Executive Compensation with Environmental and Social Performance*

Pierre Chaigneau† Nicolas Sahuguet‡

Queen's University HEC Montréal

December 30, 2024

Abstract

How to incentivize a manager to create value and be socially responsible? A manager can predict how his decisions will affect measures of social performance, which are (randomly) biased. He will therefore game an incentive system that relies on these measures. Still, we show that the compensation contract is based on measures of social performance when the level of social investments preferred by the board exceeds the one that maximizes the stock price. When these measures are used, social investments are distorted because of gaming, and the sensitivity of pay to social performance is reduced to mitigate this effect. When there are several social performance measures with i.i.d. biases, relying on multiple measures based on different methodologies will generally mitigate inefficiencies due to gaming. This implies that harmonization of social performance measurement can backfire.

*For useful comments and suggestions we thank the Editor (Marcus Opp), an Associate Editor, a referee, as well as Jeremy Bertomeu, Jean de Bettignies, Milo Bianchi, Jörg Budde (discussant), Alexander David, Alexander Dyck, Alex Edmans, Thomas Geelen (discussant), Christian Hofmann, Nicolas Inostroza, Jan Mahrt-Smith, Lucas Mahieux, Marcus Opp (discussant), Mariana Sailer (discussant), Günter Strobl, Katrin Tinn, John Van Reenen, and participants at the Accounting and Economics Society webinar, Accounting Research Workshop, Asia Meeting of the Econometric Society, CSFN Conference on Governance and Sustainability, Erasmus Corporate Governance Conference, Québec Political Economy Conference, SFS Cavalcade, University of Alberta, University of Calgary, and University of Toronto.

†Corresponding author. Address: Smith School of Business, Goodes Hall, 143 Union Street West, K7L 2P3, Kingston, ON, Canada. Email: pierre.chaigneau@queensu.ca. Tel: 613 533 2312.

‡Address: Applied Economics Department, HEC Montréal, 3000 chemin de la côte Sainte Catherine, H3T 2A7 Montréal, QC, Canada. Email: nicolas.sahuguet@hec.ca. Tel: 514 340 6031.

1 Introduction

There has recently been a marked increase in the propensity of firms to use social and environmental measures of performance in executive compensation, with three-quarters of S&P 500 companies now integrating ESG (Environmental, Social and Governance) performance metrics into incentive plans.[1] Some executive compensation contracts rely on ESG scores and ratings (Cohen et al. (2023)) which are "third-party assessment[s] of corporations' ESG performance" (Berg, Kölbel, and Rigobon (2022)).

An overwhelming majority of investors supports the inclusion of such performance metrics in incentive plans. Some institutional investors even focus their engagement on the inclusion of these metrics.[2] Tying executive compensation to ESG metrics is advocated on the basis that compensation is a "powerful tool" to achieve various company goals, not just financial objectives.[3] ESG-based compensation is arguably an integral part of stakeholder capitalism: "Paying people principally in stock is antithetical to the whole notion that the shareholder ... is one of many stakeholders."[4] Linking greenhouse gases emissions reductions to executive compensation has even been proposed as a potential solution to the climate crisis.[5]

Yet, prominent scholars have argued that metrics of social and environmental performance should not be used for incentive purposes (Edmans (2021, 2023), Bebchuk and Tallarita (2022)).[6] They make two main points that cannot be easily dismissed, even in well-governed firms, and which are not captured by standard models of multitasking. First, they argue that executives will tend to "hit the target but miss the point", i.e. they will improve a measure of social or environmental performance even when they are aware that it does not improve actual social or

[1]Source: ESG Performance Metrics in Executive Pay, Harvard Law School Forum on Corporate Governance, January 15 2024.

[2]The 2021 Global Benchmark Policy Survey from ISS Governance finds that 86% of investors "believe [that] incorporating non-financial Environmental, Social, and/or Governance-related metrics into executive compensation programs is an appropriate way to incentivize executives." For example, see the ESG Engagement Campaign from Alliance Bernstein (April 2021), and BlackRock's report "Our 2021 Stewardship Expectations: Global Principles and Market-level Voting Guidelines".

[3]See the Keynote Address by SEC Commissioner Allison Herren Lee on Climate, ESG, and the Board of Directors at the 2021 Society for Corporate Governance National Conference, reproduced on the Harvard Law School Forum on Corporate Governance, June 30 2021.

[4]Source: How to pay executives in the age of stakeholder capitalism, *Financial Times*, December 14 2022. The quote is from Judy Samuelson, who founded the business and society programme at the Aspen Institute.

[5]Source: How to make climate progress: tie it to CEO pay, *Wall Street Journal*, April 17 2024.

[6]Even before the advent of ESG-based compensation, Shleifer and Vishny (1997) warned about the opportunities for self-dealing associated with incentive pay, and Tirole (2006) conjectured that the stakeholder society would be "best promoted through ... a fixed wage rather than performance-based incentives."

environmental outcomes. We will refer to this behavior as "gaming". Second, they argue that there is disagreement about the measurement of social and environmental performance, as documented in Berg, Kölbel, and Rigobon (2022). Not only is measurement imperfect, but there is no consensus on the measurement methodology.

Are explicit incentives based on measures of Social and Environmental Performance (SEP) truly a puzzle or can they be justified even if they lead managers to game the incentive system and there is disagreement about performance measurement? To answer this question, this paper analyzes a principal-agent model of multitasking that takes these factors into account.

A manager can exert effort to improve the firm's profits and also invest resources to improve its SEP at a cost. A socially responsible board offers a compensation contract to encourage the manager to improve both profitability and SEP. SEP is not directly observed, but it is imperfectly measured. We assume that, given his practical experience running the firm and his understanding of SEP measurement, the manager can anticipate how his decisions will affect its SEP measure. Thus, a manager with SEP-based incentives will tend to invest more (less) when the SEP measure is easy (hard) to improve, even if this leads to minimal (major) social and environmental impact. We refer to the discrepancy between the measured impact and the actual impact of social and environmental investments as the performance measure's "bias". For example, reaching the same level of carbon intensity can imply a substantial environmental impact for some firms but not for others. Since the manager knows the bias at the time of making investment decisions, SEP-based incentives will lead him to game the incentive system.

The manager's effort and investment decisions will affect the firm's stock price. Since some investors are socially responsible, we allow for a positive sensitivity of the stock price to the firm's estimated social output, with a slope which can be microfounded as investors' preference for social output. Signals of SEP are therefore already incorporated in the stock price. If the board prefers the level of social investments that maximizes the stock price, then we show that the manager's compensation only depends on the stock price.[7] Thus, explicit incentives based on SEP are only used when the board does not maximize the stock price. There are two cases.

First, if the level of social investment that the board prefers is lower than the one that maximizes the stock price, then the manager's compensation depends on the stock price and profits.

[7] In our model, the value of additional performance metrics is not due to risk sharing or rent extraction since the manager is risk neutral and can be kept at his reservation level of utility.

Intuitively, the board prefers a lower level of social investment than the one that maximizes the stock price, and profits-based compensation discourages such costly investment.

Second, if the level of social investment that the board prefers exceeds the one that maximizes the stock price, then the manager's compensation depends on the stock price and SEP measures. These measures are used to supplement incentives for social investments already embedded in stock price-based compensation. To mitigate inefficiencies due to gaming, the sensitivity of compensation to SEP measures is decreasing in the variance of their bias.[8] For these two reasons, the sensitivity of managerial compensation to SEP measures can be quite low.

In practice, there is not just one measure of SEP, even if we consider a single dimension of SEP (e.g. worker safety). Part of the reason is that there is no consensus on SEP measurement. Accordingly, we extend the model to analyze the outcome with multiple SEP measures, for example multiple ESG scores provided by multiple ESG raters. When the biases in ESG scores are independent and identically distributed, we show that increasing the number of scores reduces investment distortions due to gaming. Intuitively, if scores are constructed differently, it is harder for a manager to game multiple scoring methodologies than to game a single methodology.

However, simply adding measures of SEP will not necessarily reduce investment distortions. Indeed, consider two SEP measures with a different quality, as measured by the variance of their "bias", and whose biases can be correlated. Even though the availability of an additional SEP measure can reduce investment distortions due to gaming, it can alternatively worsen them if the additional measure has a low quality and a high correlation with the first measure. This is in contrast to contracting models in which the value of a performance measure is always nonnegative (e.g. Holmström (1979)). Intuitively, the stock price incorporates an additional SEP measure to the extent that it is informative, regardless of the potentially detrimental incentive effects. Since it affects the stock price, which is useful for incentive provision, this measure cannot simply be "ignored" by the board.

These results have normative implications for the heterogeneity of ESG scores and ratings, which is often criticized on the basis that it reflects disagreement between ESG raters (Chatterji et al. (2016), Berg, Kölbel, and Rigobon (2022), and Christensen, Serafeim, and Sikochi (2022)).

[8]The low quality of these measures is not a problem per se. In section 4, we study a setting in which the manager does not know the measures' biases at the time of making investment decisions. We find that the first-best outcome can then always be obtained. Moreover, in this case, the sensitivity of compensation to SEP measures does not depend on their quality, and it is higher than in the baseline model with "gaming".

In our model, a low correlation between ESG scores that reflects different scoring methodologies rather than low-quality scores can help reduce gaming and investment distortions, and increase the social impact of the firm.[9] Multiple ESG score providers can be useful, all the more that they rely on different methodologies and data, because they limit the possibility of gaming. This has implications for the ongoing debate on the regulation and harmonization of ESG ratings, and more generally for the measurement of corporate social and environmental performance.[10]

An empirical implication of our model is that an increase in socially responsible investing makes it less likely that firms will use SEP-based compensation. Intuitively, stock price-based compensation already provides strong incentives to improve the firm's social performance. Conversely, suppose that investors become less socially responsible because of changing investor sentiment. Then, supposing also that the degree of social responsibility preferred by boards does not change, the model predicts a rise in the use of SEP-based incentives. This can contribute to explain recent trends which might otherwise appear paradoxical. Indeed, it has been noted that ESG-based incentives are on the rise even though investors seem to be less concerned about ESG as suggested by declining inflows into ESG funds.[11]

This paper contributes to our understanding of SEP-based compensation. On the positive side, it potentially explains several stylized facts. First, it can explain that only a fraction of firms that claim to be socially responsible use explicit incentives based on SEP measures (Walker (2022)). Second, the model predicts that the use of SEP-based incentives is concentrated among firms whose boards are more socially responsible. This is consistent with the finding that SEP-based compensation is more common in firms with stated environmental pledges and in countries with "greater social sensitivity toward sustainability" (Cohen et al. (2023)), since directors and the shareholders they represent tend to be drawn from the local population. Third, our model can rationalize the low sensitivity of compensation to SEP measures (Rajan, Ramella, and Zingales (2023)): since SEP-based compensation is only used to complement social incentives already embedded in stock-based compensation, it can be effective even though it is relatively small. This contrasts with

[9] Specifically, ESG scores can have a low correlation either because their "biases" have a low correlation (the interpretation is that scores are built using different methodologies or data), or because these scores are mostly driven by idiosyncratic noise and therefore largely uninformative.

[10] See: Regulatory Solutions: A Global Crackdown on ESG Greenwash, Harvard Law School Forum on Corporate Governance, June 23 2022; EU watchdog says ESG rating firms need rules to stop 'greenwashing', *Reuters* February 12 2020, where Steven Maijoor, chair of the European Securities and Markets Authority (ESMA) is quoted as saying that "ESG rating agencies should be regulated and supervised appropriately by public sector authorities."

[11] Source: 76% of companies link pay to ESG performance in rising trend: WTW, CFO Dive, Jan 24 2024.

the perspective that a low sensitivity of compensation to SEP is either window dressing or not the outcome of optimal contracting (Walker (2022)). Fourth, our results are consistent with the finding that a stronger sensitivity of pay to SEP is associated with better social and environmental outcomes (Flammer, Hong, and Minor (2019), Cohen et al. (2023)), as measured by decreases in emissions and ESG ratings: in our model, a stronger preference for social output at the board level will lead to a higher weight on SEP measures in the compensation contract, which will lead to more social investments and better social performance.

On the normative side, we emphasize that wanting to encourage SEP is not a sufficient condition for the inclusion of SEP measures in incentive plans. Instead, boards only include these measures in incentive plans when the target level of SEP exceeds the one that maximizes the stock price. This perspective suggests that boards should better understand the determinants of their firm's stock market value before considering SEP incentives. If they use these incentives, they should consider relying on multiple measures of each relevant dimension of SEP to mitigate gaming. Moreover, our result that SEP-based incentives are only used when the board sets contracts that do not maximize the stock price differs from the common view that SEP-based incentives increase firm value, i.e. that firms are "doing well by doing good" (Falck and Heblich (2007), Flammer, Hong, and Minor (2019), Mischke, Woetzel, and Birshan (2021)).

This seemingly goes against the traditional view of the board's fiduciary duty. However, the board's fiduciary duty to act in the best interests of shareholders does not always involve maximizing profits[12] or even firm value, especially if shareholders have nonpecuniary preferences (Hart and Zingales (2017)). Finally, this perspective is consistent with a broader view of the board's responsibility, such as the "enlarged fiduciary duty" proposed by Tirole (2001). It is also consistent with "constituency statutes" adopted by several U.S. states that allow directors to consider stakeholders' interests when making business decisions (Flammer and Kacperczyk (2015)).

An important difference with respect to other contracting models in which the agent uses his private information to game incentives is that, instead of being exogenous, the distortionary effect of performance measures is endogenous in our model. Specifically, it depends on the divergence of preferences between the board and stock markets: distortions are worse when the board (optimally) includes explicit SEP-based incentives in the contract.

[12] Source: Fiduciary Duties of the Board of Directors, available at https://law.stanford.edu/wp-content/uploads/2023/01/Fiduciary-Duties-of-the-Board-of-Directors.pdf

2 Literature and social performance measures

2.1 ESG ratings and scores

ESG scores and ratings, which are "third-party assessment[s] of corporations' ESG performance" (Berg, Kölbel, and Rigobon (2022)), were originally developed to allow investors to screen companies for ESG (Environmental, Social and Governance) performance.

ESG raters typically provide several different types of measures of social and environmental performance. They provide an aggregate rating for a firm, as well as separate scores that reflect its performance on various dimensions of SEP: "category scores represent a rating agency's assessment of a certain ESG category. They are based on different sets of indicators that each rely on different measurement protocols." (Berg, Kölbel, and Rigobon (2022)) These categories include greenhouse gases emissions, workplace safety, board composition, etc.

In order to be measures of SEP activities that are comparable across firms and therefore useful to investors, ESG ratings and scores are highly standardized with publicly known formulas. For example, when describing their ESG scores, S&P Global mentions: "We publish our S&P Global ESG Score methodology on our website." Likewise, Bloomberg's ESG Scores are "fully transparent including methodology & company-reported data underlying each score."[13] Other measures of SEP, such as carbon intensity or board diversity, also share this feature.

This standardization leaves them open to gaming. Indeed, it is widely acknowledged that a firm can improve its ESG ratings by engaging in actions that improve perceptions of its SEP rather than its actual SEP (Walker (2022), Duchin, Gao, and Xu (2023)). In practice, some executive compensation contracts include ESG ratings and scores as performance metrics (Cohen et al. (2023), Table 3).

2.2 Related literature

The literature on compensation based on measures of social or environmental performance, including ESG-based compensation, is still in its infancy (we review current practices in section 7.3). The empirical findings of Homroy, Mavruk, and Nguyen (2023) suggest that boards use ESG-based compensation not because it increases shareholder wealth, but because shareholders

[13]Sources: Transparency and Impact: The Essential Principles of ESG, by Douglas L. Peterson, President & Chief Executive Officer of S&P Global, and Bloomberg Professional Services, www.bloomberg.com/explore/esg/.

demand ESG in addition to wealth. This is consistent with our modelization of the board's preferences. More generally, the empirical evidence on this type of compensation suggests that it solves an agency problem (Flammer, Hong, and Minor (2019), Ikram, Li, and Minor (2019), Homroy, Mavruk, and Nguyen (2023), and Pawliczek, Carter, and Zhong (2023)). On the contrary, Bebchuk and Tallarita (2022) have criticized companies that choose a narrow set of ESG metrics which correspond to easily achievable objectives but do not capture what is most consequential to their stakeholders. They also criticize the use of vague metrics that help corporate leaders retain discretion over the final amount of compensation. Even if these metrics can be poorly used, the notions that ESG-based compensation is a manifestation of poor governance, that it facilitates rent extraction, or that ESG targets are easily achieved, are inconsistent with empirical findings (Ikram, Li, and Minor (2019), Cohen et al. (2023), Homroy, Mavruk, and Nguyen (2023)).

What is driving ESG contracting? There is some evidence that the adoption of "say-on-pay" voting laws and greater shareholder engagement lead to an increase in ESG contracting (Cohen et al. (2023), Pawliczek, Carter, and Zhong (2022)). In particular, ESG mutual funds are more likely to support shareholder proposals about redesigning executive compensation to include environmental and social objectives (Dikolli et al. (2023)). The board's preferences also play an important role: Dyck et al. (2022) find that board renewal results in changes to the firm's ESG policies.[14] Overall, empirical findings are consistent with the view that ESG compensation tends to solve an agency problem, and is driven by boards and shareholders, which is the premise of our analysis.

It is well-known that paying agents for measured performance may not lead them to act in the principal's best interests (Kerr (1975)). Starting with Holmström and Milgrom (1991), the multi-tasking literature formally analyzes incentive provision when the agent has several tasks.[15] A strand of the literature focuses on the (un)bundling of tasks (Dewatripont, Jewitt, and Tirole (2000)). Another strand emphasizes the importance of maintaining balanced incentives across tasks. In

[14]Gillan, Koch, and Starks (2021) review an earlier empirical literature that documents correlations between corporate social performance (usually measured with ESG scores) and levels of executive compensation.

[15]This framework has been widely applied, including to balanced scorecards (Budde (2007), Dikolli, Hofmann, and Kulp (2009), Kvaløy and Olsen (2022)), short-term and long-term performance (Holmström and Tirole (1993), Gryglewicz, Mayer, and Morellec (2020)), symmetric tasks (Bond and Gomes (2009)), shifting effort between tasks (Smith (2002)), contracting with limited liability (Laux (2001)), dynamic considerations (Szydlowski (2019), Hoffmann and Pfeil (2021)), pay-for-performance with intrinsic motivation (Benabou and Tirole (2016)), information aggregation into the stock price (Paul (1992)), and performance measure congruity (Feltham and Xie (1994), Datar, Kulp, and Lambert (2001), Baker (2002), Bonham and Riggs-Cragun (2021)).

the standard multitasking model of Holmström and Milgrom (1991), inefficiencies emanate from imperfect or inexistent ex-post performance measurement for some tasks, which prevents adequate incentive provision on other tasks if efforts on various tasks are substitutes. In Feltham and Xie (1994), inefficiencies emanate from noisy ex-post performance measurement on several tasks, which can induce deviations from the first-best (multidimensional) action to reduce the risk borne by a risk averse agent. In our model, by contrast, the costs of various actions are independent and the manager is risk neutral.

A fundamental difference between our model and most multitasking models is that we let the manager be privately informed about the effect of his social investment decision on the firm's social performance. Indeed, we assume that the manager observes the "noise" in SEP before making his investment decision: he can predict the future realization of social performance. This introduces new possibilities for the manager to game a performance-based contract.

More closely related are the multitasking models of Prendergast (2002) and Ederer, Holden, and Meyer (2018), in which the agent has private information about either the payoff or the effort cost of various tasks.[16] In both models, tasks are substitutes: either the agent can choose only one task, or her marginal cost of effort on a task depends on overall effort. As a result, the agent shifts effort toward less costly tasks. In these models, an action on some dimension does not affect performance on other dimensions. In our model, by contrast, the social investment reduces profitability,[17] the stock price aggregates performance on both dimensions (social and financial), and the performance measures used in the contract depend on the divergence of preferences between the board and stock markets.[18] As a result, the distortionary effect of performance measures is endogenous in our model.

The assumption that the agent has more information than the principal at the effort choice stage is also made in the standard moral hazard models of Baker (1992) and Edmans and Gabaix (2011). In Baker (1992), this results in a lower piece-rate because effort is costly. In our model, lower incentives on the social dimension are not driven by costly effort (social investment is not

[16]Prendergast (2002) studies whether to use output-based contracts ("delegation") or input-based contracts. Ederer, Holden, and Meyer (2018) study the optimality of "opaque" incentive schemes which randomize the rewards.

[17]This is precisely why monetary incentives, which are required to motivate value-enhancing effort, must be offset by SEP-based incentives, which distort social investments because of gaming.

[18]Whenever these preferences diverge, the board does not want the manager to maximize the stock price. By contrast, in Baker (2002), stock price-based incentives are undistorted, and incentives are only distorted when a market measure of firm value is unavailable.

costly to the agent), but by the need to provide incentives on the profit dimension, which requires countervailing incentives on the social dimension, which results in gaming. Our multiplicative specification of the production technology for social output in equation (1) differs from Baker's (1992), whose specification is more general but less tractable, and from Edmans and Gabaix (2011) whose specification is additive. As a result, in their model, the optimal effort does not depend on the noise in the performance measure. By contrast, in our model, the optimal investment closely depends on the noise in the performance measure. In sum, the multiplicative specification that we postulate allows to model the gaming of incentives in a tractable way. It could be used in other applications.

The possibility of gaming is related to the vast literature on manipulation, but it involves different mechanisms. In these models, a signal such as a "report" affects a contractible performance measure or a decision. Manipulation alters the value of this signal to inflate the pay or the private benefits of a manager (e.g. Beyer, Guttman, and Marinovic (2014), Bertomeu, Darrough, and Xue (2017), Caskey and Laux (2019)). The manager introduces a bias in the signal. The optimal contract is designed to prevent or curb manipulation, and it account for signal jamming. By contrast, "gaming" refers to the manager using his private information to increase investment when the impact on measurable performance is the greatest, and conversely to decrease investment when this impact is minimal. The manager optimally responds to the bias in the signal when making his investment decision. The optimal contract is designed to mitigate this distortion in resource allocation.

Contracting based on ratings is studied in other papers. Parlour and Rajan (2020) show that including credit ratings in a contract may be optimal even when they are uninformative because they are contractible. Hörner and Lambert (2021) study how to optimally design a rating for incentive purposes, whereas we take ESG scores and ratings as given.

Our paper takes an optimal contracting perspective to the provision of incentives for corporate investment in social goods. Baron (2008) analyzes the balance between "profit incentives" and "social incentives". In his model, there is no uncertainty about the firm's technology for generating social output, and no stock price-based incentives. Bonham and Riggs-Cragun (2022) take a broader perspective, and study the use of contracts, taxes, and disclosure regulation to encourage ESG activities which are imperfectly measured. Likewise, there is no stock price in their model.

