $\gamma_M$  in the baseline model, it follows from Proposition 3 that for  $\alpha$  slightly below  $\tilde{\alpha}$ , the owner delegates the control rights.

Furthermore, for  $\gamma_M = 0$  (i.e.,  $\alpha = -\gamma_M$  in the extended model), we have

$$\lim_{\beta_O \rightarrow 0} \Delta U_O = -\frac{\beta_M \gamma_O^2 (\beta_M - 2\phi_M)}{2\phi_M^2 \phi_O} > 0,$$

and therefore the owner retains the control rights when  $\alpha$  and  $\beta_O$  are sufficiently small.

Therefore, there exists an  $\hat{\alpha} < \tilde{\alpha}$  where for  $\hat{\alpha}$ ,  $R_M^\alpha < R_O$ , such that delegation switches from the owner to the manager as  $\alpha$  increases. At this threshold, as discussed in the proof of Lemma 2, the organization's sustainability drops because of the strictly increased effort by the manager.  $\square$

*Proof of Proposition 11.* From Proposition 4 it follows that for  $\gamma_M \in (0, \hat{\gamma}_M)$ , the owner delegates the control rights while for  $\gamma_M \in (\hat{\gamma}_M, \sqrt{\phi_M^2 - \beta_M^2})$  the owner retains the control rights.

When  $\gamma_O = 0$ , then

$$\frac{\partial U_O(q_O(O), q_M(O), O)}{\partial \gamma_M} = 0.$$

Furthermore, when  $\gamma_O = 0$ , then

$$\frac{\partial U_O(q_O(M), q_M(M), M)}{\partial \gamma_M} = \frac{\beta_O^2 \gamma_M \left( -\beta_M^3 \beta_O + \gamma_M^2 \phi_O \left( \sqrt{\beta_M^2 + \gamma_M^2} - \phi_M \right) + \beta_M^2 \phi_O \sqrt{\beta_M^2 + \gamma_M^2} \right)}{(\beta_M^2 + \gamma_M^2)^{3/2} (\beta_M \beta_O - \phi_M \phi_O)^2}.$$

For  $\gamma_M > 0$ , this function has the same sign as

$$g(\gamma_M) = -\beta_M^3 \beta_O + \gamma_M^2 \phi_O \left( \sqrt{\beta_M^2 + \gamma_M^2} - \phi_M \right) + \beta_M^2 \phi_O \sqrt{\beta_M^2 + \gamma_M^2}.$$

Substituting  $\gamma_M = \sqrt{C^2 - \beta_M^2}$ , we get

$$\tilde{g}(C) = -\beta_M^3 \beta_O + C^2(C - \phi_M)\phi_O + \beta_M^2 \phi_M \phi_O.$$

This function has at most three  $C$  for which it is zero (because it is a third-order polynomial). The same is then true for  $g(\gamma_M)$ .

Furthermore,  $g(0) > 0$  and  $g(\phi_M) > 0$ , and therefore  $g(\gamma_M)$  has at most two solutions such that  $g(\gamma_M) = 0$  for  $\gamma_M \in \left[0, \sqrt{\phi_M^2 - \beta_M^2}\right]$ . A third is not possible because it contradicts the fact that  $g(0) > 0$  and  $g(\phi_M) > 0$ . A single solution that crosses zero is not possible since  $g(0) > 0$  and  $g(\phi_M) > 0$ . No solution or a single solution that reaches zero but does not cross it is not possible because this would imply that  $\frac{\partial U_O(q_O(M), q_M(M), M)}{\partial \gamma_M} > 0$  for (almost all)  $\gamma_M \in \left(0, \sqrt{\phi_M^2 - \beta_M^2}\right)$ , which contradicts the fact that for  $\gamma_M > \hat{\gamma}_M$ , we have  $\Delta U_O > 0$  and the owner retains the control rights (see Proposition 4). Therefore,  $g(\gamma_M) = 0$  has two solutions in the interval  $\left[0, \sqrt{\phi_M^2 - \beta_M^2}\right]$ , which we denote by  $\tilde{\gamma}_M$  and  $\gamma'_M > \tilde{\gamma}_M$ . As a consequence,  $\frac{\partial U_O(q_O(M), q_M(M), M)}{\partial \gamma_M}$  is strictly positive for  $\gamma_M \in (0, \tilde{\gamma}_M)$  and  $\gamma_M \in \left(\gamma'_M, \sqrt{\phi_M^2 - \beta_M^2}\right)$ , and strictly negative for  $\gamma_M \in (\tilde{\gamma}_M, \gamma'_M)$ .