Bucourt and Inostroza (2023) study stock price-based incentives for ESG when investors have heterogeneous preferences for ESG. In their model, there is no agency problem and no contracting: the manager chooses her ESG effort to maximize shareholder value.

3 The model

3.1 Technology

Consider a firm run by a manager and controlled by a socially responsible board on behalf of shareholders. The production technology involves both a value-increasing action ("effort") and a social investment decision. At $t = 0$, a risk neutral manager chooses unobservable effort $e \in \{\underline{e}, \overline{e}\}$ at private cost $C(e)$, with $C(\underline{e}) = 0$ and $C(\overline{e}) = c_e > 0$. He also makes an observable social investment decision, y, that improves social and environmental outcomes but decreases the firm's profits. Cash flows or "profits" $\tilde{x}$ and "social output" $\tilde{v}$ are respectively defined as:[19]

$$\begin{aligned} \tilde{x} &= e - \theta y^2 + \tilde{\epsilon}_x \quad && \text{where } \tilde{\epsilon}_x \sim \mathcal{N}(0, \sigma_x^2) \\ \tilde{v} &= \eta y + \tilde{\epsilon}_y && \text{where } \tilde{\epsilon}_y \sim \mathcal{N}(0, \sigma_y^2) \end{aligned} \tag{1}$$

where θ is a positive constant. Cash flows are realized at $t = 2$ and paid out to shareholders. The effect of the social investment on social output depends on η, which is the unobserved realization of the random variable $\tilde{\eta} \sim \mathcal{N}(\bar{\eta}, \sigma_\eta^2)$, where $\bar{\eta} \geq 0$. It represents the firm's "social productivity".[20] All random variables, $\tilde{\eta}$, $\tilde{\epsilon}_x$, $\tilde{\epsilon}_y$, are independent.

The board's objective function is:

$$\mathbb{E}\left[\tilde{x} + \alpha_B \tilde{v} - \tilde{W}\right], \tag{2}$$

where the social output $\tilde{v}$ of the firm is weighted by $\alpha_B \geq 0$, and W is the manager's contractual

[19] In previous versions of the paper, we allowed for two different dimensions of social performance with an additively separable impact on profits and overall social output. The results were qualitatively unchanged. In section A.1 of the Online Appendix, we extend the model to allow for different dimensions of social performance.

[20] The productivity of a social investment can be negative. This means that allocating resources to increase provision of this type of social goods in this firm would decrease its social output. This is for tractability but can sometimes be justified. For example, if gender parity is an objective and the firm currently employs more men ($\eta > 0$), it means that hiring more women ($y > 0$) will help achieve this objective; however, if the firm currently employs more women ($\eta < 0$), it means that hiring more men ($y < 0$) will help achieve this objective.

payment. The notion that shareholders partly internalize the externalities generated by the firm is consistent with the empirical findings of Homroy, Mavruk, and Nguyen (2023) on the use of SEP measures in executive pay. We discuss further this specification of the board's preferences in section 7.4.

3.2 Measures of social performance and stock price

Social output υ and the social investment y are non-contractible, but an imperfect measure of social performance is realized at $t = 1$. This measure m of the firm's SEP, that we will refer to as an "ESG score" (see section 2.1 for institutional details) for brevity but without loss of generality, is such that:

$$\tilde{m} \equiv \tilde{\gamma} y \qquad \text{where } \tilde{\gamma} \sim \mathcal{N}(\eta, \sigma_\gamma^2), \tag{3}$$

where η is the realization of $\tilde{\eta}$. Equivalently, the measure can be decomposed as: $\tilde{m} = (\tilde{\eta} + \tilde{\zeta})y$, where $\tilde{\zeta} \sim \mathcal{N}(0, \sigma_\gamma^2)$, which is independent from $\tilde{\eta}$, is the unobservable noise in the ESG score. With $\sigma_\gamma^2 > 0$, γ is a noisy measure of the firm's social productivity. The difference between γ and η is simply the difference between the measured social impact and the actual social impact of social investments.

A publicly observable financial report z, which is imperfectly informative about the firm's profitability, is realized at $t = 1$ such that: $\tilde{z} = e - \theta y^2 + \epsilon_x + \tilde{\epsilon}_z$ where $\tilde{\epsilon}_z \sim \mathcal{N}(0, \sigma_z^2)$.

The stock price is also realized at $t = 1$. To account for the potential impact of corporate social performance on the stock price (see Barber, Morse, and Yasuda (2021), and our discussion in section 7.4), we assume an exogenous stock price function that puts a weight α_I on social performance relative to profits:

$$p \;=\; \mathbb{E}\left[\tilde{x} + \alpha_I \tilde{\upsilon} \,|\, y, z, m\right] \tag{4}$$

where the conditional expectation is taken with respect to the information available at $t = 1$, including the social investment y, the report z, and the ESG score m. Thus, when $\alpha_B = \alpha_I$, the board's preferences are perfectly aligned with stock price maximization. When $\alpha_B > \alpha_I$ ($\alpha_B < \alpha_I$), the board puts more (less) weight on social performance than the stock price. In Appendix A, we provide a microfoundation for equation (4) based on portfolio choice by socially responsible investors.

3.3 Contracting

At $t = 0$, the board offers a compensation contract to a risk neutral manager who has an outside option worth $\bar{W} \geq 0$. The manager receives a fixed wage w, and his compensation is linear in the following contractible performance measures: firm profits (with sensitivity β_x), stock price (with sensitivity β_p), and ESG score (with sensitivity β_m), similar to Holmström and Milgrom (1991) and Holmström and Tirole (1993).

We assume that the sensitivity of pay to profits and to the ESG score must be nonnegative, i.e. $\beta_x \geq 0$ and $\beta_m \geq 0$.[21] The former ($\beta_x \geq 0$) can be motivated similarly to Innes (1990): the manager would destroy output otherwise. The latter ($\beta_m \geq 0$) can be motivated by the public outcry that would likely result if a manager's pay were decreasing in a measure of a firm's SEP – similar to the political constraints mentioned by Jensen and Murphy (1990).

We assume that $\overline{e} - c_e - \bar{W} > 0$, i.e. a firm which hires a manager who exerts high effort can be profitable, and that the cost of high effort for the manager, c_e, is sufficiently low that it is optimal to induce high effort in all settings considered. The discount rate is zero.

4 Contracting without gaming

In this section, we assume that the manager does not observe γ at the time of making effort and investment decisions: he does not have any information that would allow him to "game" incentives. The timeline of the model is depicted in Figure 1.

[21]As will be clear in Proposition 1, the sensitivity of compensation to the stock price will be positive ($\beta_p \geq 0$), although it could be negative without the constraint that $\beta_x \geq 0$.

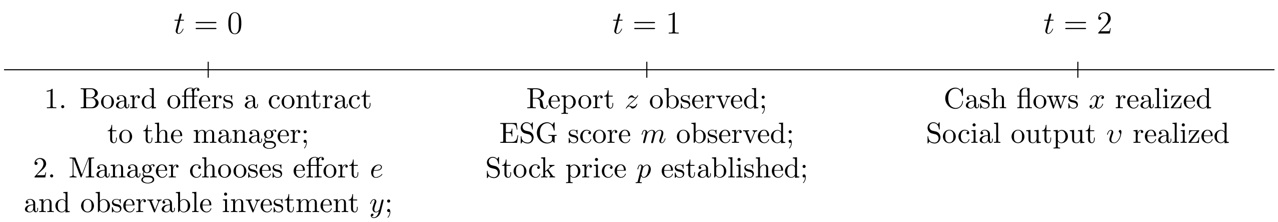

Figure 1: Timeline of the model without gaming.

Lemma 1 determines the $t = 1$ stock price.

Lemma 1 *The stock price p in equation (4) is such that:*

$$\mathbb{E}\left[\tilde{x} \,|\, y, z, m\right] = \hat{e} - \theta y^2 + \frac{\sigma_x^2}{\sigma_x^2 + \sigma_z^2}\left(z - \hat{e} + \theta y^2\right) \quad (5)$$

$$\mathbb{E}\left[\tilde{v} \,|\, y, z, m\right] = \frac{\sigma_\eta^2}{\sigma_\eta^2 + \sigma_\gamma^2} m + \frac{\sigma_\gamma^2}{\sigma_\eta^2 + \sigma_\gamma^2} y\bar{\eta} \quad (6)$$

The stock price is additive in two terms. The first is expected profits. It depends on the manager's expected effort $\hat{e}$, social investment y, and the report z. The report z is informative about profits $\tilde{x}$. As a result, it affects the stock price. This makes the stock price sensitive to the manager's actual effort (as opposed to the manager's expected effort $\hat{e}$, although the two coincide in equilibrium), so that the stock price provides effort incentives.

The second term is the expected social output of the firm. The perceived SEP of the firm as reflected in its ESG score affects the stock price when $\alpha_I > 0$ (see equation 4). This is consistent with the fact that investors rely on these scores and ratings for their investment decisions (Pagano et al. (2018), Berg, Kölbel, and Rigobon (2022)). Lemma 1 shows that, with $\alpha_I > 0$, there is a positive relation between the ESG score and the stock price in equilibrium, consistently with the empirical evidence (Berg et al. (2021)). Intuitively, a higher score is good news about the firm's social productivity. However, the ESG score is a noisy measure, so that it does not reveal the firm's actual social productivity. As a result, the expected social output of the firm ($\mathbb{E}[\tilde{v} \,|\, y, z, m]$) does not only rely on its ESG score m and its social investment y, but also on the prior belief $\bar{\eta}$ about the productivity of this investment.

Lemma 2 *When γ is not observable at $t = 0$, the first-best outcome involves $e = \overline{e}$ and $y^* = \frac{\alpha_B}{2}\frac{\bar{\eta}}{\theta}$. It is achieved with the following contract:*

$$\beta_m = \alpha_B \frac{c_e}{\overline{e} - \underline{e}}, \quad \beta_x = \frac{c_e}{\overline{e} - \underline{e}}, \quad \beta_p = 0, \tag{7}$$

and a fixed wage w such that the manager is at his reservation level of utility given these values of $\{\beta_x, \beta_p, \beta_m\}$ and high effort.

The contract described in Lemma 2 provides effort incentives via a positive sensitivity of pay to profits ($\beta_x > 0$). To counteract the resulting incentives to increase profits rather than social performance, the manager's compensation is also contingent on social performance ($\beta_m > 0$). The first-best can always be obtained, even if social performance is poorly measured at $t = 1$ (high σ_γ). Intuitively, there is no asymmetric information ex-ante, so that the contract can be set to provide first-best investment incentives, elicit effort, and keep the manager at his reservation utility. Without gaming, imperfect performance measurement does not lead to inefficiencies.

Moreover, in this case, there are no benefits to using stock price-based compensation. This being said, stock price-based compensation may alternatively and equivalently be used: when $\alpha_B \geq \alpha_I$, the first-best optimal outcome can also be induced with the following contract that involves stock price-based compensation instead of profits-based compensation:

$$\beta_m = (\alpha_B - \alpha_I)\left(1 + \frac{\sigma_z^2}{\sigma_x^2}\right)\frac{c_e}{\overline{e} - \underline{e}}, \quad \beta_x = 0, \quad \beta_p = \left(1 + \frac{\sigma_z^2}{\sigma_x^2}\right)\frac{c_e}{\overline{e} - \underline{e}}, \tag{8}$$

and the fixed wage w adjusts to satisfy the participation constraint. In sum, when the manager does not have private information, the contracting problem has a simple solution that does not involve any inefficiencies.

5 Contracting with gaming

We henceforth assume that, at $t = 0$, after contracting but before the effort and investment decisions, the manager observes the nonverifiable variable γ. This timing assumption is similar to the assumption in Baker (1992) and Edmans and Gabaix (2011) that the agent chooses his action after observing the noise. In our model, it parsimoniously captures the notion that, because of his on-the-job expertise and his understanding of the SEP measures' methodology, the manager understands how investment decisions will affect SEP measures. It allows to take into account a frequent criticism leveled at SEP-based incentives (that they will encourage "gaming"). It represents a departure from standard models of contracting and multitasking with noisy ex-post performance measurement in which the principal is aware of any biases that the manager may have at the contracting phase. The timeline of the model is depicted in Figure 2.

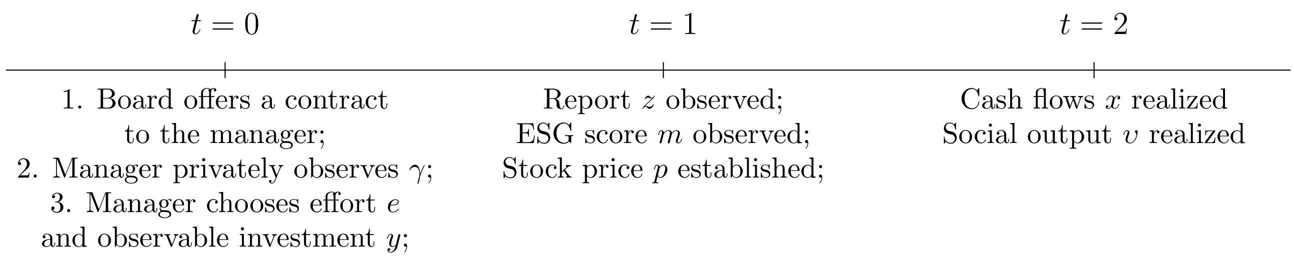

Figure 2: Timeline of the model with gaming.

We solve the model by backward induction. Since investors observe the same variables at $t = 1$ as in the previous section, the stock price is as in Lemma 1.[22] For publicly traded firms, the stock price is established on stock markets. For private firms, it represents the implied stock valuation established in a sale or a funding round.

We now consider incentive provision and contracting. The board wants to incentivize effort and social investment. To this end, it can use three performance measures: the firm's stock price, its profits, and its ESG score.

As a preliminary step, we define the first-best outcome, which provides a useful benchmark. The first-best outcome refers to the outcome without information asymmetries and without an agency problem, i.e. when there is no incentive constraint in the optimization problem. At the second-best, let $y(\gamma)$ be the social investment optimally chosen by the manager given his contract and his signal γ. Let ϕ denote the density function of $\tilde{\eta}$, and φ denote the conditional density function of $\tilde{\gamma}$.

[22]The only potential difference would have been if the manager's actions at $t = 0$ (e.g. choice of social investment y) were informative about something not observed by investors at $t = 1$. However, investors observe both y and the ESG score m at $t = 1$, so that they can infer γ.

Lemma 3 *When γ is observable at $t = 0$, the first-best social investment is:*

$$y^*(\gamma) = \frac{\alpha_B}{2}\frac{\mathbb{E}[\tilde{\eta}|\gamma]}{\theta} \quad \text{where} \quad \mathbb{E}[\tilde{\eta}|\gamma] = \frac{\sigma_\eta^2}{\sigma_\eta^2 + \sigma_\gamma^2}\gamma + \frac{\sigma_\gamma^2}{\sigma_\eta^2 + \sigma_\gamma^2}\bar{\eta} \tag{9}$$

At the second-best, when the manager privately observes γ at $t = 0$, the board sets the manager's contract to ensure participation, induce effort, and minimize the agency cost in equation (10):

$$\theta \int_\eta \int_\gamma (y(\gamma) - y^*(\gamma))^2 \varphi(\gamma|\eta)\phi(\eta) d\gamma d\eta \tag{10}$$

We refer to the difference between the board's objective function at the first-best and at the second-best as the "agency cost". Given that the manager is risk neutral and there are no constraints on contracting, the agency cost is not driven by inefficient risk sharing or rent extraction.[23] Instead, the agency cost reflects the extent of the inefficiency in resource allocation. As can be seen in equation (10), it measures the deviation between the social investment and its first-best level, which is defined in equation (9).

Lemma 3 shows that minimizing the agency cost is equivalent to minimizing the sum across SEP dimensions of the expected quadratic distance between the incentive-compatible social investment $y(\gamma)$ and the first-best social investment $y^*(\gamma)$ multiplied by the cost parameter θ. This distance is a measure of the agency cost: it measures how close the board can get to the first-best outcome described in equation (9). The agency cost is proportional to the monetary cost of social investments, as measured by θ. Intuitively, if social investments were costless, there would be no tradeoff between social investments and profits. On the contrary, the more costly social investments are, the more expensive is any deviation from the first-best level.

The contracting problem is a priori not simple for several reasons. First, a socially responsible board ($\alpha_B > 0$) would like to provide incentives for effort as well as for social investment. In doing so, it faces a multitasking problem in which the sensitivity of pay to performance must be high enough to elicit effort, and the balance of incentives matters. Second, because of the manager's private information about the firm's technology for social output and his understanding of the ESG scoring methodology, ESG score-based incentives will result in "gaming". Third, the level

[23] Risk sharing and rent extraction effects are arguably of second-order importance in large firms, where managerial equity holdings account for only 0.34% of firm equity for the median CEO (Edmans, Gabaix, and Jenter (2017)), while annual CEO compensation typically accounts for less than 0.1% of the market capitalization of S&P500 firms.

of social investment that maximizes the stock price does not necessarily maximize the board's objective function. Fourth, the sensitivity of the manager's compensation to profits and the ESG score must be nonnegative.

To focus on the case of a socially responsible board and simplify the exposition, we henceforth assume $\alpha_B > 0$. As a second preliminary step, Lemma 4 highlights the important role played by the constraint that the sensitivity of pay to profits be non-negative. It shows that, without this constraint, the first-best outcome can be achieved even when the manager privately observes γ at $t = 0$.

Lemma 4 *Without a nonnegativity constraint on β_x, the board can achieve the first-best outcome ($y(\gamma) = y^*(\gamma)\ \forall\gamma$ and $e = \overline{e}$) by offering a contract such that:*

$$\beta_m = 0, \quad \beta_x = \left(\frac{\alpha_I}{\alpha_B} - 1\right)\beta_p \quad \textit{and} \quad \beta_p = \frac{c_e}{\overline{e} - \underline{e}}\left(\frac{\alpha_I}{\alpha_B} + \frac{\sigma_x^2}{\sigma_x^2 + \sigma_z^2} - 1\right)^{-1} \tag{11}$$

If the board's preferred level of social investment is lower than the one that maximizes the stock price ($\alpha_B < \alpha_I$), then only relying on stock price-based incentives would lead to excessive spending on social investment from the board's perspective. To counterbalance stock price-based incentives, the board then provides profits-based incentives ($\beta_x > 0$), which discourage costly social investments. On the contrary, if the board's preferred level of social investment is higher than the one that maximizes the stock price ($\alpha_B > \alpha_I$), then only relying on stock price-based incentives would lead to inadequate spending on social investment from the board's perspective. Encouraging social investments further can then be achieved by punishing the manager for achieving high profits, i.e. $\beta_x < 0$. The sensitivity of compensation to the stock price increases as needed to still provide effort incentives. In any case, the agency cost is zero: the firm's social investment, $y(\gamma)$, is state-by-state equal to the first-best level $y^*(\gamma)$ defined in equation (9).

We now describe the second-best optimal linear contract, with a nonnegativity constraint on β_x, in Proposition 1.

Proposition 1

- *If $\alpha_B \leq \alpha_I$, then $y(\gamma) = y^*(\gamma)\ \forall\gamma$, and the optimal linear contract is defined by:*

$$\beta_m = 0, \quad \beta_x = \left(\frac{\alpha_I}{\alpha_B} - 1\right)\frac{c_e}{\overline{e} - \underline{e}}\left(\frac{\alpha_I}{\alpha_B} + \frac{\sigma_x^2}{\sigma_x^2 + \sigma_z^2} - 1\right)^{-1}, \quad \beta_p = \frac{c_e}{\overline{e} - \underline{e}}\left(\frac{\alpha_I}{\alpha_B} + \frac{\sigma_x^2}{\sigma_x^2 + \sigma_z^2} - 1\right)^{-1}.$$

- *If $\alpha_B > \alpha_I$, then generically $y(\gamma) \neq y^*(\gamma)$, the expected social investment is below the first-best level, i.e. $\mathbb{E}[y(\tilde{\gamma})] < \mathbb{E}[y^*(\tilde{\gamma})]$, and the optimal linear contract is defined by:*

$$\beta_m = (\alpha_B - \alpha_I)\frac{c_e}{\bar{e} - \underline{e}}\left(1 + \frac{\sigma_z^2}{\sigma_x^2}\right)\left(\frac{\sigma_\gamma^2}{\sigma_\eta^2 + \sigma_\gamma^2}\frac{\bar{\eta}^2}{\sigma_\gamma^2 + \sigma_\eta^2 + \bar{\eta}^2} + \frac{\sigma_\eta^2}{\sigma_\eta^2 + \sigma_\gamma^2}\right), \quad \beta_x = 0, \quad \beta_p = \frac{c_e}{\bar{e} - \underline{e}}\left(1 + \frac{\sigma_z^2}{\sigma_x^2}\right).$$

The structure of the contract described in Proposition 1 can be explained as follows. First, the fixed wage adjusts to keep the manager at his reservation level of utility including compensation for the cost of effort. As a result, the manager does not derive rents. Second, the sensitivity of pay to the performance measures available has implications for the effort and investment decisions. A positive sensitivity of pay to profits encourages effort and discourages social investments. A positive sensitivity of pay to the ESG score tends to encourage social investment.[24] A positive sensitivity of pay to the stock price encourages the manager to exert effort and to choose the social investment that maximizes the stock price.

To start, consider the case when the board prefers the social investment that maximizes the stock price, i.e. $\alpha_B = \alpha_I$. In this case, the manager's compensation is only contingent on the stock price: $\beta_m = 0$, $\beta_x = 0$, and $\beta_p = \frac{c_e}{\bar{e}-\underline{e}}\left(1 + \frac{\sigma_z^2}{\sigma_x^2}\right)$. When managerial compensation is only sensitive to the stock price, a self-interested manager optimally chooses the social investment that maximizes the stock price, in line with the board's preference. The sensitivity of compensation to the stock price is then determined to provide adequate incentives for effort.

Now consider the case when the level of social investment that the board prefers is lower than the one which maximizes the stock price ($\alpha_B < \alpha_I$). For example, as further discussed in section 7.4, this could be the case if investors enjoy sufficiently strong warm-glow utility from investing in socially responsible firms, or if there are many socially responsible funds with impact mandates (Oehmke and Opp (2024)). In this case, the manager's compensation is contingent on profits and the stock price. A positive sensitivity of compensation to profits discourages social investment relative to the level that would maximize the stock price. Since both profits-based and stock price-based compensation encourage managerial effort, a positive sensitivity of compensation to profits reduces the sensitivity of compensation to the stock price required to elicit managerial effort. In this case, the board can still induce a level of social investment y that corresponds to the first-best

[24]Because of the imperfection of ESG scores, they do not *always* encourage social investment. In the case when the sign of γ is opposite the sign of η, imperfect measurement that can be anticipated ex-ante gives rise to counterproductive incentives.

level y^* state-by-state (for any realization of $\tilde{\gamma}$).

Finally, consider the case when the level of social investment that the board prefers is higher than the one which maximizes the stock price ($\alpha_B > \alpha_I$). In particular, this is the relevant case if investors do not intrinsically value holding equity in a socially responsible firm, so that SEP does not have a stock price impact ($\alpha_I = 0$), even though they would be better off if the firm reduced its externalities. Then a board with $\alpha_B > 0$ solves a coordination problem. More generally, whenever $\alpha_B > \alpha_I$, the manager's compensation is contingent on the stock price and the ESG score. A positive sensitivity of compensation to the ESG score encourages social investment relative to the level that would maximize the stock price.

Ideally, the board would like the social investment to depend on its average productivity $\bar{\eta}$, and on the signal γ that the manager receives on its actual productivity (see the first-best investment policy in equation (9)). The stock market's valuation of the firm's social investments combines these two aspects (see Lemma 1). Moreover, the stock price aggregates information about the firm's social output efficiently for investment purposes, since it reflects the effect of the ESG score on beliefs about the productivity of the firm's social investment. However, the stock price's aggregation of information about the firm's social output and its profits does not correspond to the board's preference when $\alpha_B \neq \alpha_I$. When $\alpha_B > \alpha_I$, stock price-based incentives are excessively tilted toward profits maximization relative to the board's preference. This cannot be remedied by a negative sensitivity of managerial compensation to profits: the nonnegativity constraint on profits-based compensation binds.

This can be partly remedied by also making managerial compensation contingent on the firm's ESG score, so as to complement the social investment incentives already embedded in stock price-based compensation. However, by definition, ESG score m only depends on social investment y and the signal γ. The latter is an imperfect signal of the firm's technology for social output (since $\sigma_\gamma > 0$), and it is known ex-ante by the manager. For example, when γ is high, the manager understands that increasing social investment y will have a large impact on ESG score m, likely over and beyond the actual impact (ηy). Thus, making compensation sensitive to the ESG score will lead the manager to be excessively responsive to realizations of the signal γ, which can be viewed as "gaming". This inefficiency reduces the sensitivity of the manager's compensation to the ESG score: this sensitivity is decreasing in the noisiness of scores, as parameterized by σ_γ.[25]

[25]In the model without gaming in section 4, when $\alpha_B > \alpha_I$ and stock price-based compensation is used to elicit

As a result, the firm generally underinvests in social investments relative to what the board would prefer.[26] Intuitively, the difficulty of aligning interests with respect to SEP lowers the second-best level of expected social investment below the first-best level. In summary, the more noisy the ESG score is, the more distorted are the manager's incentives for social investments, and the less the board incentivizes social investments.

The deviation from the first-best outcome depending on whether $\alpha_B \gtreqless \alpha_I$ is asymmetric. When $\alpha_B \leq \alpha_I$, it is possible to reach the first-best outcome using available compensation instruments – with profits-based and stock price-based compensation. On the contrary, when $\alpha_B > \alpha_I$, it is impossible to do so. In this latter case, the firm will use ESG score-based as well as stock price-based compensation, and social investment will be distorted for the reasons mentioned in preceding paragraphs.

This result relies on a wedge between the observability and the contractibility of social investments, so that the stock price aggregates information in a way that cannot be replicated by a contract. This is consistent with the notion that the stock price provides incremental information that is useful for contracting (Holmström and Tirole (1993)). For simplicity, we have assumed that social investments are not contractible, but this result would still hold under the more realistic assumption that social investments are not fully contractible. The logic is reminiscent of Parlour and Rajan (2020). In their model, a credit rating does not provide any new information, but it allows to write better contracts since it is contractible. In our model, the stock price plays this role.

Overall, the main forces at play in the model can be summarized as follows. Equity-based and profits-based compensation are used as substitutes to elicit managerial effort. At the same time, a board that wants to encourage social investments ($\alpha_B > 0$) starts by relying on stock price-based incentives (if the stock price is sensitive to perceived SEP, i.e. $\alpha_I > 0$). The social incentives thus provided can be adjusted by changing the relative weights of stock price-based and profits-based compensation while preserving effort incentives. Only at the point when β_x cannot be decreased further (i.e. $\beta_x = 0$) does the board start using measures of SEP. Indeed, since it will foster inefficient gaming, using these measures in a contract is less efficient than relying on equity-based

effort, β_m in equation (8) is strictly higher than β_m in the third case of Proposition 1. Morever, in the model without gaming, the level of β_m is independent of the noisiness σ_γ of SEP measurement.

[26]Specifically, the expected level of social investment is below its expected level at the first-best, see Proposition 1. The firm can end up either underinvesting or overinvesting, depending on the realization of $\tilde{\gamma}$.

compensation.

Proposition 1 generates empirical implications about the use of SEP measures in incentive programs. First, it suggests that firms should not necessarily use explicit incentives based on SEP measures, even if their boards are socially responsible. Indeed, measures of SEP are only used when the board's preferred level of social investments exceeds the one that maximizes the stock price. This suggests that compensation based on SEP measures and socially responsible investors are substitutes rather than complements. Second, the intrinsic quality of SEP measures, captured by σ_γ in the model, does not matter for the inclusion of these measures in the contract. However, it matters for the sensitivity of pay to these measures when they are used. The worse the quality of these measures (i.e. the higher σ_γ), the lower is the sensitivity of pay to these measures (see Proposition 1).

6 Multiple ESG scores

We now extend the model to analyze the case with multiple ESG scores. In practice, several ESG rating agencies provide ESG scores and ratings. These scores are informative, in the sense that they affect firms' stock prices (Berg et al. (2021)), but there is evidence of substantial divergence across these ratings, including on the measurement of the same dimension of SEP (Berg, Kölbel, and Rigobon (2022)).

We now assume that there are N ESG raters. Each rater j provides ESG score m_j, which is defined as $m_j \equiv y\gamma_j$. We will present results based on two different assumptions on the joint distribution of these scores.

To start, we analyze the effect of increasing the number of ESG scores with uncorrelated noise terms. Denoting $\tilde{\zeta}_j \sim \mathcal{N}(0, \sigma_\gamma^2)$, for any score j we can decompose the score as $\tilde{m}_j = y(\tilde{\eta} + \tilde{\zeta}_j)$, where the subscript j indicates that the noise term $\tilde{\zeta}_j$ is different for each rating j. We assume that the noise terms $\tilde{\zeta}_j$ in ESG scores are independent and identically distributed (i.i.d.). This implies that ESG scores are positively correlated because of their dependence on $\tilde{\eta}$. Let $\bar{\gamma} = \frac{1}{N}\sum_{j=1}^{N}\gamma_j$ be the average signal on the firm's social productivity generated by ESG scores, and redefine the agency cost for this part of the paper accordingly. This average signal is a sufficient statistic for the mean of the distribution (see the proof of Proposition 2).

Proposition 2 *For $\alpha_B \leq \alpha_I$, social investment is equal to its first-best level. As N increases, social investment and expected social output converge asymptotically to their first-best levels:*

$$\lim_{N\to\infty} y(\bar{\gamma}) = \frac{\alpha_B}{2\theta}\bar{\gamma} \qquad \text{and} \qquad \lim_{N\to\infty} \mathbb{E}[\tilde{\eta}\tilde{y}] = \frac{\alpha_B}{2\theta}\left(\bar{\eta}^2 + \sigma_\eta^2\right). \tag{12}$$

Proposition 2 implies that increasing the number of ESG scores allows to overcome the agency problem. The reason is as follows. In this setting, the average score is a sufficient statistic. Therefore, the manager considers the average score when making the investment decision. Additional scores reduce the variance of the bias of the average score, which increases its quality, and decreases the manager's propensity to game the ESG scoring system. Intuitively, with a single score, incentives are distorted in one direction only. With additional scores which are imperfectly correlated with the first score, incentives are still distorted but not necessarily in the same direction. Proposition 2 shows that, as the number of ESG scores increases, investment distortions vanish in the limit.

As the number of scores increases, the total sensitivity of compensation to ESG scores increases (see equation (64) in the Appendix) when these scores are used in the contract (i.e. $\alpha_B > \alpha_I$). Intuitively, multiple scores reduce the possibility of gaming, which allows to provide higher-powered incentives.

Figure 3 depicts the speed of convergence of the agency cost toward zero in some cases. It illustrates how it depends on model parameters when the noise terms of ESG scores are i.i.d.. When scores are highly noisy (see the left panel), the speed of convergence is slow. When scores are not very noisy (see the middle and right panels), the speed of convergence is fast, so that the marginal benefit of an additional score decreases quickly with the number of scores. In sum, the practical relevance of the mechanism at play in Proposition 2 depends on the noisiness of ESG scores.

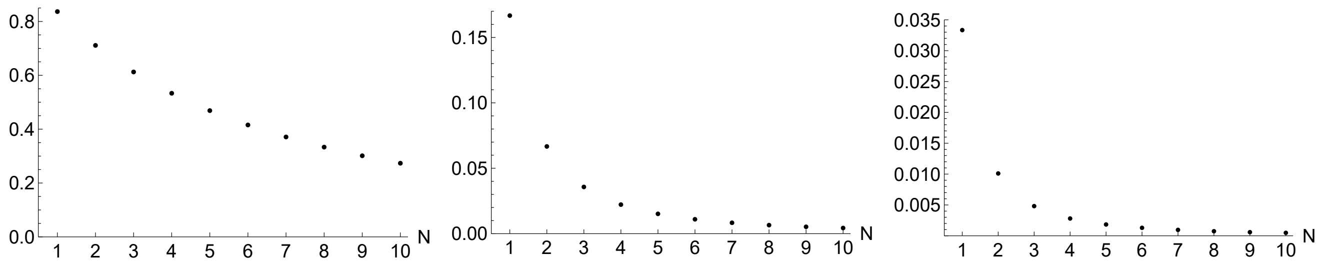

Figure 3: Each panel depicts the agency cost (as derived in equation (67) in the Appendix) as a function of the number N of ESG scores for $\alpha_B > \alpha_I$ with the normalization $(\alpha_B - \alpha_I)^2/2\theta = 1$. Left panel: $\sigma_\eta^2 = 1$, $\sigma_\gamma^2 = 4$, $\bar{\eta} = 1$; middle panel: $\sigma_\eta^2 = 1$, $\sigma_\gamma^2 = 1$, $\bar{\eta} = 1$; right panel: $\sigma_\eta^2 = 2$, $\sigma_\gamma^2 = 1$, $\bar{\eta} = 1$.

The mechanism can be contrasted with the one in Ederer, Holden, and Meyer (2018). In their model, to avoid gaming, the principal randomizes over which dimension the agent will be paid. A risk averse agent then works on all dimensions. In our model, to avoid gaming, the principal uses several performance measures. Their mechanism relies on randomization and risk aversion with multiple tasks, whereas ours relies on multiple performance measures with a single investment decision.

An important question related to SEP measurement if whether adding an ESG score with a lower quality than an existing score can be beneficial. Is it better to have one high-quality ESG score, or is it valuable to complement it with an additional score whose bias can be correlated with the former score's bias and which can have a lower quality (higher noise)?

We now analyze the effect of adding to an existing score some possibly more noisy and correlated ESG score. There are two scores, $\tilde{m}_1 = \tilde{\gamma}_1 y$ and $\tilde{m}_2 = \tilde{\gamma}_2 y$, that follow a multivariate normal distribution. As above, we can decompose the score as $\tilde{m}_j = (\tilde{\eta} + \tilde{\zeta}_j)y$, where the subscript j indicates that the noise term $\tilde{\zeta}_j$ is different for each score j. Denote by $\rho \in (-1, 1)$ the correlation coefficient between the noise terms $\tilde{\zeta}_1$ and $\tilde{\zeta}_2$. The noise terms $\tilde{\zeta}_1$ and $\tilde{\zeta}_2$ follow a multivariate normal distribution with mean vector $(0, 0)^T$ and covariance matrix:

$$\begin{pmatrix} \sigma_\gamma^{1^2} & \rho\sigma_\gamma^1\sigma_\gamma^2 \\ \rho\sigma_\gamma^1\sigma_\gamma^2 & \sigma_\gamma^{2^2} \end{pmatrix}$$

Even with $\rho < 0$, the ESG scores $\tilde{m}_1$ and $\tilde{m}_2$ can still be positively correlated because of their common dependence on $\tilde{\eta}$.

We show in the proof of Proposition 3 that neither γ_1 nor γ_2 is a sufficient statistic for $\tilde{\eta}$, i.e. the additional ESG score is informative. However, it is not necessarily valuable. The informativeness principle (Holmström (1979)), which states that any informative signal is useful for contracting, does not apply in our setting without risk aversion or rent extraction.[27] Accordingly, we now analyze the effect of adding an ESG score on social investment and expected social output when the manager's compensation can linearly depend on both scores.

Proposition 3 *For $\alpha_B \leq \alpha_I$, social investment is equal to its first-best level. Moreover, expected social output is higher with two scores rather than one if the precision $1/\sigma_\gamma^{2\,2}$ of the second score is sufficiently high or the correlation ρ between the scores' noise terms is sufficiently low.*

Proposition 3 shows that the result from Proposition 1 that the first-best social investment can be induced if and only if $\alpha_B \leq \alpha_I$ is robust to the addition of an ESG score.[28] Proposition 3 also shows that adding an ESG score can increase the expected social output of the firm under some conditions. However, numerical examples depicted in Figure 4 show that adding an ESG score can also be detrimental. Intuitively, an additional score which is not perfectly correlated with the first score is informative, even if its quality is low, and it consequently affects the stock price. Since the stock price is used to provide incentives, the board cannot just "ignore" this additional score. The problem is that the bias of an additional score can be so strong (high σ_γ^2) and its correlation ρ with the first score so high that adding this score will tend to worsen the bias induced by SEP measures. Indeed, Figure 4 illustrates that introducing an additional ESG score can increase the agency cost if this ESG score is noisier and sufficiently highly correlated with the first score (see the blue area).[29]

[27]If the agent were risk averse, the contract would be designed not just to provide efficient effort and investment incentives, but also based on risk sharing considerations. Then, for the additional score to be useful, a necessary (but not always sufficient) condition is that it be incrementally informative with respect to the initial score and the stock price in a sufficient statistic sense.

[28]The first-best investment with two ESG scores, $y^*(\gamma_1, \gamma_2)$, is defined in equation (73) in the Appendix. The proof of Proposition 3 also shows that, in the limit case as $\rho \to -1$, expected social output tends to the same level as with perfect performance measurement ($\sigma_\gamma^j = 0$), regardless of $\alpha_B \gtrless \alpha_I$ and even if both scores are noisy.

[29]The black area in Figure 4 is the set of parameter values such that the nonnegativity constraint on the sensitivity of pay to SEP is binding for one of the two ESG scores. With two ESG scores, this constraint can be binding when the correlation ρ between the noise terms of the two scores, ζ_1 and ζ_2, is sufficiently high. Intuitively, the optimal contract without this nonnegativity constraint would have a positive weight on the less noisy score, and a negative weight on the more noisy score to partly filter out the noise.

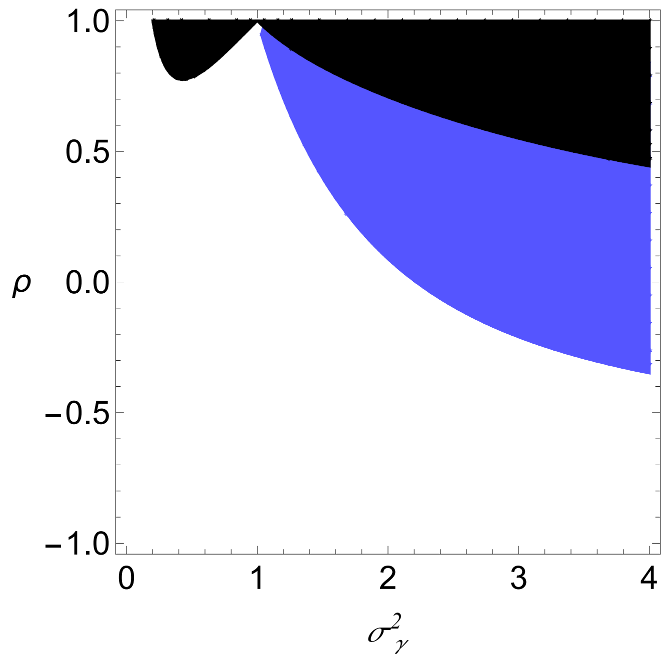

Figure 4: The blue area is the set of values of σ^2_γ and ρ such that the agency cost is lower with one rather than two ESG scores. The black area is the set of values of σ^2_γ and ρ such that the constraint that $\beta^j_m \geq 0$ is binding for $j = 1$ or $j = 2$, where β^j_m is the sensitivity of the manager's pay to ESG score j. We have: $\bar{\eta} = 1$, $\theta = 1$, $\alpha_B = 2$, $\alpha_I = 1$, $\sigma_\eta = 1$, $\sigma^1_\gamma = 1$.

In practice, the correlation between ESG ratings is on average 0.54 (Berg, Kölbel, and Rigobon (2022)), and it is even lower at a disaggregated level. This allows to put in perspective the practical relevance of the blue area in Figure 4. Indeed, 0.54 can be viewed as an upper bound for ρ (part of the correlation between ESG scores is driven by their common dependence on η, whereas ρ is only the correlation between the noise terms in ESG scores). If $\rho = 0.54$ and the standard deviation of the first ESG score is $\sigma^1_\gamma = 1$, the standard deviation σ^2_γ of the second ESG score would need to be greater than 1.34 for the blue area to be relevant, i.e. noise in the second ESG score must be at least one-third higher than in the first score. If $\rho = 0.27$, the blue area is relevant for $\sigma^2_\gamma > 1.67$, i.e. noise in the second ESG score must be at least two-thirds higher than in the first score.[30]

The results in this section have implications for SEP measurement "convergence" or "harmo-

[30]To take another perspective, when the two ESG scores are as noisy, so that both noise terms have a standard deviation of $\sigma^j_\gamma \equiv \sigma_\gamma$, the correlation coefficient between ESG scores $\tilde{m}_1$ and $\tilde{m}_2$ is simply: $\text{corr}_{\tilde{m}_1,\tilde{m}_2} = \frac{cov(\tilde{m}_1,\tilde{m}_2)}{\sqrt{var(\tilde{m}_1)}\sqrt{var(\tilde{m}_2)}} = \frac{\sigma^2_\eta+\rho\sigma^2_\gamma}{\sigma^2_\eta+\sigma^2_\gamma}$. When in addition the variability of the ESG scores around their mean is driven as much by fundamentals and noise (i.e. $\sigma_\eta = \sigma_\gamma$), and $\rho = 0$, we have $\text{corr}_{m_1,m_2} = \frac{1}{2}$, which is close to the observed correlation of 0.54 between ESG ratings.

nization", which is debated and frequently advocated.[31] On the one hand, Proposition 2 shows that different ESG scores with i.i.d. noise terms can be beneficial for corporate governance purposes. If regulation or "harmonization" of ESG measurement results in a decrease in the number of available ESG scores and ratings, our results suggest that it may result in more gaming of the remaining ESG scores and ratings by managers, for example via greenwashing. On the other hand, ESG scores with sufficiently low quality are detrimental, because they affect the stock price and distort managerial incentives. Together, these results suggest that regulatory efforts should focus on improving ESG scores' quality.

Harmonization of SEP measurement might admittedly improve the quality of ESG scores. Our results suggest that this increase would need to be sufficiently high to offset the increased gaming effect resulting from a decrease in the number of scores that we highlighted. In other words, an improvement in the quality of ESG scores is necessary but insufficient for harmonization to be beneficial. In the case when the noise in scores is i.i.d. as in Proposition 2, a hypothetical unique standardized score would need to have a precision which is higher than N times the precision of an individual score if it replaces N different scoring methodologies. This sets a high bar for harmonization via regulation.

7 Discussion

7.1 Multiple dimensions of SEP

For simplicitly, the model presented in this paper only included one dimension of SEP. In practice, companies do not only invest along one social or environmental dimension. In section A.1 of the Online Appendix, we extend the model to incorporate different dimensions of SEP. As long as these dimensions do not interact (e.g. investing in greenhouse gas reductions does not affect the cost of improving working conditions), we show that the main results continue to hold. We point out that our results are then relevant for a dimension of SEP instead of overall SEP. In

[31] There is a regulatory push for uniform standards in SEP measurement and reporting, including government mandated CSR (Corporate Social Responsibility) reporting in the European Union (Fiechter, Hitz, and Lehmann (2022)). As stated by Larcker et al. (2022): "The major credit rating agencies . . . are subject to regulation by the Securities and Exchange Commission which requires covered firms to adhere to certain policies, procedures, and protections to reduce conflicts of interest and improve market confidence in their quality. Should ESG ratings be subject to similar requirements?" This harmonization argument is also made in influential academic research (Berg, Kölbel, and Rigobon (2022)).

particular, gaming can be mitigated with multiple SEP scores on each dimension of SEP.

7.2 SEP and financial performance

Our model postulates a tradeoff between improving SEP (social output v) and financial performance (profits x). This means that improving social performance involves the firm going "over and beyond" what would maximize its bottom line. For example, improving worker safety will to a large extent also improve financial performance. However, beyond a certain point, it will not. Our model considers safety improvements beyond this point, that have a negative effect on financial performance (safety improvements that also improve financial performance are a "no-brainer", and do not require any degree of social responsibility). This presents a challenge for the empirical literature, and more generally for CSR assessment, since it is not easy to disentangle social investments that improve the bottom line from those that do not.

7.3 Contracting with SEP

In practice, SEP incentives are mostly incorporated in annual incentive plans and annual bonuses (Walker (2022)). This is consistent with our assumptions about contracting. SEP incentives can also take other forms. For example, Duke Energy issued performance shares that vest in part depending on the firm's SEP performance (Walker (2022)). Moreover, there is a growing tendency to incorporate these metrics in long-term incentive plans (Tonello (2024)).

Large firms are much more likely to incorporate ESG metrics into their incentive plans. As noted by the Managing Director of ESG Research at The Conference Board (Tonello (2024)), "in 2023, only 23.8% of companies with annual revenue under \$100 million reported using ESG performance metrics, but the percentage grows to 83.6% among companies with annual revenue of \$50 billion and over." The same report notes that ESG performance metrics are commonly used in the utilities and energy sectors, but are much less common in information technology, consumer staples, financials, and communication services. Empirical studies have documented similar patterns (Flammer, Hong, and Minor (2019), Ikram, Li, and Minor (2023)).