Observe, that the threshold above which the owner delegates the control rights,  $\hat{\gamma}_M$ , must satisfy  $\hat{\gamma}_M \in (\tilde{\gamma}_M, \gamma'_M)$ . It cannot be the case that  $\hat{\gamma}_M \leq \tilde{\gamma}_M$ . The reason is that for  $\gamma_M \in (0, \tilde{\gamma}_M]$ , we have  $\Delta U_O < 0$  because  $\frac{\partial U_O(q_O(O), q_M(O), O)}{\partial \gamma_M} = 0$  and  $\frac{\partial U_O(q_O(M), q_M(M), M)}{\partial \gamma_M} > 0$ . Furthermore, it cannot be the case that  $\hat{\gamma}_M \geq \gamma'_M$  because for  $\gamma_M = \sqrt{\phi_M^2 - \beta_M^2}$ , we have  $\Delta U_O > 0$  (see Proposition 4) and for  $\gamma_M \in \left(\gamma'_M, \sqrt{\phi_M^2 - \beta_M^2}\right]$ , we have  $\frac{\partial U_O(q_O(O), q_M(O), O)}{\partial \gamma_M} = 0$  while  $\frac{\partial U_O(q_O(M), q_M(M), M)}{\partial \gamma_M} > 0$ , which implies that  $\Delta U_O > 0$  for  $\gamma_M \in \left(\gamma'_M, \sqrt{\phi_M^2 - \beta_M^2}\right]$ . Therefore,  $\hat{\gamma}_M \in (\tilde{\gamma}_M, \gamma'_M)$ .

For  $\gamma_A \in (0, \tilde{\gamma}_M)$ , the owner delegates the control rights because  $\hat{\gamma}_M \in (\tilde{\gamma}_M, \gamma'_M)$  and therefore  $\frac{\partial \max_{d \in \{O, M\}} U_O(q_O(d), q_M(d), d)}{\partial \gamma_M} = \frac{\partial U_O(q_O(M), q_M(M), M)}{\partial \gamma_M} > 0$ . For  $\gamma_M \in [\tilde{\gamma}_M, \hat{\gamma}_M)$ , the owner still delegates the control rights but  $\frac{\partial \max_{d \in \{O, M\}} U_O(q_O(d), q_M(d), d)}{\partial \gamma_M} = \frac{\partial U_O(q_O(M), q_M(M), M)}{\partial \gamma_M} \leq 0$ . At  $\hat{\gamma}_M$ , the owner is indifferent between delegating or retaining the control rights and  $\frac{\partial U_O(q_O(M), q_M(M), M)}{\partial \gamma_M} \leq 0$  and  $\frac{\partial U_O(q_O(O), q_M(O), O)}{\partial \gamma_M} = 0$ . For  $\gamma_M > \hat{\gamma}_M$ , the owner retains the control rights and therefore  $\frac{\partial \max_{d \in \{O, M\}} U_O(q_O(d), q_M(d), d)}{\partial \gamma_M} = \frac{\partial U_O(q_O(O), q_M(O), O)}{\partial \gamma_M} = 0$ , which proves the first result of the proposition.