Companies typically rely on a combination of several metrics. Across all industries, the most commonly used metrics involve human capital management. In the energy and utilities sectors, environmental performance metrics are also heavily used (Tonello (2024)). Moreover, the proportion of companies that use environmental metrics such as carbon footprint, emission reduction, and energy efficiency, is rapidly growing. Institutional investors and proxy advisors are increasingly pushing companies to adopt quantitative rather than qualitative metrics to improve accountability and transparency.[32] Recently, some companies have removed controversial ESG metrics, most notably Diversity, Equity and Inclusion (DEI) targets, from their incentive plans, but the trend for the inclusion of other SEP metrics is still positive.[33]

7.4 Preferences for SEP

Preferences for SEP play a crucial role in the model. They drive its predictions.

We assumed that shareholders (as represented by the board) and stock market investors intrinsically value the social and environmental impact of the firm. This can be justified based on the empirical evidence that investors have social and environmental concerns, and that they are willing to sacrifice financial return to this end (Riedl and Smeets (2017), Hartzmark and Sussman (2019), Barber, Morse, and Yasuda (2021), Bauer, Ruof, and Smeets (2021), Bolton and Kacperczyk (2021), Humphrey et al. (2023), Haber et al. (2022), Baker, Egan, and Sarkar (2022), Heeb et al. (2023)).

The modelization of social preferences of the board (in equation (2)) and of investors is similar to the one in Pástor, Stambaugh, and Taylor (2021), Broccardo, Hart, and Zingales (2022), Friedman, Heinle, and Luneva (2022), Geelen, Hajda, and Starmans (2022), Goldstein et al. (2022), and Dewatripont and Tirole (2022). This is also similar to the modelization of altruism in Gaynor et al. (2023), and consistent with the empirical evidence on ESG-linked compensation (Homroy, Mavruk, and Nguyen (2023)).

A discrepancy between parameters α_B and α_I can arise because the board's interests are not fully aligned with shareholders', or because the firm's shareholders are not necessarily the same economic agents as investors who actively trade on financial markets. If investors have warm-glow preferences and intrinsically value holding equity in a socially responsible firm, of if they represent fund managers with impact mandates for socially responsible investments, then their stock market investments have a stock price impact captured by $\alpha_I > 0$. On the contrary, if

[32] Source: 2024 Proxy Roundup: ESG Metrics in Incentive Compensation Plans, Harvard Law School Forum on Corporate Governance, August 5 2024.

[33] Source: Companies drop DEI targets from bonus plans on pressure from conservatives, *Financial Times*, July 21 2024.

atomistic investors are affected by externalities, which they take as given, but do not intrinsically value socially responsible investments, then $\alpha_I = 0$ even though they would be better off if the firm reduced externalities. In this latter case, a board with $\alpha_B > 0$ solves a coordination problem by encouraging externalities mitigation above and beyond the level that would maximize the stock price (Oehmke and Opp (2024)).

Section A.2 of the Online Appendix presents alternative interpretations of the model. In particular, α_B might not reflect the intrinsic preferences of the board. Instead, the board might set managerial incentives "as if" it were socially responsible (with $\alpha_B > 0$) as a way to commit to socially responsible policies. This could allow the firm to raise funding at a lower cost or to hire socially responsible employees.

7.5 Heterogeneous preferences with respect to SEP

In section A.1.2 of the Online Appendix, we analyze the outcome when the board and stock market investors don't put the same weight on all SEP measures. For example, the board might care more about the firm's carbon emissions but less about working conditions than stock market investors. In this case, the manager's contract can be more complex than in the baseline model. Indeed, it can simultaneously include stock price-based compensation, profits-based compensation to discourage excessive investments in working conditions (from the board's perspective), and compensation contingent on the firm's carbon intensity to encourage related investments (carbon capture, green technologies, etc.) above the level that would maximize the stock price. This will be the case when the board cares slightly more about carbon emissions than investors. By contrast, a board that cares much more about carbon emissions will be very concerned with providing efficient investment incentives on this dimension, even at the cost of not sufficiently discouraging excessive investment on other dimensions. Thus, it will not use profits-based compensation, even though its negative effect on carbon emissions could in principle be offset by increasing the sensitivity of pay to measures of carbon intensity. The reason is that a high sensitivity of pay to carbon intensity would substantially encourage gaming and inefficient investment decisions on this dimension, which is especially concerning for a board that cares a lot about carbon emissions. In sum, contract complexity, as proxied by the different types of performance measures used, does not necessarily rise when the divergence in preferences increases.

8 Conclusion

Criticisms levelled at ESG-based compensation cannot be answered by existing models of multitasking in which performance is measured ex-post with noise: as is well-known, imperfect measurement does not rule out performance-based compensation. To address these criticisms, we let the manager know the bias in performance measurement at the time of making investment decisions, which allows him to game SEP-based incentives. We also study the outcome with heterogeneous SEP measurement.

We show that ESG-linked compensation will only be used when the level of social investments preferred by the board exceeds the one that maximizes the stock price, i.e., when the board does not maximize the stock price. This contradicts the notion that SEP-based compensation will increase firm value. Moreover, even though it is sometimes second-best optimal, incorporating SEP measures in the contract distorts the manager's incentives for social investments. This distortion reduces the optimal sensitivity of compensation to SEP measures. Finally, we have shown that heterogeneity in social performance measurement can be useful to mitigate gaming.

The model has normative implications for the regulation of SEP measures including ESG scores and ratings. Indeed, it is harder for the manager to game SEP incentives when there are different ESG raters that use a variety of data and methodologies to produce ESG scores and ratings of a similar quality. This suggests that the harmonization of SEP measures may have a counterproductive effect. This should be taken into account at a time when a uniform standard is considered. For example, the objective of the International Sustainability Standards Board (ISSB) is to develop a global standard for sustainability reporting.

Finally, it is worth noting that our optimal contracting perspective did not consider several factors. On the firm side, it did not consider other motivations for providing SEP-based incentives, including public scrutiny and political pressures ("virtue signaling"). On the manager side, it did not consider factors such as opportunistic bargaining, self-dealing, and rent extraction. In poorly governed firms, compensation based on SEP measures could be used to undeservedly inflate managerial compensation (Bebchuk and Tallarita (2022)). Likewise, heterogeneous SEP measurement could allow managers to opportunistically cherry-pick the metrics that shed the best light on their performance. By delineating the circumstances in which SEP measures-based compensation is optimal, our results can help identify instances in which this type of compensation exacerbates rather than ameliorates the agency problem.

References

Baker, G.P., 1992. Incentive contracts and performance measurement. *Journal of Political Economy*, 100, 598-614.

Baker, G., 2002. Distortion and risk in optimal incentive contracts. *Journal of Human Resources*, 37, 728-751.

Baker, M., Egan, M.L., Sarkar, S.K., 2022. How do investors value ESG? Working paper, Harvard Business School.

Barber, B.M., Morse, A., Yasuda, A., 2021. Impact investing. *Journal of Financial Economics*, 139, 162-185.

Baron, D.P., 2008. Managerial contracting and corporate social responsibility. *Journal of Public Economics*, 92, 268-288.

Bauer, R., Ruof, T., Smeets, P., 2021. Get real! Individuals prefer more sustainable investments. *Review of Financial Studies*, 34, 3976-4043.

Bebchuk, L.A., Tallarita, R., 2022. The perils and questionable promise of ESG-based compensation. *Journal of Corporation Law*, 48, 37-75.

Bénabou, R., Tirole, J., 2016. Bonus culture: Competitive pay, screening, and multitasking. *Journal of Political Economy*, 124, 305-370.

Berg, F., Kölbel, J.F., Pavlova, A., Rigobon, R., 2021. ESG confusion and stock returns: Tackling the problem of noise. Working paper, MIT.

Berg, F., Kölbel, J.F., Rigobon, R., 2022. Aggregate confusion: The divergence of ESG ratings. *Review of Finance*, 26, 1315-1344.

Bertomeu, J., Darrough, M., Xue, W., 2017. Optimal conservatism with earnings manipulation. *Contemporary Accounting Research,* 34, 252-284.

Beyer, A., Guttman, I., Marinovic, I., 2014. Optimal contracts with performance manipulation. *Journal of Accounting Research*, 52, 817-847.

Bolton, P., Kacperczyk, M., 2021. Do investors care about carbon risk? *Journal of Financial Economics*, 142, 517-549.

Bond, P., Gomes, A., 2009. Multitask principal–agent problems: Optimal contracts, fragility, and effort misallocation. *Journal of Economic Theory*, 144, 175-211.

Bonham, J., Riggs-Cragun, A., 2021. Contracting on what firm owners value. Working paper, University of Chicago.

Bonham, J., Riggs-Cragun, A., 2022. Motivating ESG activities through contracts, taxes and disclosure regulation. Working paper, University of Chicago.

Broccardo, E., Hart, O. and Zingales, L., 2022. Exit versus voice. *Journal of Political Economy*, 130, 3101-3145

Bucourt, N., Inostroza, N., 2023. ESG investing and managerial incentives. Working paper, University of Toronto.

Budde, J., 2007. Performance measure congruity and the balanced scorecard. *Journal of Accounting Research*, 45, 515-539.

Caskey, J., Laux, V., 2017. Corporate governance, accounting conservatism, and manipulation. *Management Science*, 63, 424-437.

Chatterji, A.K., Durand, R., Levine, D.I., Touboul, S., 2016. Do ratings of firms converge? Implications for managers, investors and strategy researchers. *Strategic Management Journal*, 37, 1597-1614.

Christensen, D.M., Serafeim, G., Sikochi, A., 2022. Why is corporate virtue in the eye of the beholder? The case of ESG ratings. *The Accounting Review*, 97, 147-175.

Cohen, S., Kadach, I., Ormazabal, G., Reichelstein, S., 2023. Executive compensation tied to ESG performance: International evidence. *Journal of Accounting Research*, forthcoming.

Datar, S., Kulp, S.C., Lambert, R.A., 2001. Balancing performance measures. *Journal of Accounting Research*, 39, 75-92.

Dewatripont, M., Jewitt, I., Tirole, J., 2000. Multitask agency problems: Focus and task clustering. *European Economic Review*, 44, 869-877.

Dewatripont, M., Tirole, J., 2022. The morality of markets. Working paper, ECARES.

Dikolli, S.S., Hofmann, C., Kulp, S.L., 2009. Interrelated performance measures, interactive effort, and incentive weights. *Journal of Management Accounting Research*, 21, 125-149.

Dikolli, S.S., Frank, M.M., Guo, Z. and Lynch, L.J., 2023. ESG mutual fund voting on executive compensation shareholder proposals. *Journal of Management Accounting Research*, 35, 51-74.

Duchin, R., Gao, J., Xu, Q., 2022. Sustainability or greenwashing: Evidence from the asset market for industrial pollution. Working paper, Boston College.

Dyck, A., Lins, K.V., Roth, L., Towner, M., Wagner, H.F., 2022. Renewable governance: Good for the environment? *Journal of Accounting Research*, forthcoming.

Ederer, F., Holden, R., Meyer, M., 2018. Gaming and strategic opacity in incentive provision. *RAND Journal of Economics*, 49, 819-854.

Edmans, A., 2021, Why companies shouldn't tie CEO pay to ESG metrics. *Wall Street Journal*, June 27.

Edmans, A., 2023. The end of ESG. *Financial Management*, 52, 3-17.

Edmans, A., Gabaix, X., 2011. Tractability in incentive contracting. *Review of Financial Studies*, 24, 2865-2894.

Edmans, A., Gabaix, X., Jenter, D., 2017. Executive compensation: A survey of theory and evidence. *The Handbook of the Economics of Corporate Governance*, 1, 383-539.

Falck, O., Heblich, S., 2007. Corporate social responsibility: Doing well by doing good. *Business horizons*, 50, 247-254.

Feltham, G.A., Xie, J., 1994. Performance measure congruity and diversity in multi-task principal/agent relations. *The Accounting Review*, 69, 429-453.

Fiechter, P., Hitz, J.M., Lehmann, N., 2022. Real effects of a widespread CSR reporting mandate: Evidence from the European Union's CSR Directive. *Journal of Accounting Research*, 60, 1499-1549.

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

Flammer, C., Kacperczyk, A., 2016. The impact of stakeholder orientation on innovation: Evidence from a natural experiment. *Management Science*, 62, 1982-2001.

Friedman, H.L., Heinle, M.S., Luneva, I., 2022. A theoretical framework for ESG reporting to investors. Working paper, UCLA.

Gaynor, M., Mehta, N., Richards-Shubik, S., 2023. Optimal contracting with altruistic agents: Medicare payments for dialysis drugs. *American Economic Review*, 113, 1530-1571.

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.

Geelen, T., Hajda, J. and Starmans, J., 2023. Sustainable organizations. Working paper, Pennsylvania State University.

Goldstein, I., Kopytov, A., Shen, L., Xiang, H., 2022. On ESG investing: Heterogeneous preferences, information, and asset prices. Working paper, University of Pennsylvania.

Gryglewicz, S., Mayer, S., Morellec, E., 2020. Agency conflicts and short-versus long-termism in corporate policies. *Journal of Financial Economics*, 136, 718-742.

Hart, O.D., Zingales, L., 2017. Companies Should Maximize Shareholder Welfare Not Market Value. *Journal of Law, Finance, and Accounting*, 2, 247-274.

Hartzmark, S.M., Sussman, A.B., 2019. Do investors value sustainability? A natural experiment examining ranking and fund flows. *Journal of Finance*, 74, 2789-2837.

Heeb, F., Kölbel, J.F., Paetzold, F., Zeisberger, S., 2023. Do investors care about impact? *Review of Financial Studies*, 36, 1737-1787.

Hoffmann, F., Pfeil, S., 2021. Dynamic multitasking and managerial investment incentives. *Journal of Financial Economics*, 142, 954-974.

Holmström, B., 1979. Moral hazard and observability. *Bell Journal of Economics*, 10, 74-91.

Holmström, B., Milgrom, P., 1991. Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. *Journal of Law, Economics & Organization*, 7, 24-52.

Holmström, B., Tirole, J., 1993. Market liquidity and performance monitoring. *Journal of Political Economy*, 101, 678-709.

Homroy, S., Mavruk, T., Nguyen, D., 2023. ESG-Linked Compensation, CEO Skills, and Shareholders' Welfare. *Review of Corporate Finance Studies*, forthcoming.

Hörner, J., Lambert, N.S., 2021. Motivational ratings. *Review of Economic Studies*, 88, 1892-1935.

Humphrey, J., Kogan, S., Sagi, J.S., Starks, L.T., 2023, The asymmetry in responsible investing preferences. NBER Working Paper 29288.

Ikram, A., Li, Z.F., Minor, D., 2019. CSR-contingent executive compensation contracts. *Journal of Banking & Finance*, 105655.

Innes, R.D., 1990. Limited liability and incentive contracting with ex-ante action choices. *Journal of Economic Theory*, 52, 45-67.

Jensen, M.C., Murphy, K.J., 1990. Performance pay and top-management incentives. *Journal of Political Economy*, 98, 225-264.

Kerr, S., 1975. On the folly of rewarding A, while hoping for B. *Academy of Management Journal*, 18, 769-783.

Kvaløy, O., Olsen, T.E., 2022. Balanced scorecards: A relational contract approach. *Journal of Accounting Research*, forthcoming.

Larcker, D.F., Pomorski, L., Tayan, B., Watts, E.M., 2022. ESG ratings: A compass without direction. Working paper, Stanford University.

Laux, C., 2001. Limited-liability and incentive contracting with multiple projects. *RAND Journal of Economics*, 32, 514-526.

Mischke, J., Woetzel, J., Birshan, M., 2021. The necessity of doing well by doing good. *Milken Institute Review*, April.

Oehmke, M., Opp, M.M., 2023. A theory of socially responsible investment. *Review of Economic Studies*, forthcoming.

Pagano, M.S., Sinclair, G., Yang, T., 2018. Understanding ESG ratings and ESG indexes. *Research Handbook of Finance and Sustainability.* Edward Elgar Publishing.

Parlour, C.A., Rajan, U., 2020. Contracting on credit ratings: Adding value to public information. *Review of Financial Studies*, 33, 1412-1444.

Pástor, L., Stambaugh, R.F., Taylor, L.A., 2021. Sustainable investing in equilibrium. *Journal of Financial Economics*, 142, 550-571.

Paul, J.M., 1992. On the efficiency of stock-based compensation. *Review of Financial Studies*, 5, 471-502.

Pawliczek, A., Carter, M.E., Zhong, R.I., 2023. Say on ESG: The adoption of say-on-pay laws, ESG contracting, and firm ESG performance. Working paper, University of Colorado Boulder.

Prendergast, C., 2002. The tenuous trade-off between risk and incentives. *Journal of Political Economy*, 110, 1071-1102.

Rajan, R., Ramella, P., Zingales, L., 2023. What purpose do corporations purport? Evidence from letters to shareholders. Working paper, University of Chicago.

Riedl, A., Smeets, P., 2017. Why do investors hold socially responsible mutual funds? *Journal of Finance*, 72, 2505-2550.

Shleifer, A., Vishny, R.W., 1997. A survey of corporate governance. *Journal of Finance*, 52, 737-783.

Smith, M.J., 2002. Gaming nonfinancial performance measures. *Journal of Management Accounting Research*, 14, 119-133.

Szydlowski, M., 2019. Incentives, project choice, and dynamic multitasking. *Theoretical Economics*, 14, 813-847.

Tirole, J., 2001. Corporate Governance. *Econometrica*, 69, 1-35.

Tirole, J., 2006. *The Theory of Corporate Finance.* Princeton university press.

Tonello, M., 2024, ESG Performance Metrics in Executive Pay, *Harvard Law School Forum on Corporate Governance*, January 15.

Walker, D.I., 2022. The economic (in)significance of executive pay ESG incentives. *Stanford Journal of Law, Business & Finance*, 27, 317-350

Appendix

A Stock price

This section analyzes the portfolio optimization problem of socially conscious investors with "warm-glow" preferences. This allows to microfound the stock price function postulated in equation (4).

Assume that investors can invest their wealth either at the riskfree rate, which is zero, or in the firm's stock. They are risk neutral, and they have preferences over the firm's cash flows and its social output: they assign a weight of $\alpha_I \geq 0$ to the firm's social output relative to its cash flows (see section 7.4).[34] They trade at $t = 1$ after observing the financial report z, social investment y, and ESG score m. The firm's stock price is set by market clearing.

All investors have the same information and therefore do not learn from the stock price. They are price takers. A risk neutral investor chooses the quantity q of stock to buy conditional on his information available at $t = 1$ to maximize:

$$\mathbb{E}\left[(\tilde{x} + \alpha_I \tilde{v} - p)\, q | y, z, m\right] = q\left(\mathbb{E}\left[\tilde{x}|y, z, m\right] + \alpha_I \mathbb{E}\left[\tilde{v}|y, z, m\right]\right) - pq \tag{13}$$

Maximizing the expression in equation (13) with respect to q shows that the only stock price compatible with market clearing for the firm stock is:

$$p = \mathbb{E}\left[\tilde{x}|y, z, m\right] + \alpha_I \mathbb{E}\left[\tilde{v}|y, z, m\right] \tag{14}$$

This is as in equation (4).

B Proofs

Proof of Lemma 1:

The expected effort of the manager given the contract offered is denoted by $\hat{e}$. The stock price depends on belief updating at $t = 1$, after the observation of the report z, investment y, and the ESG score m. Since $m = \gamma y$ and y is also observable, m reveals γ. Using the standard formula for the conditional

[34] This is the simplest specification of investors' social preferences. If investors have heterogeneous social preferences, the parameter α_I is the social weight of the marginal investor. The marginal investor, who is indifferent between buying the stock or not, is such that the stock market clears (i.e. total investor demand equals supply), see Bucourt and Inostroza (2023). In this type of model, there must be constraints on portfolio choice for the stock market to clear, for example a short-selling constraint and a borrowing constraint.

expectation with normal distributions:

$$\mathbb{E}[\tilde{x}|y,z,m] = \mathbb{E}[\tilde{x}] + \frac{cov(\tilde{z},\tilde{x})}{var(\tilde{z})}\left(z - \mathbb{E}[\tilde{z}]\right) = \hat{e} - \theta y^2 + \frac{\sigma_x^2}{\sigma_x^2 + \sigma_z^2}\left(z - \hat{e} + \theta y^2\right) \tag{15}$$

$$\mathbb{E}[\tilde{v}|y,z,m] = \mathbb{E}\left[\tilde{\eta}\tilde{y} + \tilde{\epsilon}_y|y,z,m\right] = y\mathbb{E}\left[\tilde{\eta}|m,y\right] \tag{16}$$

where:

$$\mathbb{E}\left[\tilde{\eta}|m,y\right] = \mathbb{E}\left[\tilde{\eta}|\gamma\right] = \mathbb{E}\left[\tilde{\eta}\right] + \frac{cov\left(\tilde{\eta},\tilde{\gamma}\right)}{var\left(\tilde{\gamma}\right)}\left(\gamma - \mathbb{E}[\tilde{\gamma}]\right) = \bar{\eta} + \frac{\sigma_\eta^2}{\sigma_\eta^2 + \sigma_\gamma^2}\left(\gamma - \bar{\eta}\right) \tag{17}$$

Substituting into equation (16):

$$\mathbb{E}[\tilde{v}|y,z,m] = \left(\bar{\eta} + \frac{\sigma_\eta^2}{\sigma_\eta^2 + \sigma_\gamma^2}\left(\gamma - \bar{\eta}\right)\right)y = \frac{\sigma_\eta^2}{\sigma_\eta^2 + \sigma_\gamma^2}m + \frac{\sigma_\gamma^2}{\sigma_\eta^2 + \sigma_\gamma^2}y\bar{\eta} \tag{18}$$

■

Proof of Lemma 2:

The principal's optimization problem at the first-best (without incentive compatibility constraints) is:

$$\max_{e,w,y,\beta_x,\beta_p,\beta_m} \mathbb{E}[V(y,e)] \quad \text{s.t.} \quad \mathbb{E}[u(y,e)] \geq \overline{U},\ \beta_x \geq 0, \beta_p \geq 0 \tag{19}$$

where:

$$\mathbb{E}[V(y,e)] = \mathbb{E}\left[\tilde{x} + \alpha_B\tilde{v} - (w + \beta_x\tilde{x} + \beta_p\tilde{p} + \beta_m\tilde{m})\right] \tag{20}$$

$$\mathbb{E}[u(y,e)] = \mathbb{E}\left[w + \beta_x\tilde{x} + \beta_p\tilde{p} + \beta_m\tilde{m}\right] - C(e) \tag{21}$$

The principal's objective function is concave in y, and the first-best optimal value of y is:

$$y^* = \frac{\alpha_B}{2}\frac{\bar{\eta}}{\theta} \tag{22}$$

The manager must also receive an expected payment equal to $\overline{U} + c_e$.

Now consider the second-best outcome with an incentive problem. The optimization program is as above but it also includes the constraints that effort and investment are optimally chosen by the manager: $\{e^\star, y^\star\} = \arg\max_{e,y}\mathbb{E}[u(y,e)]$. The manager's objective function given his contract $\{w, \beta_x, \beta_p, \beta_m\}$ is:

$$\mathbb{E}\left[w + \beta_x\left(e - \theta y^2 + \tilde{\epsilon}_x\right) + \beta_p\tilde{p} + \beta_m m\right] - C(e) \tag{23}$$

For $\beta_x + \beta_p > 0$, the manager's objective function is concave in y. The first-order condition (FOC) with respect to y is:

$$\beta_x \left(-2\theta y\right) + \beta_p \left(-2\theta y + \alpha_I \bar{\eta}\right) + \beta_m \bar{\eta} = 0 \quad \Leftrightarrow \quad y = \frac{\beta_m + \beta_p \alpha_I}{\beta_x + \beta_p} \frac{\bar{\eta}}{2\theta} \tag{24}$$

Thus, the first-best investment decision can be elicited at the second-best simply by setting:

$$\frac{\beta_m + \beta_p \alpha_I}{\beta_x + \beta_p} = \alpha_B. \tag{25}$$

The manager will optimally exert high effort ($e = \bar{e}$) if and only if:

$$\begin{aligned}
& \mathbb{E}\left[u(y,e)|e = \bar{e}\right] - c_e \geq \mathbb{E}\left[u(y,e)|e = \underline{e}\right] \\
\Leftrightarrow \quad & w + \beta_x \mathbb{E}\left[e - \theta y^2 + \tilde{\epsilon}_x | e = \bar{e}\right] + \beta_p \mathbb{E}\left[\tilde{x} + \alpha_I \tilde{v} \middle| e = \bar{e}\right] + \beta_m \mathbb{E}\left[y\tilde{\gamma}\right] - c_e \\
& \geq w + \beta_x \mathbb{E}\left[e - \theta y^2 + \tilde{\epsilon}_x | e = \underline{e}\right] + \beta_p \mathbb{E}\left[\tilde{x} + \alpha_I \tilde{v} \middle| e = \underline{e}\right] + \beta_m \mathbb{E}\left[y\tilde{\gamma}\right] \\
\Leftrightarrow \quad & \beta_x + \beta_p \frac{\sigma_x^2}{\sigma_x^2 + \sigma_z^2} \geq \frac{c_e}{\bar{e} - \underline{e}}.
\end{aligned} \tag{26}$$

Finally, given the values of $\{\beta_x, \beta_p, \beta_m\}$ above, the fixed wage w adjusts so that the manager's expected payment is equal to $\overline{U} + c_e$. This establishes that the first-best outcome can be achieved at the second-best. ■

Proof of Lemma 3:

The contract is characterized by a fixed payment w to the manager and by the sensitivity of pay to performance with respect to $\tilde{x}$, $\tilde{p}$, and $\tilde{m}$, which is given by $\{\beta_x, \beta_p, \beta_m\}$, respectively. The principal's optimization problem is:

$$\max_{e,w,\beta_x,\beta_p,\beta_m} \mathbb{E}[V(y,e)] \quad \text{s.t.} \quad \{e^\star, y^\star\} = \arg\max_{e,y} \mathbb{E}[u(y,e)],\ \mathbb{E}[u(x,y,e)] \geq \overline{U},\ \beta_x \geq 0, \beta_p \geq 0 \tag{27}$$

where $\mathbb{E}[V(y,e)]$ and $\mathbb{E}[u(y,e)]$ are defined in equations (20) and (21). The manager's objective function given his contract $\{w, \beta_x, \beta_p, \beta_m\}$ is:

$$\mathbb{E}\left[w + \beta_x\left(e - \theta y^2 + \tilde{\epsilon}_x\right) + \beta_p \tilde{p} + \beta_m m | \gamma\right] - C(e) \tag{28}$$

where the stock price $\tilde{p}$ is as in Lemma 1. For $\beta_x + \beta_p > 0$, the manager's objective function is concave in y. The FOC with respect to y is:

$$
\begin{aligned}
& & \beta_x(-2\theta y)+\beta_p\left(-2\theta y+\alpha_I\left(\frac{\sigma_\eta^2}{\sigma_\eta^2+\sigma_\gamma^2}\gamma+\frac{\sigma_\gamma^2}{\sigma_\eta^2+\sigma_\gamma^2}\bar{\eta}\right)\right)+\beta_m\gamma=0 \\
\Leftrightarrow \quad y(\gamma) &=& \frac{\beta_m\gamma+\beta_p\alpha_I\left(\frac{\sigma_\eta^2}{\sigma_\eta^2+\sigma_\gamma^2}\gamma+\frac{\sigma_\gamma^2}{\sigma_\eta^2+\sigma_\gamma^2}\bar{\eta}\right)}{\beta_x+\beta_p}\frac{1}{2\theta} \\
&=& \frac{\left(\beta_m+\beta_p\alpha_I\frac{\sigma_\eta^2}{\sigma_\eta^2+\sigma_\gamma^2}\right)\gamma+\beta_p\alpha_I\frac{\sigma_\gamma^2}{\sigma_\eta^2+\sigma_\gamma^2}\bar{\eta}}{\beta_x+\beta_p}\frac{1}{2\theta} \qquad (29)
\end{aligned}
$$

Given his contract, the manager will optimally exert high effort ($e=\bar{e}$) if and only if:

$$
\begin{aligned}
& \mathbb{E}\left[u(y,e)|e=\bar{e}\right]-c_e\geq\mathbb{E}\left[u(y,e)|e=\underline{e}\right] \\
\Leftrightarrow \quad & w+\beta_x\mathbb{E}\left[e-\theta y^2+\tilde{\epsilon}_x|e=\bar{e}\right]+\beta_p\mathbb{E}\left[\mathbb{E}\left[\tilde{x}|y,z,m\right]+\alpha_I\mathbb{E}\left[\tilde{v}|y,z,m\right]\Big|e=\bar{e}\right]+\beta_m y\gamma-c_e \\
& \geq w+\beta_x\mathbb{E}\left[e-\theta y^2+\tilde{\epsilon}_x|e=\underline{e}\right]+\beta_p\mathbb{E}\left[\mathbb{E}\left[\tilde{x}|y,z,m\right]+\alpha_I\mathbb{E}\left[\tilde{v}|y,z,m\right]\Big|e=\underline{e}\right]+\beta_m y\gamma \\
\Leftrightarrow \quad & \beta_x+\beta_p\frac{\sigma_x^2}{\sigma_x^2+\sigma_z^2}\geq\frac{c_e}{\bar{e}-\underline{e}} \qquad (30)
\end{aligned}
$$

The fixed component of pay, w, is set to guarantee the manager's participation (before the manager observes γ) for $e=\bar{e}$:

$$
\begin{aligned}
\mathbb{E}[u(y,\bar{e})] &=& \bar{W} \\
\Leftrightarrow \quad w+\beta_x\mathbb{E}\left[\bar{e}-\theta y^2+\tilde{\epsilon}_x\right]+\beta_p\mathbb{E}[\tilde{p}]+\beta_m\mathbb{E}[\tilde{m}] &=& \bar{W}+C(\bar{e}) \qquad (31)
\end{aligned}
$$

For a given γ, the principal's objective function in equation (20) can be rewritten as:

$$
\begin{aligned}
\mathbb{E}[V(y,e)] &=& \mathbb{E}\left[(1-\beta_x)\left(\bar{e}-\theta y^2+\tilde{\epsilon}_x\right)+\alpha_B\left(\tilde{\eta}y+\tilde{\epsilon}_y\right)-\left(w+\beta_p\tilde{p}+\beta_m\tilde{m}\right)|\gamma\right] \\
&=& \mathbb{E}\left[\bar{e}-\theta y^2+\tilde{\epsilon}_x+\alpha_B\left(\tilde{\eta}y+\tilde{\epsilon}_y\right)-\left(\bar{W}+C(\bar{e})\right)|\gamma\right] \\
&=& \bar{e}-\theta y^2+\alpha_B\mathbb{E}\left[\tilde{\eta}|\gamma\right]y-\left(\bar{W}+C(\bar{e})\right) \qquad (32)
\end{aligned}
$$

We derive the "first-best" outcome without an agency problem (i.e. the incentive constraint in equation (30) can be ignored and information is symmetric). The board chooses social investment y to maximize its objective function in equation (32), where

$$
\mathbb{E}\left[\tilde{\eta}|\gamma\right]=\mathbb{E}\left[\tilde{\eta}\right]+\frac{cov(\tilde{\eta},\tilde{\gamma})}{var(\tilde{\gamma})}(\gamma-\mathbb{E}[\tilde{\gamma}])=\bar{\eta}+\frac{\sigma_\eta^2}{\sigma_\eta^2+\sigma_\gamma^2}(\gamma-\bar{\eta})=\frac{\sigma_\eta^2}{\sigma_\eta^2+\sigma_\gamma^2}\gamma+\frac{\sigma_\gamma^2}{\sigma_\eta^2+\sigma_\gamma^2}\bar{\eta} \qquad (33)
$$

The objective function is concave in y, so that the first-best optimum is given by the FOC:

$$y^*(\gamma) = \frac{\alpha_B}{2}\frac{1}{\theta}\left(\frac{\sigma_\eta^2}{\sigma_\eta^2+\sigma_\gamma^2}\gamma + \frac{\sigma_\gamma^2}{\sigma_\eta^2+\sigma_\gamma^2}\bar{\eta}\right) \tag{34}$$

We now consider the second-best outcome with an agency problem. The objective function of the board is still given by equation (32). For given γ and y, define:

$$f(\gamma, y) \equiv -\theta y^2 + \alpha_B \mathbb{E}[\tilde{\eta}|\gamma] y \tag{35}$$

The first and second derivatives with respect to y are respectively:

$$f_y(\gamma, y) = -2\theta y + \alpha_B \mathbb{E}[\tilde{\eta}|\gamma] \tag{36}$$

$$f_{yy}(\gamma, y) = -2\theta \tag{37}$$

For a given e, maximizing the objective function of the board is equivalent to maximizing:

$$\mathbb{E}\left[f(\gamma, y)\right] = \int_\eta \int_\gamma f(\gamma, y)\varphi(\gamma|\eta)\phi(\eta) d\gamma d\eta \tag{38}$$

Thus, for a given effort e and a given γ, the value of y that maximizes the board's objective function is the value of y that maximizes the expression in equation (38). By definition of $y^*(\gamma)$, for a given γ, the function $f(\gamma, y)$ is maximized by setting $y(\gamma) = y^*(\gamma)$. The value of y that maximizes the board's objective function is the value of y that maximizes:

$$\max_y \mathbb{E}\left[(f(\gamma, y) - f(\gamma, y^*(\gamma)))\right] = \mathbb{E}\left[(f(\gamma, y) - f(\gamma, y^*(\gamma)))\right] \tag{39}$$

The function $f(\gamma, y)$ is quadratic in y. Therefore, for a given γ, a second-order Taylor expansion around $y^*(\gamma)$ is exact.

$$\begin{aligned} f(\gamma, y) &= f(\gamma, y^*(\gamma)) + f_y(\gamma, y^*(\gamma))(y - y^*(\gamma)) + \frac{1}{2} f_{yy}(\gamma, y^*(\gamma))(y - y^*(\gamma))^2 \\ &= f(\gamma, y^*(\gamma)) + \left(-2\theta y^*(\gamma) + \alpha_B \mathbb{E}[\tilde{\eta}|\gamma]\right)(y - y^*(\gamma)) - \theta(y - y^*(\gamma))^2 \end{aligned}$$

Thus:

$$
\begin{aligned}
f(\gamma,y)-f(\gamma,y^*(\gamma)) &= \left(-2\theta y^*(\gamma)+\alpha_B\left(\frac{\sigma_\eta^2}{\sigma_\eta^2+\sigma_\gamma^2}\gamma+\frac{\sigma_\gamma^2}{\sigma_\eta^2+\sigma_\gamma^2}\bar{\eta}\right)\right)(y(\gamma)-y^*(\gamma))-\theta(y(\gamma)-y^*(\gamma))^2 \\
&= 2\theta\left(\left(-y^*(\gamma)+\frac{\alpha_B}{2}\frac{1}{\theta}\left(\frac{\sigma_\eta^2}{\sigma_\eta^2+\sigma_\gamma^2}\gamma+\frac{\sigma_\gamma^2}{\sigma_\eta^2+\sigma_\gamma^2}\bar{\eta}\right)\right)(y(\gamma)-y^*(\gamma))-\frac{1}{2}(y(\gamma)-y^*(\gamma))^2\right) \\
&= 2\theta\left(-\frac{1}{2}(y(\gamma)-y^*(\gamma))^2\right) \\
&= -\theta\left(y(\gamma)-y^*(\gamma)\right)^2 \qquad (40)
\end{aligned}
$$

where we used equations (33) and (34) to get the first and third equalities, respectively. ■

Proof of Lemma 4:

Plugging the investment $y(\gamma)$ optimally chosen by the manager and the first-best investment $y^*(\gamma)$ from equations (29) and (34), and using Lemma 3, the board's optimization problem for a given effort is:

$$
\begin{aligned}
&\min_{\beta_x,\beta_p,\beta_m}\theta\int_\eta\int_\gamma\left(\frac{1}{2}\frac{1}{\theta}\frac{\left(\beta_m+\beta_p\alpha_I\frac{\sigma_\eta^2}{\sigma_\eta^2+\sigma_\gamma^2}\right)\gamma+\beta_p\alpha_I\frac{\sigma_\gamma^2}{\sigma_\eta^2+\sigma_\gamma^2}\bar{\eta}}{\beta_x+\beta_p}-\frac{\alpha_B}{2}\frac{1}{\theta}\left(\frac{\sigma_\eta^2}{\sigma_\eta^2+\sigma_\gamma^2}\gamma+\frac{\sigma_\gamma^2}{\sigma_\eta^2+\sigma_\gamma^2}\bar{\eta}\right)\right)^2\varphi(\gamma|\eta)\phi(\eta)d\gamma d\eta \\
\Leftrightarrow &\min_{\beta_x,\beta_p,\beta_m}\frac{1}{\theta}\int_\eta\int_\gamma\left(\left(\frac{\beta_m}{\beta_x+\beta_p}+\left(\frac{\beta_p\alpha_I}{\beta_x+\beta_p}-\alpha_B\right)\frac{\sigma_\eta^2}{\sigma_\eta^2+\sigma_\gamma^2}\right)\gamma+\left(\frac{\beta_p\alpha_I}{\beta_x+\beta_p}-\alpha_B\right)\frac{\sigma_\gamma^2}{\sigma_\eta^2+\sigma_\gamma^2}\bar{\eta}\right)^2\varphi(\gamma|\eta)\phi(\eta)d\gamma d\eta \qquad (41)
\end{aligned}
$$

Without nonnegativity constraints on contract parameters, simply set $\beta_m=0$ and $\frac{\beta_p\alpha_I}{\beta_x+\beta_p}=\alpha_B \Leftrightarrow \beta_x=\frac{\beta_p\alpha_I}{\alpha_B}-\beta_p$, so that the expression under the integral sign in equation (41) is zero for any γ. Since this expression (a quadratic function) is nonnegative for any γ, achieving a value of zero for this expression at every γ maximizes the objective function of the principal for a given effort, and it implies that $y(\gamma)=y^*(\gamma)\ \forall\gamma$ (see equation (40)). To elicit high effort, use equation (30) to set:

$$
\frac{\beta_p\alpha_I}{\alpha_B}-\beta_p+\beta_p\frac{\sigma_x^2}{\sigma_x^2+\sigma_z^2}=\frac{c_e}{\bar{e}-\underline{e}} \quad\Leftrightarrow\quad \beta_p=\frac{c_e}{\bar{e}-\underline{e}}\left(\frac{\alpha_I}{\alpha_B}+\frac{\sigma_x^2}{\sigma_x^2+\sigma_z^2}-1\right)^{-1} \qquad (42)
$$

Proof of Proposition 1:

We still rely on equation (41). Conditional on η, the random variable $\tilde{\gamma}$, with PDF $\varphi(\gamma|\eta)$, is normally distributed with mean η and variance σ_γ^2. Thus, $\int_\gamma \gamma\varphi(\gamma|\eta)d\gamma=\eta$. This implies:

$$
\int_\eta\int_\gamma \gamma\varphi(\gamma|\eta)\phi(\eta)d\gamma d\eta=\int_\eta\left(\int_\gamma\gamma\varphi(\gamma|\eta)d\gamma\right)\phi(\eta)d\eta=\int_\eta\eta\phi(\eta)d\eta=\bar{\eta} \qquad (43)
$$

Moreover, $\int_\gamma \gamma^2 \varphi(\gamma|\eta)d\gamma = \mathbb{E}[\tilde{\gamma}^2|\eta] = var(\tilde{\gamma}|\eta) + (\mathbb{E}[\tilde{\gamma}|\eta])^2 = \sigma_\gamma^2 + \eta^2$. This implies:

$$\int_\eta \int_\gamma \gamma^2 \varphi(\gamma|\eta)\phi(\eta)d\gamma d\eta = \int_\eta \left(\sigma_\gamma^2 + \eta^2\right)\phi(\eta)d\eta = \sigma_\gamma^2 + \int_\eta \eta^2 \phi(\eta)d\eta = \sigma_\gamma^2 + \sigma_\eta^2 + \bar{\eta}^2 \tag{44}$$

The expression in equation (41) is globally convex with respect to β_m. The FOC w.r.t. β_m is:

$$\begin{aligned}
&\frac{1}{\theta}\int_\eta \int_\gamma \frac{2}{\beta_x+\beta_p}\gamma\Bigg(\left(\frac{\beta_m}{\beta_x+\beta_p} + \left(\frac{\beta_p\alpha_I}{\beta_x+\beta_p} - \alpha_B\right)\frac{\sigma_\eta^2}{\sigma_\eta^2+\sigma_\gamma^2}\right)\gamma \\
&+ \left(\frac{\beta_p\alpha_I}{\beta_x+\beta_p} - \alpha_B\right)\frac{\sigma_\gamma^2}{\sigma_\eta^2+\sigma_\gamma^2}\bar{\eta}\Bigg)\varphi(\gamma|\eta)\phi(\eta)d\gamma d\eta = 0 \\
\Leftrightarrow\quad & \left(\frac{\beta_m}{\beta_x+\beta_p} + \left(\frac{\beta_p\alpha_I}{\beta_x+\beta_p} - \alpha_B\right)\frac{\sigma_\eta^2}{\sigma_\eta^2+\sigma_\gamma^2}\right)\left(\sigma_\gamma^2+\sigma_\eta^2+\bar{\eta}^2\right) = \left(\alpha_B - \frac{\beta_p\alpha_I}{\beta_x+\beta_p}\right)\frac{\sigma_\gamma^2}{\sigma_\eta^2+\sigma_\gamma^2}\bar{\eta}^2 = 0 \\
\Leftrightarrow\quad & \beta_m = (\alpha_B(\beta_x+\beta_p) - \beta_p\alpha_I)\left(\frac{\sigma_\gamma^2}{\sigma_\eta^2+\sigma_\gamma^2}\frac{\bar{\eta}^2}{\sigma_\gamma^2+\sigma_\eta^2+\bar{\eta}^2} + \frac{\sigma_\eta^2}{\sigma_\eta^2+\sigma_\gamma^2}\right),
\end{aligned} \tag{45}$$

where we used equations (43) and (44).

We establish an intermediary result: when the nonnegativity constraint on β_m does not bind, substituting for $\frac{\beta_m}{\beta_x+\beta_p}$ in equation (41):

$$\begin{aligned}
&\frac{1}{\theta}\int_\eta \int_\gamma \Bigg(\Bigg(\left(\alpha_B - \frac{\beta_p\alpha_I}{\beta_x+\beta_p}\right)\left(\frac{\sigma_\gamma^2}{\sigma_\eta^2+\sigma_\gamma^2}\frac{\bar{\eta}^2}{\sigma_\gamma^2+\sigma_\eta^2+\bar{\eta}^2} + \frac{\sigma_\eta^2}{\sigma_\eta^2+\sigma_\gamma^2}\right) \\
&+ \left(\frac{\beta_p\alpha_I}{\beta_x+\beta_p} - \alpha_B\right)\frac{\sigma_\eta^2}{\sigma_\eta^2+\sigma_\gamma^2}\Bigg)\gamma + \left(\frac{\beta_p\alpha_I}{\beta_x+\beta_p} - \alpha_B\right)\frac{\sigma_\gamma^2}{\sigma_\eta^2+\sigma_\gamma^2}\bar{\eta}\Bigg)^2 \varphi(\gamma|\eta)\phi(\eta)d\gamma d\eta \\
= \quad & \left(\alpha_B - \frac{\beta_p\alpha_I}{\beta_x+\beta_p}\right)^2 \frac{\sigma_\gamma^4}{(\sigma_\eta^2+\sigma_\gamma^2)^2}\theta\int_\eta \int_\gamma \left(\frac{\bar{\eta}^2}{\sigma_\gamma^2+\sigma_\eta^2+\bar{\eta}^2}\gamma - \bar{\eta}\right)^2 \varphi(\gamma|\eta)\phi(\eta)d\gamma d\eta
\end{aligned} \tag{46}$$

Start with the case $\alpha_B \leq \alpha_I$. Supposing that the nonnegativity constraint on β_m does not bind, the expression under the integral sign in equation (46) is positive and independent from the contract. It is minimized by minimizing $\left(\alpha_B - \frac{\beta_p\alpha_I}{\beta_x+\beta_p}\right)^2$. With a nonnegativity constraint on β_x, this is achieved by setting $\beta_x = \max\{\frac{\beta_p\alpha_I}{\alpha_B} - \beta_p, 0\}$. Substituting for β_x shows that the expression in equation (46) is zero as long as the nonnegativity constraint does not bind for β_x, i.e. $\frac{\beta_p\alpha_I}{\alpha_B} \geq \beta_p$, which with $\alpha_B \leq \alpha_I$ is true. This implies that $y(\gamma) = y^*(\gamma)\ \forall\gamma$ in this case (see equation (40)).