The result that the organization's sustainability improves for  $\gamma_A \in (0, \tilde{\gamma}_M)$  follows directly from Proposition 1 and the fact that  $\tilde{\gamma}_M \leq \hat{\gamma}_M$ .  $\square$



## F Project Implementation Effect

**Proposition 16** (Project Implementation Effect and Manager's Pro-Social Preferences). *Fix  $(\beta_O, \gamma_O, \beta_M)$  then there exists a threshold  $\check{\gamma}_M$  such that for  $\gamma_M > \check{\gamma}_M$  the project implementation effect is positive*

$$\mathbb{P}(\tilde{\pi} > 0 | d = M) > \mathbb{P}(\tilde{\pi} > 0 | d = O).$$

*Proof of Proposition 16.* Observe that the following inequalities are equivalent

$$\begin{aligned} & \mathbb{P}(\tilde{\pi} > 0 | d = M) > \mathbb{P}(\tilde{\pi} > 0 | d = O), \\ & 1 - \mathbb{P}(\tilde{\pi} > 0 | d = M) < 1 - \mathbb{P}(\tilde{\pi} > 0 | d = O), \\ & (1 - q_O(M))(1 - q_M(M)) < (1 - q_O(O))(1 - q_M(O)), \\ & \frac{(1 - q_O(M))(1 - q_M(M))}{(1 - q_O(O))(1 - q_M(O))} < 1. \end{aligned}$$

We can rewrite

$$\begin{aligned} & \frac{(1 - q_O(M))(1 - q_M(M))}{(1 - q_O(O))(1 - q_M(O))} \\ &= \frac{\phi_O \left( \phi_M - \sqrt{\beta_M^2 + \gamma_M^2} \right) \left( \beta_M \beta_O + \gamma_M \gamma_O - \sqrt{\beta_M^2 + \gamma_M^2} \sqrt{\beta_O^2 + \gamma_O^2} + \phi_M \sqrt{\beta_O^2 + \gamma_O^2} - \phi_M \phi_O \right)}{\phi_M \left( \phi_O - \sqrt{\beta_O^2 + \gamma_O^2} \right) \left( \beta_M \beta_O + \gamma_M \gamma_O - \sqrt{\beta_M^2 + \gamma_M^2} \sqrt{\beta_O^2 + \gamma_O^2} + \phi_O \sqrt{\beta_M^2 + \gamma_M^2} - \phi_M \phi_O \right)}. \end{aligned}$$

As  $\gamma_M$  increases such that  $\sqrt{\beta_M^2 + \gamma_M^2} = u_M(\pi_M, s_M) \rightarrow \phi_M$  then the numerator goes to zero while the denominator converges to a non-zero number. As a result

$$\lim_{\gamma_M \rightarrow \sqrt{\phi_M^2 - \beta_M^2}} \frac{(1 - q_O(M))(1 - q_M(M))}{(1 - q_O(O))(1 - q_M(O))} = 0.$$

Given that  $\frac{(1 - q_O(M))(1 - q_M(M))}{(1 - q_O(O))(1 - q_M(O))} > 0$  is sufficiently well-behaved and continuous in  $\gamma_M$  there exists a threshold  $\check{\gamma}_M$  such that for  $\gamma_M > \check{\gamma}_M$ ,  $\frac{(1 - q_O(M))(1 - q_M(M))}{(1 - q_O(O))(1 - q_M(O))} < 1$  and the project implementation effect is positive.  $\square$

**Proposition 17** (Project Implementation Effect and Owner's Pro-Social Preferences). *Fix  $(\beta_O, \beta_M, \gamma_M)$  then there exists a threshold  $\check{\gamma}_O$  such that for  $\gamma_O > \check{\gamma}_O$  the project implementation effect is negative*

$$\mathbb{P}(\tilde{\pi} < 0 | d = M) < \mathbb{P}(\tilde{\pi} > 0 | d = O).$$

*Proof of Proposition 17.* The proof is the same as in Proposition 16 when one interchanges the owner and manager.  $\square$