Now supposing instead that the nonnegativity constraint on β_m binds, i.e. $\beta_m = 0$, substituting for β_m in equation (41):

$$\frac{1}{\theta}\int_\eta\int_\gamma\left(\left(\frac{\beta_p\alpha_I}{\beta_x+\beta_p}-\alpha_B\right)\frac{\sigma_\eta^2}{\sigma_\eta^2+\sigma_\gamma^2}\gamma+\left(\frac{\beta_p\alpha_I}{\beta_x+\beta_p}-\alpha_B\right)\frac{\sigma_\gamma^2}{\sigma_\eta^2+\sigma_\gamma^2}\bar{\eta}\right)^2\varphi(\gamma|\eta)\phi(\eta)d\gamma d\eta \tag{47}$$

Also substituting for β_x shows that the expression under the integral sign of equation (47) is zero for any γ as long as the nonnegativity constraint does not bind for β_x, i.e. $\frac{\beta_p\alpha_I}{\alpha_B}\geq\beta_p$, which with $\alpha_B\leq\alpha_I$ is always true. Since the expression in equation (47) (a quadratic function) is nonnegative for any γ, achieving a value of zero for this expression at every γ maximizes the objective function of the principal for a given effort, and it implies that $y(\gamma)=y^*(\gamma)\ \forall\gamma$ (see equation (40)).

Using these results and equation (9), expected social output when $\alpha_B\leq\alpha_I$ is:

$$\begin{aligned}\mathbb{E}[\tilde{\eta}\tilde{y}] &= \mathbb{E}\left[\frac{\alpha_B}{2}\frac{1}{\theta}\left(\frac{\sigma_\eta^2}{\sigma_\eta^2+\sigma_\gamma^2}\tilde{\gamma}+\frac{\sigma_\gamma^2}{\sigma_\eta^2+\sigma_\gamma^2}\bar{\eta}\right)\tilde{\eta}\right]\\ &= \frac{\alpha_B}{2}\frac{1}{\theta}\left(\frac{\sigma_\eta^2}{\sigma_\eta^2+\sigma_\gamma^2}\left(\bar{\eta}^2+\sigma_\eta^2\right)+\frac{\sigma_\gamma^2}{\sigma_\eta^2+\sigma_\gamma^2}\bar{\eta}^2\right)\\ &= \frac{\alpha_B}{2\theta}\left(\frac{\sigma_\eta^4}{\sigma_\eta^2+\sigma_\gamma^2}+\bar{\eta}^2\right)\end{aligned} \tag{48}$$

Now consider the case $\alpha_B>\alpha_I$. Then the nonnegativity constraint on β_m does not bind (see equation (45) with $\beta_x\geq 0$), and the principal minimizes equation (46). This is equivalent to minimizing $\alpha_B-\frac{\beta_p\alpha_I}{\beta_x+\beta_p}$ subject to $\beta_x\geq 0$, which yields $\beta_x=0$. To induce effort, β_p is set as in equation (30) with $\beta_x=0$. Finally, β_m is given in equation (45) with $\beta_x=0$ and the value of β_p derived in the previous step. From equation (29):

$$\begin{aligned}y(\gamma) &= \frac{1}{2\theta}\left(\left(\frac{\beta_m}{\beta_p}+\alpha_I\frac{\sigma_\eta^2}{\sigma_\eta^2+\sigma_\gamma^2}\right)\gamma+\alpha_I\frac{\sigma_\gamma^2}{\sigma_\eta^2+\sigma_\gamma^2}\bar{\eta}\right)\\ &= \frac{1}{2\theta}\left(\left((\alpha_B-\alpha_I)\left(\frac{\sigma_\gamma^2}{\sigma_\eta^2+\sigma_\gamma^2}\frac{\bar{\eta}^2}{\sigma_\gamma^2+\sigma_\eta^2+\bar{\eta}^2}+\frac{\sigma_\eta^2}{\sigma_\eta^2+\sigma_\gamma^2}\right)+\alpha_I\frac{\sigma_\eta^2}{\sigma_\eta^2+\sigma_\gamma^2}\right)\gamma+\alpha_I\frac{\sigma_\gamma^2}{\sigma_\eta^2+\sigma_\gamma^2}\bar{\eta}\right)\end{aligned}$$

So that, also using equation (9):

$$\begin{aligned}y(\gamma)-y^*(\gamma) &= \frac{1}{2\theta}\underbrace{(\alpha_B-\alpha_I)}_{>0}\frac{\sigma_\gamma^2}{\sigma_\eta^2+\sigma_\gamma^2}\left(\frac{\bar{\eta}^2}{\sigma_\gamma^2+\sigma_\eta^2+\bar{\eta}^2}\gamma-\bar{\eta}\right)\\ \Rightarrow\quad \mathbb{E}[y(\tilde{\gamma})]-\mathbb{E}[y^*(\tilde{\gamma})] &= \frac{1}{2\theta}\underbrace{(\alpha_B-\alpha_I)}_{>0}\frac{\sigma_\gamma^2}{\sigma_\eta^2+\sigma_\gamma^2}\left(\frac{\bar{\eta}^2}{\sigma_\gamma^2+\sigma_\eta^2+\bar{\eta}^2}-1\right)\bar{\eta}\end{aligned} \tag{49}$$

Moreover:

$$\begin{aligned}
\mathbb{E}[\tilde{\eta}\tilde{y}] &= \mathbb{E}\left[\frac{1}{2\theta}\left(\left((\alpha_B-\alpha_I)\left(\frac{\sigma_\gamma^2}{\sigma_\eta^2+\sigma_\gamma^2}\frac{\bar{\eta}^2}{\sigma_\gamma^2+\sigma_\eta^2+\bar{\eta}^2}+\frac{\sigma_\eta^2}{\sigma_\eta^2+\sigma_\gamma^2}\right)+\alpha_I\frac{\sigma_\eta^2}{\sigma_\eta^2+\sigma_\gamma^2}\right)\tilde{\gamma}+\alpha_I\frac{\sigma_\gamma^2}{\sigma_\eta^2+\sigma_\gamma^2}\bar{\eta}\right)\tilde{\eta}\right] \\
&= \frac{1}{2\theta}\left(\left((\alpha_B-\alpha_I)\left(\frac{\sigma_\gamma^2}{\sigma_\eta^2+\sigma_\gamma^2}\frac{\bar{\eta}^2}{\sigma_\gamma^2+\sigma_\eta^2+\bar{\eta}^2}+\frac{\sigma_\eta^2}{\sigma_\eta^2+\sigma_\gamma^2}\right)+\alpha_I\frac{\sigma_\eta^2}{\sigma_\eta^2+\sigma_\gamma^2}\right)\left(\sigma_\eta^2+\bar{\eta}^2\right)+\alpha_I\frac{\sigma_\gamma^2}{\sigma_\eta^2+\sigma_\gamma^2}\bar{\eta}^2\right) \quad (50)
\end{aligned}$$

■

Proof of Proposition 2:

With N ESG scores whose noise terms are i.i.d., the average score is a sufficient statistic for the mean of the distribution, i.e. η (from a standard application of the factorization theorem). Define:

$$\bar{m} \equiv \frac{1}{N}\sum_{j=1}^{N} m_j \tag{51}$$

We have:

$$\frac{1}{N}\sum_{j=1}^{N}\tilde{m}_j = \frac{1}{N}\sum_{j=1}^{N} y\left(\eta+\tilde{\zeta}_j\right) = \eta y + \frac{y}{N}\sum_{j=1}^{N}\tilde{\zeta}_j \tag{52}$$

where the random variable $\frac{1}{N}\sum_{j=1}^{N}\tilde{\zeta}_j$ is normally distributed with the following mean and variance:

$$\begin{aligned}
\mathbb{E}\left[\frac{1}{N}\sum_{j=1}^{N}\tilde{\zeta}_j\right] &= \frac{1}{N}\sum_{j=1}^{N}\mathbb{E}\left[\tilde{\zeta}_j\right] = 0 && (53)\\
var\left(\frac{1}{N}\sum_{j=1}^{N}\tilde{\zeta}_j\right) &= \frac{1}{N^2}\sum_{j=1}^{N} var\left(\tilde{\zeta}_j\right) = \frac{1}{N}\sigma_\gamma^2 && (54)
\end{aligned}$$

where we used the i.i.d. assumption about the noise terms. Therefore, the unconditional variance of $\tilde{\bar{m}}/y$ is: $var\left(\tilde{\bar{m}}/y\right) = \sigma_\eta^2 + \frac{1}{N}\sigma_\gamma^2$.

We follow the same steps as in previous proofs and accordingly we omit or abbreviate some steps. Using the standard formula for the conditional expectation with normal distributions, beliefs are updated at $t = 1$ as follows:

$$\mathbb{E}[\tilde{v}|y,z,m] \;=\; \mathbb{E}\left[\tilde{\eta}\tilde{y} + \tilde{\epsilon}_y|y,\bar{m}\right] \;=\; y\mathbb{E}\left[\tilde{\eta}|y,\bar{m}\right] \tag{55}$$

$$\mathbb{E}\left[\tilde{\eta}|y,\bar{m}\right] \;=\; \mathbb{E}\left[\tilde{\eta}\right] + \frac{cov\left(\tilde{\eta},\tilde{\bar{m}}/y\right)}{var\left(\tilde{\bar{m}}/y\right)}\left(\bar{m}/y - \mathbb{E}[\tilde{\bar{m}}/y]\right) = \bar{\eta} + \frac{\sigma_\eta^2}{\sigma_\eta + \sigma_\gamma^2/N}\left(\bar{m}/y - \bar{\eta}\right) \tag{56}$$

Substituting into equation (55):

$$\mathbb{E}[\tilde{v}|y,z,m] \;=\; y\left(\bar{\eta} + \frac{\sigma_\eta^2}{\sigma_\eta^2 + \sigma_\gamma^2/N}\left(\bar{m}/y - \bar{\eta}\right)\right) = \frac{\sigma_\eta^2}{\sigma_\eta^2 + \sigma_\gamma^2/N}\bar{m} + \frac{\sigma_\gamma^2/N}{\sigma_\eta^2 + \sigma_\gamma^2/N}y\bar{\eta} \tag{57}$$

The stock price p is set accordingly.

Since the average ESG score $\bar{m}$ is a sufficient statistic, we consider contracts based on the average score $\bar{m}$ rather than on each individual score. The average score is defined as: $\bar{m} = y\bar{\gamma}$ with $\bar{\gamma} = \frac{1}{N}\sum_{j=1}^{N}\gamma_j$. Rewriting the manager's objective function accordingly gives:

$$\max_{e,y}\mathbb{E}\left[w + \beta_x\left(e - \theta y^2\right) + \beta_p\tilde{p} + \beta_m\bar{m}\middle|\bar{\gamma}\right] - C(e) \tag{58}$$

$$\Rightarrow \quad y(\bar{\gamma}) = \frac{\left(\beta_m + \beta_p\alpha_I\frac{\sigma_\eta^2}{\sigma_\eta^2+\sigma_\gamma^2/N}\right)\bar{\gamma} + \beta_p\alpha_I\frac{\sigma_\gamma^2/N}{\sigma_\eta^2+\sigma_\gamma^2/N}\bar{\eta}}{\beta_x + \beta_p}\frac{1}{2\theta} \tag{59}$$

Following the same steps as in the proof of Lemma 3, maximizing the board's objective function is equivalent to minimizing the expected quadratic distance between $y(\bar{\gamma})$ and $y^*(\bar{\gamma})$, where:

$$\mathbb{E}\left[\tilde{\eta}|\bar{\gamma}\right] \;=\; \frac{\sigma_\eta^2}{\sigma_\eta^2 + \sigma_\gamma^2/N}\bar{\gamma} + \frac{\sigma_\gamma^2/N}{\sigma_\eta^2 + \sigma_\gamma^2/N}\bar{\eta} \tag{60}$$

$$y^*(\bar{\gamma}) \;=\; \frac{\alpha_B}{2}\frac{1}{\theta}\left(\frac{\sigma_\eta^2}{\sigma_\eta^2 + \sigma_\gamma^2/N}\bar{\gamma} + \frac{\sigma_\gamma^2/N}{\sigma_\eta^2 + \sigma_\gamma^2/N}\bar{\eta}\right) \tag{61}$$

This problem is:

$$\min_{\beta_x,\beta_p,\beta_m}\int_\eta\int_{\bar{\gamma}}\left(\frac{1}{2}\frac{1}{\theta}\frac{\left(\beta_m + \beta_p\alpha_I\frac{\sigma_\eta^2}{\sigma_\eta^2+\sigma_\gamma^2/N}\right)\bar{\gamma} + \beta_p\alpha_I\frac{\sigma_\gamma^2/N}{\sigma_\eta^2+\sigma_\gamma^2/N}\bar{\eta}}{\beta_x + \beta_p}\right.$$
$$\left.-\frac{\alpha_B}{2}\frac{1}{\theta}\left(\frac{\sigma_\eta^2}{\sigma_\eta^2 + \sigma_\gamma^2/N}\bar{\gamma} + \frac{\sigma_\gamma^2/N}{\sigma_\eta^2 + \sigma_\gamma^2/N}\bar{\eta}\right)\right)^2\varphi(\bar{\gamma}|\eta)\phi(\eta)d\bar{\gamma}d\eta \tag{62}$$

We have: $\int_{\bar{\gamma}} \bar{\gamma}^2 \varphi(\bar{\gamma}|\eta) d\bar{\gamma} = \mathbb{E}[\tilde{\bar{\gamma}}^2|\eta] = var(\tilde{\bar{\gamma}}|\eta) + (\mathbb{E}[\tilde{\bar{\gamma}}|\eta])^2 = \sigma_\gamma^2/N + \eta^2$. This implies:

$$\int_\eta \int_{\bar{\gamma}} \bar{\gamma}^2 \varphi(\bar{\gamma}|\eta)\phi(\eta) d\bar{\gamma} d\eta = \int_\eta \left(\sigma_\gamma^2/N + \eta^2\right) \phi(\eta) d\eta = \sigma_\gamma^2/N + \int_\eta \eta^2 \phi(\eta) d\eta = \sigma_\gamma^2/N + \sigma_\eta^2 + \bar{\eta}^2 \quad (63)$$

The expression in equation (62) is globally convex with respect to β_m. The FOC w.r.t. β_m is:

$$\frac{\beta_m}{\beta_x + \beta_p} = \left(\alpha_B - \frac{\beta_p \alpha_I}{\beta_x + \beta_p}\right) \left(\frac{\sigma_\gamma^2/N}{\sigma_\eta^2 + \sigma_\gamma^2/N} \frac{\bar{\eta}^2}{\sigma_\gamma^2/N + \sigma_\eta^2 + \bar{\eta}^2} + \frac{\sigma_\eta^2}{\sigma_\eta^2 + \sigma_\gamma^2/N} \right) \quad (64)$$

This shows that the principal's objective function is minimized by minimizing $\left(\alpha_B - \frac{\beta_p \alpha_I}{\beta_x + \beta_p}\right)^2$. With a nonnegativity constraint on β_x, this is achieved by setting $\beta_x = \max\{\frac{\beta_p \alpha_I}{\alpha_B} - \beta_p, 0\}$. As above, for $\alpha_B \leq \alpha_I$, we have $\beta_p > 0$, $\beta_x = \frac{\beta_p \alpha_I}{\alpha_B} - \beta_p \geq 0$, β_m as defined as in equation (64) is equal to zero, and social investment is:

$$y(\bar{\gamma}) = \frac{\beta_p \alpha_I \left(\frac{\sigma_\eta^2}{\sigma_\eta^2 + \sigma_\gamma^2/N} \tilde{\bar{\gamma}} + \frac{\sigma_\gamma^2/N}{\sigma_\eta^2 + \sigma_\gamma^2/N} \bar{\eta} \right)}{\beta_x + \beta_p} \frac{1}{2\theta} = \frac{\alpha_B}{2} \frac{1}{\theta} \left(\frac{\sigma_\eta^2}{\sigma_\eta^2 + \sigma_\gamma^2/N} \tilde{\bar{\gamma}} + \frac{\sigma_\gamma^2/N}{\sigma_\eta^2 + \sigma_\gamma^2/N} \bar{\eta} \right),$$

which is the same as $y^*(\bar{\gamma})$ as defined in equation (61). Moreover, expected social output when $\alpha_B \leq \alpha_I$ is:

$$\mathbb{E}[\tilde{\eta}\tilde{y}] = \frac{\beta_p \alpha_I \left(\frac{\sigma_\eta^2}{\sigma_\eta^2 + \sigma_\gamma^2/N} \mathbb{E}[\tilde{\eta}\tilde{\bar{\gamma}}] + \frac{\sigma_\gamma^2/N}{\sigma_\eta^2 + \sigma_\gamma^2/N} \bar{\eta} \mathbb{E}[\tilde{\eta}] \right)}{\beta_x + \beta_p} \frac{1}{2\theta} = \frac{\alpha_B}{2\theta} \left(\frac{\sigma_\eta^4}{\sigma_\eta^2 + \sigma_\gamma^2/N} + \bar{\eta}^2 \right) \quad (65)$$

In the limit:

$$\lim_{N \to \infty} \mathbb{E}[\tilde{\eta}\tilde{y}] = \frac{\alpha_B}{2\theta} \left(\bar{\eta}^2 + \sigma_\eta^2\right)$$

For $\alpha_B > \alpha_I$, we have $\beta_x = 0$ and β_p is as in equation (30) In this case, $\beta_m > 0$ is as in equation (64). Substituting for β_x and β_p in equation (64) gives β_m. Substituting for β_m, β_p, and β_x in equation (59):

$$\begin{aligned} y(\bar{\gamma}) = \frac{1}{2\theta} \Bigg(& (\alpha_B - \alpha_I) \left(\frac{\sigma_\gamma^2/N}{\sigma_\eta^2 + \sigma_\gamma^2/N} \frac{\bar{\eta}^2}{\sigma_\gamma^2/N + \sigma_\eta^2 + \bar{\eta}^2} + \frac{\sigma_\eta^2}{\sigma_\eta^2 + \sigma_\gamma^2/N} \right) \bar{\gamma} \\ & + \alpha_I \left(\frac{\sigma_\eta^2}{\sigma_\eta^2 + \sigma_\gamma^2/N} \bar{\gamma} + \frac{\sigma_\gamma^2/N}{\sigma_\eta^2 + \sigma_\gamma^2/N} \bar{\eta} \right) \Bigg) \end{aligned}$$

We have:

$$y(\bar{\gamma}) - y^*(\bar{\gamma}) = \frac{1}{2\theta}\left((\alpha_B - \alpha_I)\left(\frac{\sigma_\gamma^2/N}{\sigma_\eta^2 + \sigma_\gamma^2/N}\frac{\bar{\eta}^2}{\sigma_\gamma^2/N + \sigma_\eta^2 + \bar{\eta}^2}\bar{\gamma} - \frac{\sigma_\gamma^2/N}{\sigma_\eta^2 + \sigma_\gamma^2/N}\bar{\eta}\right)\right) \tag{66}$$

The agency cost is derived as in Lemma 3. It is defined as a function of the number N of scores:

$$\begin{aligned} AC(N) &= \theta \int_\eta \int_\gamma \left(y(\bar{\gamma}) - y^*(\bar{\gamma})\right)^2 \varphi(\bar{\gamma}|\eta)\phi(\eta) d\bar{\gamma} d\eta \\ &= \frac{(\alpha_B - \alpha_I)^2}{2\theta} \int_\eta \int_\gamma \left(\frac{\sigma_\gamma^2/N}{\sigma_\eta^2 + \sigma_\gamma^2/N}\frac{\bar{\eta}^2}{\sigma_\gamma^2/N + \sigma_\eta^2 + \bar{\eta}^2}\bar{\gamma} - \frac{\sigma_\gamma^2/N}{\sigma_\eta^2 + \sigma_\gamma^2/N}\bar{\eta}\right)^2 \varphi(\bar{\gamma}|\eta)\phi(\eta) d\bar{\gamma} d\eta \quad (67) \end{aligned}$$

In the limit, the agency cost is zero: $\lim_{N\to\infty} AC(N) = 0$. Moreover, expected social output when $\alpha_B > \alpha_I$ is:

$$\begin{aligned} \mathbb{E}[\tilde{\eta}\tilde{y}] = \frac{1}{2\theta}\Bigg(&(\alpha_B - \alpha_I)\left(\frac{\sigma_\gamma^2/N}{\sigma_\eta^2 + \sigma_\gamma^2/N}\frac{\bar{\eta}^2}{\sigma_\gamma^2/N + \sigma_\eta^2 + \bar{\eta}^2} + \frac{\sigma_\eta^2}{\sigma_\eta^2 + \sigma_\gamma^2/N}\right)\left(\bar{\eta}^2 + \sigma_\eta^2\right) \\ &+\alpha_I\left(\bar{\eta}^2 + \frac{\sigma_\eta^2}{\sigma_\eta^2 + \sigma_\gamma^2/N}\sigma_\eta^2\right)\Bigg) \end{aligned}$$

In the limit:

$$\lim_{N\to\infty} \mathbb{E}[\tilde{\eta}\tilde{y}] = \frac{1}{2\theta}\left((\alpha_B - \alpha_I)\left(\bar{\eta}^2 + \sigma_\eta^2\right) + \alpha_I\left(\bar{\eta}^2 + \sigma_\eta^2\right)\right) = \frac{\alpha_B}{2\theta}\left(\bar{\eta}^2 + \sigma_\eta^2\right)$$

■

Proof of Proposition 3:

We calculate the conditional distribution of $\tilde{\eta}$ after the observations of $\{\gamma_1, \gamma_2\}$. The variables $\tilde{\eta}, \tilde{\gamma}_1, \tilde{\gamma}_2$ follow a multivariate normal distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ such that:

$$\boldsymbol{\mu} = \begin{pmatrix} \mathbb{E}[\tilde{\eta}] \\ \mathbb{E}[\tilde{\gamma}_1] \\ \mathbb{E}[\tilde{\gamma}_2] \end{pmatrix} = \begin{pmatrix} \bar{\eta} \\ \bar{\eta} \\ \bar{\eta} \end{pmatrix}, \qquad \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix} \tag{68}$$

where

$$\mathbf{\Sigma}_{11} = var\left(\tilde{\eta}\right), \quad \mathbf{\Sigma}_{12} = \begin{pmatrix} cov\left(\tilde{\eta}, \tilde{\gamma}_1\right) & cov\left(\tilde{\eta}, \tilde{\gamma}_2\right) \end{pmatrix},$$

$$\mathbf{\Sigma}_{21} = \mathbf{\Sigma}_{12}^T, \quad \mathbf{\Sigma}_{22} = \begin{pmatrix} var\left(\tilde{\gamma}_1\right) & cov\left(\tilde{\gamma}_1, \tilde{\gamma}_2\right) \\ cov\left(\tilde{\gamma}_2, \tilde{\gamma}_1\right) & var\left(\tilde{\gamma}_2\right) \end{pmatrix}$$

$$\Rightarrow \qquad \mathbf{\Sigma}_{11} = \sigma_\eta^2, \quad \mathbf{\Sigma}_{12} = \begin{pmatrix} \sigma_\eta^2 & \sigma_\eta^2 \end{pmatrix}, \quad \mathbf{\Sigma}_{21} = \mathbf{\Sigma}_{12}^T, \quad \mathbf{\Sigma}_{22} = \begin{pmatrix} \sigma_\eta^2 + \sigma_{\gamma_1}^2 & \sigma_\eta^2 + \rho\sigma_{\gamma_1}\sigma_{\gamma_2} \\ \sigma_\eta^2 + \rho\sigma_{\gamma_1}\sigma_{\gamma_2} & \sigma_\eta^2 + \sigma_{\gamma_2}^2 \end{pmatrix}$$

The posterior distribution of $\tilde{\eta}$ after observing $\{\gamma_1, \gamma_2\}$ is normal with mean and variance:

$$\begin{aligned}
\mathbb{E}[\tilde{\eta}|\gamma_1, \gamma_2] &= \bar{\eta} + \mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1} \begin{pmatrix} \gamma_1 - \bar{\eta} \\ \gamma_2 - \bar{\eta} \end{pmatrix} \\
&= \bar{\eta} + \begin{pmatrix} \sigma_\eta^2 & \sigma_\eta^2 \end{pmatrix} \begin{pmatrix} \frac{\sigma_\eta^2 + \sigma_{\gamma_2}^2}{\sigma_{\gamma_1}^2\sigma_{\gamma_2}^2 - \rho^2\sigma_{\gamma_1}^2\sigma_{\gamma_2}^2 + \sigma_\eta^2(\sigma_{\gamma_1}^2 + \sigma_{\gamma_2}^2 - 2\rho\sigma_{\gamma_1}\sigma_{\gamma_2})} & \frac{-\sigma_\eta^2 - \rho\sigma_{\gamma_1}\sigma_{\gamma_2}}{\sigma_{\gamma_1}^2\sigma_{\gamma_2}^2 - \rho^2\sigma_{\gamma_1}^2\sigma_{\gamma_2}^2 + \sigma_\eta^2(\sigma_{\gamma_1}^2 + \sigma_{\gamma_2}^2 - 2\rho\sigma_{\gamma_1}\sigma_{\gamma_2})} \\ \frac{-\sigma_\eta^2 - \rho\sigma_{\gamma_1}\sigma_{\gamma_2}}{\sigma_{\gamma_1}^2\sigma_{\gamma_2}^2 - \rho^2\sigma_{\gamma_1}^2\sigma_{\gamma_2}^2 + \sigma_\eta^2(\sigma_{\gamma_1}^2 + \sigma_{\gamma_2}^2 - 2\rho\sigma_{\gamma_1}\sigma_{\gamma_2})} & \frac{\sigma_\eta^2 + \sigma_{\gamma_1}^2}{\sigma_{\gamma_1}^2\sigma_{\gamma_2}^2 - \rho^2\sigma_{\gamma_1}^2\sigma_{\gamma_2}^2 + \sigma_\eta^2(\sigma_{\gamma_1}^2 + \sigma_{\gamma_2}^2 - 2\rho\sigma_{\gamma_1}\sigma_{\gamma_2})} \end{pmatrix} \begin{pmatrix} \gamma_1 - \bar{\eta} \\ \gamma_2 - \bar{\eta} \end{pmatrix} \\
&= \bar{\eta} + \frac{\sigma_\eta^2(\sigma_{\gamma_2}^2 - \rho\sigma_{\gamma_1}\sigma_{\gamma_2})}{\sigma_{\gamma_1}^2\sigma_{\gamma_2}^2 - \rho^2\sigma_{\gamma_1}^2\sigma_{\gamma_2}^2 + \sigma_\eta^2(\sigma_{\gamma_1}^2 + \sigma_{\gamma_2}^2 - 2\rho\sigma_{\gamma_1}\sigma_{\gamma_2})}\left(\gamma_1 - \bar{\eta}\right) \\
&\quad + \frac{\sigma_\eta^2(\sigma_{\gamma_1}^2 - \rho\sigma_{\gamma_1}\sigma_{\gamma_2})}{\sigma_{\gamma_1}^2\sigma_{\gamma_2}^2 - \rho^2\sigma_{\gamma_1}^2\sigma_{\gamma_2}^2 + \sigma_\eta^2(\sigma_{\gamma_1}^2 + \sigma_{\gamma_2}^2 - 2\rho\sigma_{\gamma_1}\sigma_{\gamma_2})}\left(\gamma_2 - \bar{\eta}\right)
\end{aligned}$$

$$\begin{aligned}
\sigma_{\eta|\gamma_1,\gamma_2}^2 &= \sigma_\eta^2 - \mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}\mathbf{\Sigma}_{21} \\
&= \sigma_\eta^2 - \frac{\sigma_\eta^4\left(\sigma_{\gamma_1}^2 + \sigma_{\gamma_2}^2 - 2\rho\sigma_{\gamma_1}\sigma_{\gamma_2}\right)}{\sigma_\eta^2\sigma_{\gamma_1}^2 + \sigma_\eta^2\sigma_{\gamma_2}^2 + \sigma_{\gamma_1}^2\sigma_{\gamma_2}^2 - \rho^2\sigma_{\gamma_1}^2\sigma_{\gamma_2}^2 - 2\rho\sigma_\eta^2\sigma_{\gamma_1}\sigma_{\gamma_2}}
\end{aligned}$$

The stock price p is set accordingly.

Rewriting the manager's objective function in the case with two ESG scores, and denoting by β_m^j the sensivity of pay to score j, gives:

$$\max_{e,y} \mathbb{E}\left[w + \beta_x\left(e - \theta y^2\right) + \beta_p\tilde{p} + \sum_{j=1,2}\beta_m^j m_j \middle| \gamma_1, \gamma_2\right] - C(e) \tag{69}$$

For $\beta_x + \beta_p > 0$, the manager's objective function is concave in y. The FOC with respect to y

is, for $j \neq k$:

$$\beta_x(-2\theta y) - 2\beta_p\theta y + \alpha_I\beta_p\left(\bar{\eta} + \sum_{k=1,2}\frac{\sigma_\eta^2(\sigma_{\gamma_j}^2 - \rho\sigma_{\gamma_k}\sigma_{\gamma_j})}{\sigma_{\gamma_k}^2\sigma_{\gamma_j}^2 - \rho^2\sigma_{\gamma_k}^2\sigma_{\gamma_j}^2 + \sigma_\eta^2(\sigma_{\gamma_k}^2 + \sigma_{\gamma_j}^2 - 2\rho\sigma_{\gamma_k}\sigma_{\gamma_j})}(\gamma_k - \bar{\eta})\right) + \sum_{k=1,2}\beta_m^k\gamma_k = 0$$

$$\Leftrightarrow \quad y(\gamma_1,\gamma_2) = \frac{1}{2\theta}\frac{\sum_{k=1,2}\beta_m^k\gamma_k + \alpha_I\beta_p\left(\bar{\eta} + \sum_{k=1,2}\frac{\sigma_\eta^2(\sigma_{\gamma_j}^2 - \rho\sigma_{\gamma_k}\sigma_{\gamma_j})}{\sigma_{\gamma_k}^2\sigma_{\gamma_j}^2 - \rho^2\sigma_{\gamma_k}^2\sigma_{\gamma_j}^2 + \sigma_\eta^2(\sigma_{\gamma_k}^2 + \sigma_{\gamma_j}^2 - 2\rho\sigma_{\gamma_k}\sigma_{\gamma_j})}(\gamma_k - \bar{\eta})\right)}{\beta_x + \beta_p} \tag{70}$$

Given his contract, the manager will optimally exert high effort ($e = \bar{e}$) if and only if equation (30) is satisfied. The fixed component of pay, w, is set to guarantee the manager's participation for $e = \bar{e}$:

$$\begin{aligned} \mathbb{E}[u(y,\bar{e})] &= \bar{W} \\ \Leftrightarrow \quad w + \beta_x\mathbb{E}\left[\bar{e} - \theta y^2 + \tilde{\epsilon}_x\right] + \beta_p\mathbb{E}[\tilde{p}] + \sum_{j=1,2}\beta_m^j\mathbb{E}[\tilde{m}_j] &= \bar{W} + C(\bar{e}) \end{aligned} \tag{71}$$

Thus, the board's objective function with high effort can be rewritten as:

$$\begin{aligned} \mathbb{E}[V(y,\bar{e})] &= \mathbb{E}\left[(1-\beta_x)\left(\bar{e} - \theta y^2 + \tilde{\epsilon}_x\right) + \alpha_B(\tilde{\eta}y + \tilde{\epsilon}_y) - \left(w + \beta_p\tilde{p} + \sum_{j=1,2}\beta_m^j\tilde{m}_j\right)\right] \\ &= \mathbb{E}\left[\bar{e} - \theta y^2 + \tilde{\epsilon}_x + \alpha_B(\tilde{\eta}y + \tilde{\epsilon}_y) - \left(\bar{W} + C(\bar{e})\right)\right] \end{aligned} \tag{72}$$

Following the same steps as in the proof of Lemma 3, maximizing the board's objective function is equivalent to minimizing the expected quadratic distance between $y(\gamma_1,\gamma_2)$ and $y^*(\gamma_1,\gamma_2)$, where:

$$\begin{aligned} \mathbb{E}\left[\tilde{\eta}|\gamma_1,\gamma_2\right] &= \bar{\eta} + \sum_{k=1,2}\frac{\sigma_\eta^2(\sigma_{\gamma_j}^2 - \rho\sigma_{\gamma_k}\sigma_{\gamma_j})}{\sigma_{\gamma_k}^2\sigma_{\gamma_j}^2 - \rho^2\sigma_{\gamma_k}^2\sigma_{\gamma_j}^2 + \sigma_\eta^2(\sigma_{\gamma_k}^2 + \sigma_{\gamma_j}^2 - 2\rho\sigma_{\gamma_k}\sigma_{\gamma_j})}(\gamma_k - \bar{\eta}) \\ y^*(\gamma_1,\gamma_2) &= \frac{\alpha_B}{2}\frac{1}{\theta}\left(\bar{\eta} + \sum_{k=1,2}\frac{\sigma_\eta^2(\sigma_{\gamma_j}^2 - \rho\sigma_{\gamma_k}\sigma_{\gamma_j})}{\sigma_{\gamma_k}^2\sigma_{\gamma_j}^2 - \rho^2\sigma_{\gamma_k}^2\sigma_{\gamma_j}^2 + \sigma_\eta^2(\sigma_{\gamma_k}^2 + \sigma_{\gamma_j}^2 - 2\rho\sigma_{\gamma_k}\sigma_{\gamma_j})}(\gamma_k - \bar{\eta})\right) \end{aligned} \tag{73}$$

This problem is:

$$\min_{\beta_x,\beta_p,\beta_m^k} \sum_{k=1,2} \int_\eta \int_{\gamma_k} \left(\frac{1}{2}\frac{1}{\theta} \frac{\beta_m^k \gamma_k + \alpha_I \beta_p \left(\frac{\bar{\eta}}{2} + \frac{\sigma_\eta^2(\sigma_{\gamma_j}^2 - \rho\sigma_{\gamma_k}\sigma_{\gamma_j})}{\sigma_{\gamma_k}^2\sigma_{\gamma_j}^2 - \rho^2\sigma_{\gamma_k}^2\sigma_{\gamma_j}^2 + \sigma_\eta^2(\sigma_{\gamma_k}^2 + \sigma_{\gamma_j}^2 - 2\rho\sigma_{\gamma_k}\sigma_{\gamma_j})} (\gamma_k - \bar{\eta}) \right)}{\beta_x + \beta_p} \right.$$
$$\left. - \frac{\alpha_B}{2}\frac{1}{\theta} \left(\frac{\bar{\eta}}{2} + \frac{\sigma_\eta^2(\sigma_{\gamma_j}^2 - \rho\sigma_{\gamma_k}\sigma_{\gamma_j})}{\sigma_{\gamma_k}^2\sigma_{\gamma_j}^2 - \rho^2\sigma_{\gamma_k}^2\sigma_{\gamma_j}^2 + \sigma_\eta^2(\sigma_{\gamma_k}^2 + \sigma_{\gamma_j}^2 - 2\rho\sigma_{\gamma_k}\sigma_{\gamma_j})} (\gamma_k - \bar{\eta}) \right) \right)^2 \varphi(\gamma_k|\eta)\phi(\eta) d\gamma_k d\eta$$
$$\min_{\beta_x,\beta_p,\beta_m^k} \frac{1}{2}\frac{1}{\theta} \sum_{k=1,2} \int_\eta \int_{\gamma_k} \left(\left(\frac{\beta_m^k}{\beta_x + \beta_p} + \left(\frac{\alpha_I \beta_p}{\beta_x + \beta_p} - \alpha_B \right) \frac{\sigma_\eta^2(\sigma_{\gamma_j}^2 - \rho\sigma_{\gamma_k}\sigma_{\gamma_j})}{\sigma_{\gamma_k}^2\sigma_{\gamma_j}^2 - \rho^2\sigma_{\gamma_k}^2\sigma_{\gamma_j}^2 + \sigma_\eta^2(\sigma_{\gamma_k}^2 + \sigma_{\gamma_j}^2 - 2\rho\sigma_{\gamma_k}\sigma_{\gamma_j})} \right) \gamma_k \right.$$
$$\left. + \left(\frac{\alpha_I \beta_p}{\beta_x + \beta_p} - \alpha_B \right) \left(\frac{1}{2} - \frac{\sigma_\eta^2(\sigma_{\gamma_j}^2 - \rho\sigma_{\gamma_k}\sigma_{\gamma_j})}{\sigma_{\gamma_k}^2\sigma_{\gamma_j}^2 - \rho^2\sigma_{\gamma_k}^2\sigma_{\gamma_j}^2 + \sigma_\eta^2(\sigma_{\gamma_k}^2 + \sigma_{\gamma_j}^2 - 2\rho\sigma_{\gamma_k}\sigma_{\gamma_j})} \right) \bar{\eta} \right)^2 \varphi(\gamma_k|\eta)\phi(\eta) d\gamma_k d\eta \qquad (74)$$

The expression in equation (74) is globally convex with respect to β_m^k. The FOC w.r.t. β_m^k when the nonnegativity constraints on β_m^j are not binding gives:

$$\frac{\beta_m^k}{\beta_x + \beta_p} = \left(\alpha_B - \frac{\alpha_I \beta_p}{\beta_x + \beta_p} \right) \left(\frac{1}{2} \frac{\bar{\eta}^2}{\sigma_{\gamma_k}^2 + \sigma_\eta^2 + \bar{\eta}^2} \right.$$
$$\left. + \frac{\sigma_\eta^2(\sigma_{\gamma_j}^2 - \rho\sigma_{\gamma_k}\sigma_{\gamma_j})}{\sigma_{\gamma_k}^2\sigma_{\gamma_j}^2 - \rho^2\sigma_{\gamma_k}^2\sigma_{\gamma_j}^2 + \sigma_\eta^2(\sigma_{\gamma_k}^2 + \sigma_{\gamma_j}^2 - 2\rho\sigma_{\gamma_k}\sigma_{\gamma_j})} \left(1 - \frac{\bar{\eta}^2}{\sigma_{\gamma_k}^2 + \sigma_\eta^2 + \bar{\eta}^2} \right) \right) \qquad (75)$$

Letting $\Sigma_k \equiv \frac{\sigma_\eta^2(\sigma_{\gamma_j}^2 - \rho\sigma_{\gamma_k}\sigma_{\gamma_j})}{\sigma_{\gamma_k}^2\sigma_{\gamma_j}^2 - \rho^2\sigma_{\gamma_k}^2\sigma_{\gamma_j}^2 + \sigma_\eta^2(\sigma_{\gamma_k}^2 + \sigma_{\gamma_j}^2 - 2\rho\sigma_{\gamma_k}\sigma_{\gamma_j})}$ and substituting in equation (74) when β_m^k is as in equation (75):

$$\min_{\beta_x,\beta_p} \frac{1}{2}\frac{1}{\theta} \sum_{k=1,2} \int_\eta \int_{\gamma_k} \left(\left(\left(\alpha_B - \frac{\alpha_I \beta_p}{\beta_x + \beta_p} \right) \left(\frac{1}{2} \frac{\bar{\eta}^2}{\sigma_{\gamma_k}^2 + \sigma_\eta^2 + \bar{\eta}^2} + \Sigma_k \left(1 - \frac{\bar{\eta}^2}{\sigma_{\gamma_k}^2 + \sigma_\eta^2 + \bar{\eta}^2} \right) \right) \right. \right.$$
$$\left. \left. + \left(\frac{\alpha_I \beta_p}{\beta_x + \beta_p} - \alpha_B \right) \Sigma_k \right) \gamma_k + \left(\frac{\alpha_I \beta_p}{\beta_x + \beta_p} - \alpha_B \right) \left(\frac{1}{2} - \Sigma_k \right) \bar{\eta} \right)^2 \varphi(\gamma_k|\eta)\phi(\eta) d\gamma_k d\eta$$
$$\Leftrightarrow \quad \min_{\beta_x,\beta_p} \left(\alpha_B - \frac{\alpha_I \beta_p}{\beta_x + \beta_p} \right)^2 \frac{1}{2}\frac{1}{\theta} \sum_{k=1,2} \int_\eta \int_{\gamma_k} \left(\left(\left(\frac{1}{2} \frac{\bar{\eta}^2}{\sigma_{\gamma_k}^2 + \sigma_\eta^2 + \bar{\eta}^2} + \Sigma_k \left(1 - \frac{\bar{\eta}^2}{\sigma_{\gamma_k}^2 + \sigma_\eta^2 + \bar{\eta}^2} \right) \right) - \Sigma_k \right) \gamma_k \right.$$
$$\left. - \left(\frac{1}{2} - \Sigma_k \right) \bar{\eta} \right)^2 \varphi(\gamma_k|\eta)\phi(\eta) d\gamma_k d\eta \qquad (76)$$

The expression under the integral signs in equation (76) is positive and independent from the contract. The expression in equation (76) is minimized by minimizing $\left(\alpha_B - \frac{\alpha_I \beta_p}{\beta_x + \beta_p} \right)^2$. This is achieved as in the proof of Proposition 1.

First, when $\alpha_B \leq \alpha_I$ so that $\beta_m^k = 0$ for any k, substituting in equation (70):

$$
\begin{aligned}
y(\gamma_1, \gamma_2) &= \alpha_B \left(\bar{\eta} + \sum_{k=1,2} \frac{\sigma_\eta^2(\sigma_{\gamma_j}^2 - \rho\sigma_{\gamma_k}\sigma_{\gamma_j})}{\sigma_{\gamma_k}^2\sigma_{\gamma_j}^2 - \rho^2\sigma_{\gamma_k}^2\sigma_{\gamma_j}^2 + \sigma_\eta^2(\sigma_{\gamma_k}^2 + \sigma_{\gamma_j}^2 - 2\rho\sigma_{\gamma_k}\sigma_{\gamma_j})} (\gamma_k - \bar{\eta}) \right) \frac{1}{2\theta} \\
&= \alpha_B \left(\bar{\eta} + \sum_{k=1,2} \Sigma_k (\gamma_k - \bar{\eta}) \right) \frac{1}{2\theta},
\end{aligned} \tag{77}
$$

which is as in equation (73). In this case, expected social output is:

$$
\begin{aligned}
\mathbb{E}[\tilde{\eta}\tilde{y}] &= \frac{\alpha_B}{2\theta} \left(\mathbb{E}[\tilde{\eta}]\bar{\eta} + \sum_{k=1,2} \Sigma_k \left(\mathbb{E}[\tilde{\eta}\tilde{\gamma}_k] - \mathbb{E}[\tilde{\eta}]\bar{\eta} \right) \right) \\
&= \frac{\alpha_B}{2\theta} \left(\bar{\eta}^2 + \sum_{k=1,2} \frac{\sigma_\eta^2(\sigma_{\gamma_j}^2 - \rho\sigma_{\gamma_k}\sigma_{\gamma_j})}{\sigma_{\gamma_k}^2\sigma_{\gamma_j}^2 - \rho^2\sigma_{\gamma_k}^2\sigma_{\gamma_j}^2 + \sigma_\eta^2(\sigma_{\gamma_k}^2 + \sigma_{\gamma_j}^2 - 2\rho\sigma_{\gamma_k}\sigma_{\gamma_j})} \sigma_\eta^2 \right)
\end{aligned}
$$

For $\rho = -1$:

$$
\begin{aligned}
\mathbb{E}[\tilde{\eta}\tilde{y}] &= \frac{\alpha_B}{2\theta} \left(\bar{\eta}^2 + \sum_{k=1,2} \frac{\sigma_\eta^2(\sigma_{\gamma_j}^2 + \sigma_{\gamma_k}\sigma_{\gamma_j})}{\sigma_{\gamma_k}^2\sigma_{\gamma_j}^2 - \sigma_{\gamma_k}^2\sigma_{\gamma_j}^2 + \sigma_\eta^2(\sigma_{\gamma_k}^2 + \sigma_{\gamma_j}^2 + 2\sigma_{\gamma_k}\sigma_{\gamma_j})} \sigma_\eta^2 \right) \\
&= \frac{\alpha_B}{2\theta} \left(\bar{\eta}^2 + \sum_{k=1,2} \frac{\sigma_{\gamma_j}}{\sigma_{\gamma_k} + \sigma_{\gamma_j}} \sigma_\eta^2 \right) \\
&= \frac{\alpha_B}{2\theta} \left(\bar{\eta}^2 + \sigma_\eta^2 \right),
\end{aligned}
$$

which is larger than in equation (48) (the case with one ESG score) for any parameter values. Moreover, $\mathbb{E}[\tilde{\eta}\tilde{y}]$ is continuous in ρ, so that the result about ρ in Proposition 3 when $\alpha_B \leq \alpha_I$ holds by a continuity argument. As $\sigma_{\gamma_2} \to 0$:

$$
\mathbb{E}[\tilde{\eta}\tilde{y}] \to \frac{\alpha_B}{2\theta} \left(\bar{\eta}^2 + \sigma_\eta^2 \right)
$$

which is larger than in equation (48) for any parameter values. Moreover, $\mathbb{E}[\tilde{\eta}\tilde{y}]$ is continuous in σ_{γ_2}, so that the result about σ_{γ_2} in Proposition 3 when $\alpha_B \leq \alpha_I$ holds by a continuity argument.

Second, when $\alpha_B \geq \alpha_I$ so that $\beta_x = 0$, substituting in equation (70) with β_m^k as in equation

(75) when the nonnegativity constraints on β_m^j are not binding:

$$\begin{aligned} y(\gamma_1, \gamma_2) &= \frac{1}{2\theta} \sum_{k=1,2} \left((\alpha_B - \alpha_I) \left(\frac{1}{2} \frac{\bar{\eta}^2}{\sigma_{\gamma_k}^2 + \sigma_\eta^2 + \bar{\eta}^2} + \Sigma_k \left(1 - \frac{\bar{\eta}^2}{\sigma_{\gamma_k}^2 + \sigma_\eta^2 + \bar{\eta}^2} \right) \right) \gamma_k \right. \\ & \left. + \alpha_I \left(\frac{\bar{\eta}}{2} + \Sigma_k (\gamma_k - \bar{\eta}) \right) \right) \end{aligned}$$

In this case, expected social output is:

$$\begin{aligned} \mathbb{E}[\tilde{\eta}\tilde{y}] &= \frac{1}{2\theta} \sum_{k=1,2} \left((\alpha_B - \alpha_I) \left(\frac{1}{2} \frac{\bar{\eta}^2}{\sigma_{\gamma_k}^2 + \sigma_\eta^2 + \bar{\eta}^2} + \Sigma_k \left(1 - \frac{\bar{\eta}^2}{\sigma_{\gamma_k}^2 + \sigma_\eta^2 + \bar{\eta}^2} \right) \right) \left(\bar{\eta}^2 + \sigma_\eta^2 \right) \right. \\ & \left. + \alpha_I \left(\frac{\bar{\eta}^2}{2} + \Sigma_k \left(\bar{\eta}^2 + \sigma_\eta^2 - \bar{\eta}^2 \right) \right) \right) \end{aligned}$$

For $\rho = -1$:

$$\begin{aligned} \mathbb{E}[\tilde{\eta}\tilde{y}] &= \frac{1}{2\theta} \sum_{k=1,2} \left((\alpha_B - \alpha_I) \left(\frac{1}{2} \frac{\bar{\eta}^2}{\sigma_{\gamma_k}^2 + \sigma_\eta^2 + \bar{\eta}^2} + \frac{\sigma_{\gamma_k}}{\sigma_{\gamma_2} + \sigma_{\gamma_1}} \left(1 - \frac{\bar{\eta}^2}{\sigma_{\gamma_k}^2 + \sigma_\eta^2 + \bar{\eta}^2} \right) \right) \left(\bar{\eta}^2 + \sigma_\eta^2 \right) \right. \\ & \left. + \alpha_I \left(\frac{\bar{\eta}^2}{2} + \frac{\sigma_{\gamma_k}}{\sigma_{\gamma_2} + \sigma_{\gamma_1}} \left(\bar{\eta}^2 + \sigma_\eta^2 - \bar{\eta}^2 \right) \right) \right) \\ &= \frac{\alpha_B}{2\theta} \left(\bar{\eta}^2 + \sigma_\eta^2 \right) \end{aligned} \tag{78}$$

which is larger than in equation (50) for any parameter values, and the nonnegativity constraints on β_m^j are not binding. Moreover, $\mathbb{E}[\tilde{\eta}\tilde{y}]$ is continuous in ρ, so that the result about ρ in Proposition 3 when $\alpha_B > \alpha_I$ holds by a continuity argument. As $\sigma_{\gamma_2} \to 0$:

$$\begin{aligned} \mathbb{E}[\tilde{\eta}\tilde{y}] &\to \frac{1}{2\theta} \left((\alpha_B - \alpha_I) \left(\frac{\bar{\eta}^2}{\sigma_{\gamma_k}^2 + \sigma_\eta^2 + \bar{\eta}^2} + \left(1 - \frac{\bar{\eta}^2}{\sigma_{\gamma_1}^2 + \sigma_\eta^2 + \bar{\eta}^2} \right) \right) \left(\bar{\eta}^2 + \sigma_\eta^2 \right) + \alpha_I \left(\bar{\eta}^2 + \sigma_\eta^2 \right) \right) \\ &= \frac{\alpha_B}{2\theta} \left(\bar{\eta}^2 + \sigma_\eta^2 \right) \end{aligned}$$

which is larger than in equation (50) for any parameter values, and the nonnegativity constraints on β_m^j are not binding. Moreover, $\mathbb{E}[\tilde{\eta}\tilde{y}]$ is continuous in σ_{γ_2}, so that the result about σ_{γ_2} in Proposition 3 when $\alpha_B > \alpha_I$ holds by a continuity argument. ■

Online Appendix
Executive Compensation with Environmental and Social Performance

A.1 Multiple dimensions of SEP

In this section, we extend the model by assuming that there are two different dimensions of social performance, and we allow the board to have a different preference with respect to each dimension i of SEP.

Cash flows or "profits" $\tilde{x}$ and overall "social output" $\tilde{\upsilon}$ are now respectively defined as:

$$\begin{aligned} \tilde{x} &= e - \theta_1 y_1^2 - \theta_2 y_2^2 + \tilde{\epsilon}_x \quad \text{where } \tilde{\epsilon}_x \sim \mathcal{N}(0, \sigma_x^2) \\ \tilde{\upsilon} &= \eta_1 y_1 + \eta_2 y_2 + \tilde{\epsilon}_y \qquad\quad \text{where } \tilde{\epsilon}_y \sim \mathcal{N}(0, \sigma_y^2) \end{aligned} \tag{A.1}$$

where θ_1 and θ_2 are positive constants.

Social performance measures are realized at $t = 1$, and they imperfectly reflect the actual SEP of the firm. The measure of the firm's SEP on dimension i (for $i \in \{1, 2\}$), that we will refer to as an "ESG score" for brevity but without loss of generality, is:

$$m_i \equiv \gamma_i y_i \qquad \text{where } \tilde{\gamma}_i \sim \mathcal{N}(\eta_i, \sigma_\gamma^2), \tag{A.2}$$

where η_i is the realization of $\tilde{\eta}_i$. That is, with $\sigma_\gamma^2 > 0$, γ_i is a noisy measure of the firm's social productivity on dimension i. The sensitivity of compensation to score m_i is denoted by β_i. The model is otherwise as described in section 3 of the main paper, with gaming with respect to γ_1 and γ_2 as in section 5.

A.1.1 Baseline model

Lemma 5 *With one score m_i on each dimension i of social performance, the stock price is such that:*

$$\begin{aligned} \mathbb{E}[\tilde{x}|y, z, m_1, m_2] &= \hat{e} - \theta_1 y_1^2 + \theta_2 y_2^2 + \frac{\sigma_x^2}{\sigma_x^2 + \sigma_z^2}\left(z - \hat{e} + \theta_1 y_1^2 + \theta_2 y_2^2\right) \\ \mathbb{E}[\tilde{\upsilon}|y, z, m_1, m_2] &= \frac{\sigma_\eta^2}{\sigma_\eta^2 + \sigma_\gamma^2} m_1 + \frac{\sigma_\gamma^2}{\sigma_\eta^2 + \sigma_\gamma^2} y_1 \bar{\eta} + \frac{\sigma_\eta^2}{\sigma_\eta^2 + \sigma_\gamma^2} m_2 + \frac{\sigma_\gamma^2}{\sigma_\eta^2 + \sigma_\gamma^2} y_2 \bar{\eta} \end{aligned}$$

Proposition 4 describes the use of SEP-based compensation in the optimal linear contract with two dimensions of SEP performance, where β_i is the sensitivity of CEO pay to ESG score m_i.

Proposition 4

- *If $\alpha_B \leq \alpha_I$, then $y_i(\gamma_i) = y_i^*(\gamma_i)\ \forall \gamma_i$, and the optimal linear contract is defined by:*

$$\beta_i = 0, \quad \beta_x = \left(\frac{\alpha_I}{\alpha_B} - 1\right) \frac{c_e}{\bar{e} - \underline{e}} \left(\frac{\alpha_I}{\alpha_B} + \frac{\sigma_x^2}{\sigma_x^2 + \sigma_z^2} - 1\right)^{-1}, \quad \beta_p = \frac{c_e}{\bar{e} - \underline{e}} \left(\frac{\alpha_I}{\alpha_B} + \frac{\sigma_x^2}{\sigma_x^2 + \sigma_z^2} - 1\right)^{-1}.$$

- *If $\alpha_B > \alpha_I$, then generically $y_i(\gamma_i) \neq y_i^*(\gamma_i)$, the expected social investment is below the first-best level, i.e. $\mathbb{E}[y_i(\tilde{\gamma}_i)] < \mathbb{E}[y_i^*(\tilde{\gamma}_i)]$, and the optimal linear contract is defined by:*

$$\beta_i = (\alpha_B - \alpha_I) \frac{c_e}{\bar{e} - \underline{e}} \left(1 + \frac{\sigma_z^2}{\sigma_x^2}\right) \left(\frac{\sigma_\gamma^2}{\sigma_\eta^2 + \sigma_\gamma^2} \frac{\bar{\eta}^2}{\sigma_\gamma^2 + \sigma_\eta^2 + \bar{\eta}^2} + \frac{\sigma_\eta^2}{\sigma_\eta^2 + \sigma_\gamma^2}\right), \quad \beta_x = 0, \quad \beta_p = \frac{c_e}{\bar{e} - \underline{e}} \left(1 + \frac{\sigma_z^2}{\sigma_x^2}\right).$$

The two subsections that follow extend this model along two directions.

A.1.2 Heterogeneous social preferences

In this subsection, we allow the board to have a different preference with respect to each dimension i of SEP relative to profits, as measured by $\alpha_B^i \geq 0$, for $i = 1, 2$ (by contrast, in the baseline model of section A.1.1, $\alpha_B^1 = \alpha_B^2$). We also let the stock price impact of perceived SEP, as measured by $\alpha_I^i \geq 0$, potentially be different across dimensions of SEP. As argued by Starks (2023), economic agents with different values may have different priorities across areas of concern. For simplicity and tractability, in this subsection we assume $\sigma_y = 0$.

Proposition 5

(i) If $\alpha_B^i \geq \alpha_I^i$ for $i = 1, 2$, then $\beta_x = 0$, $\beta_p = \frac{c_e}{\bar{e} - \underline{e}}\left(1 + \frac{\sigma_z^2}{\sigma_x^2}\right)$, and, for $i = 1, 2$:

$$\frac{\beta_i}{\beta_p} = (\alpha_B^i - \alpha_I^i) \left(\frac{\sigma_\gamma^2}{\sigma_\eta^2 + \sigma_\gamma^2} \frac{\bar{\eta}^2}{\sigma_\gamma^2 + \sigma_\eta^2 + \bar{\eta}^2} + \frac{\sigma_\eta^2}{\sigma_\eta^2 + \sigma_\gamma^2}\right); \tag{A.3}$$

(ii) If $\alpha_B^i < \alpha_I^i$ for $i = 1, 2$, then $\beta_x > 0$;
(iii) If $\alpha_I^i = \alpha_I$ for $i = 1, 2$, and $\alpha_B^1 < \alpha_I \leq \alpha_B^2$ with $\frac{\alpha_I}{\alpha_B^1}$ sufficiently large and $\frac{\alpha_B^2}{\alpha_I}$ sufficiently close to 1, then $\beta_x > 0$, $\beta_p > 0$, $\beta_1 = 0$, and $\beta_2 > 0$.
(iv) If α_B^i is sufficiently large and $\alpha_I^i > 0$, then $\beta_x = 0$, $\beta_p = \frac{c_e}{\bar{e} - \underline{e}}\left(1 + \frac{\sigma_z^2}{\sigma_x^2}\right)$, and $\frac{\beta_i}{\beta_p}$ is as in equation (A.3).

In part (i), when the board prefers on all dimensions a higher level of social investment than the one that maximizes the stock price, there is no profits-based compensation.[35] On the contrary, in part (ii), when the board prefers on all dimensions a lower level of social investment than the one that maximizes the stock price, then it uses profits-based compensation in conjunction with stock price-based compensation. In both cases, the intuition is as in the main model.

In part (iii), the board prefers a much lower level of social investment than the one that maximizes the stock price on dimension 1, and a slightly higher level of social investment on dimension 2. Then, compensation is sensitive to profits to deter investment in dimension $i = 1$ of SEP, but it is also sensitive to ESG score $i = 2$ to encourage investment in dimension $i = 2$ of SEP. This is illustrated in Example 1 below. By contrast, this outcome is not possible in Proposition 4, where preferences for social output are homogeneous across dimensions of social output.

Example 1 *Suppose that* $\alpha_B^1 = 0$, $\alpha_I^1 = 1$, $\alpha_B^2 = 1.1$, $\alpha_I^2 = 1$, $\sigma_\eta = 1$, $\sigma_\gamma = 1$, $\sigma_x = 1$, $\sigma_z = 1$, $\bar{\eta} = 2$, $c_e/(\bar{e} - \underline{e}) = 0.1$. *Then we have* $\beta_1 = 0$, $\beta_2 = 0.03$, $\beta_x = 0.02$, *and* $\beta_p = 0.16$.

In part (iv), when the board prefers a much higher level of social investment than the one that maximizes the stock price on dimension i, then it uses SEP-based compensation on dimension i. Moreover, there is no profits-based compensation, even if the board prefers a much lower level of social investment than the one that maximizes the stock price on the other dimension (j). This is surprising because the board could use profits-based compensation to discourage social investments on dimension j ($\beta_x > 0$), and offset the effect of profits-based compensation on dimension i of SEP by incentivizing investment on this dimension with the corresponding ESG score (by further increasing β_i). However, this would lead to an (already discussed) inefficiency which is especially costly for a board that cares a lot about this dimension of SEP. Thus, contract complexity, as proxied by the different types of performance measures used, does not necessarily rise when the divergence in preferences increases. This is illustrated in Example 2 below.

[35] More generally, we show in the proof of Proposition 5 that the firm will not use profits-based compensation ($\beta_x = 0$) if a weighted sum of $\alpha_B^i - \alpha_I^i$ is positive, as opposed to $\alpha_B - \alpha_I > 0$ in the case with homogeneous preferences for social output. Letting $\beta \equiv \frac{\beta_p}{\beta_x + \beta_p}$, we show that the optimal value of β is a solution to the following problem: $\min_\beta \sum_{i=1,2} \Gamma_i \left(\alpha_B^i - \beta \alpha_I^i\right)^2$ s.t. $\beta \leq 1$ where Γ_i is a positive constant defined in the proof of Proposition 5. That is, the optimal fraction of stock price-based incentives as a proportion of total financial incentives, $\frac{\beta_p}{\beta_x + \beta_p}$, minimizes a weighted average quadratic distance between the board's preference for dimension i of SEP and investors' preference for the same dimension of SEP times $\frac{\beta_p}{\beta_x + \beta_p}$. Accordingly, we find that this ratio is optimally either equal to 1 (when the constraint $\beta_x \geq 0$ is binding), or to $\frac{\beta_p}{\beta_x + \beta_p} = \frac{\sum_{i=1,2} \Gamma_i \alpha_I^i \alpha_B^i}{\sum_{i=1,2} \Gamma_i {\alpha_I^i}^2}$. Intuitively, the optimal ratio $\frac{\beta_p}{\beta_x + \beta_p}$ is less than one when investors tend to have a stronger preference for SEP compared to the board, which results in the board giving profits-based incentives to the manager to counterbalance stock price-based incentives.

Example 2 *Suppose that* $\alpha_B^1 = 0$, $\alpha_I^1 = 1$, $\alpha_B^2 = 2$, $\alpha_I^2 = 1$, $\sigma_\eta = 1$, $\sigma_\gamma = 1$, $\sigma_x = 1$, $\sigma_z = 1$, $\bar{\eta} = 2$, $c_e/(\bar{e} - \underline{e}) = 0.1$. *Then we have* $\beta_1 = 0$, $\beta_2 = 0.17$, $\beta_x = 0$, *and* $\beta_p = 0.20$.

In Example 2, the only difference with respect to Example 1 is that the board cares even more about the second social dimension. In Example 2, a board with a strong preference for the second social dimension will not use profits-based compensation to discourage investment in the first dimension – even though it could separately encourage investment in the second dimension by further increasing β_2.

In summary, having heterogeneous preferences for social output across economic agents and across dimensions of social investments is a necessary but insufficient condition for a managerial contract to be explicitly contingent on three different types of performance measures: profits, the stock price, and some ESG scores. These results contribute to the nascent literature on the complexity of executive compensation (Murphy and Sandino (2020), Burkert et al. (2024), Albuquerque et al. (2024)).

A.1.3 Multiple ESG scores

We now extend the model with multiple dimensions of SEP performance to analyze the case with multiple ESG scores on each dimension, similar to section 6 in the main paper.

We now assume that there are N ESG raters. Each rater j provides a set of ESG scores $\{m_1^j, m_2^j\}$. ESG score on dimension i by ESG rater j is defined as $m_i^j \equiv \gamma_i^j y_i$.

We analyze the effect of increasing the number of ESG scores with uncorrelated noise terms. As in the baseline model, the distribution of ESG scores j on dimension i is centered on $\eta_i y_i$ such that $\tilde{m}_i^j \sim \mathcal{N}(\eta_i y_i, \sigma_\gamma^2)$. Denoting $\tilde{\zeta}_i^j \sim \mathcal{N}(0, \sigma_\gamma^2)$, for any score j on dimension i we can decompose the score as $\tilde{m}_i^j = y_i(\tilde{\eta}_i + \tilde{\zeta}_i^j)$, where the superscript j indicates that the noise term $\tilde{\zeta}_i^j$ is different for each rating j. We assume that the noise terms $\tilde{\zeta}_i^j$ on any given dimension are independent and identically distributed (i.i.d.), with a variance of σ_γ^2. This implies that ESG scores on dimension i are uncorrelated conditional on η_i, but they are unconditionally positively correlated because of their dependence on η_i. Let $\bar{\gamma}_i = \frac{1}{N}\sum_{j=1}^N \gamma_i^j$ be the average signal on the firm's social productivity on dimension i generated by ESG scores. This average signal is a sufficient statistic for the mean η_i of the distribution.

Proposition 6 *As the number of ESG scores is increased, social investment and expected social* *output converge asymptotically to the first-best levels of these variables:*

$$\lim_{N\to\infty} y_i(\bar{\gamma}_i) = \frac{\alpha_B}{2\theta_i}\bar{\gamma}_i \qquad \textit{and} \qquad \lim_{N\to\infty} \mathbb{E}[\tilde{\eta}_i\tilde{v}_i] = \frac{\alpha_B}{2\theta_i}\left(\bar{\eta}^2 + \sigma_\eta^2\right)\ \forall i. \tag{A.4}$$

A.2 Alternative interpretations of the model

In the main interpretation of the model, $\tilde{v}$ is the amount of "social output" (including positive externalities and reductions in negative externalities) produced by the firm at $t = 2$, and α_B and α_I are respectively the board's and investors' preference for social output relative to cash flows.

There are alternative interpretations for the specification of the board's and investors' social and environmental preferences.

The first alternative interpretation is that the board is not intrinsically socially responsible, but it uses incentive pay to commit to an investment policy that would otherwise not be in the best interests of the firm ex-post. For example, this can be useful to raise funding from socially and environmentally responsible investors at a lower cost, or to hire employees who care about these issues. To be credible, the firm must commit to be socially responsible.[36] Specifically, it should invest "as if" it were socially responsible. This can be achieved by setting a compensation contract "as if" the board had the objective function in equation (2). In this interpretation of the model, investors are socially responsible if they put a positive weight α_I on social output relative to profits. In principle, this hypothesis can explain the concomitant rise of commitments such as "sustainability pledges", and the increased reliance on social and environmental measures of firm performance in executive compensation.

The second alternative interpretation is that the board and investors are not intrinsically socially responsible, but they are aware of political and judicial pressures emanating from activists and regulators. These third parties may punish firms that generate negative externalities, for example by requesting or mandating reparations for harm caused in the past. Even though this might not affect the firm's profitability during the manager's tenure (until $t = 2$), these actions might be costly to the firm in the distant future ($t = 3$). This heightened concern can be explained by recent shifts in public opinion. According to the US Department of Justice: "criminal

[36]In some instances, when green investments are well-defined, this can alternatively be induced by raising funding via green bonds (Barbalau and Zeni (2022)). In other cases, socially responsible investments are not well-defined, i.e. they cannot be described in a contract a priori, or there is not enough information at the contracting stage to determine efficient investments.

prosecution acknowledges that environmental stewardship has become a mainstream value, such that most Americans recognize that polluting ... [is] repugnant." In 2023, the US Supreme Court allowed lawsuits by municipalities seeking to hold energy companies liable for harms caused by climate emissions to move forward.[37] This is related to the notion of "enlarged fiduciary duty" proposed by Tirole (2001), in which stakeholders could sue a firm whose actions did not "follow the mandate of the stakeholder society".

In this interpretation of the model, the random variable $\tilde{v}$ is the monetary amount in penalties imposed on the firm at $t = 3$, and α_B and α_I are respectively the board's and investors' discount factors for $t = 3$ cash flows relative to $t = 2$ cash flows. Discount factors could differ because long-term shareholders and stock market investors have different endowments and markets are incomplete (Grossman and Hart (1979)).[38] Discount factors could also differ because the firm's shareholders enjoy substantial private benefits of control (and are therefore unwilling to trade) but are more or less patient than stock market investors. For example, suppose that the firm's shareholders are more patient than the marginal stock market investor. In this case, the discount factor applied to $t = 3$ penalties relative to $t = 2$ cash flows is higher for shareholders than for the marginal stock market investor, i.e. $\alpha_B > \alpha_I$. Finally, discount factors could differ because of disagreement between the board and investors with respect to the extent of $t = 3$ penalties for social and environmental damages. In this interpretation, differences between so-called "discount factors" would reflect the different beliefs associated with $t = 3$ penalties.[39]

A.3 Proofs

Proof of Lemma 5:

Investors believe that the manager exerts some effort $\hat{e}$. With one set of ESG scores, investors update their beliefs about the firm's technology for social output after observing ESG scores as

[37]Sources: https://www.justice.gov/enrd/environmental-crime-victim-assistance/prosecution-federal-pollution-crimes and https://www.nbcnews.com/politics/supreme-court/supreme-court-rejects-oil-companies-appeals-climate-change-disputes-rcna49823

[38]Grossman and Hart (1979) note that marginal rates of substitution will then be heterogeneous across shareholders (or "investors"), i.e., each of them will have her own discount factor.

[39]It is noteworthy that disagreement across investors reduces the discount rate used for stock pricing (Yu (2011), Huang et al. (2020)). In our model, α_I is the discount factor that matters for stock pricing, i.e. disagreement across investors would increase α_I.

follows:

$$\begin{aligned}
\mathbb{E}[\tilde{x}|y,z,m_1,m_2] &= \mathbb{E}[\tilde{x}] + \frac{cov(\tilde{z},\tilde{x})}{var(\tilde{z})}\left(z - \mathbb{E}[\tilde{z}]\right) \\
&= \hat{e} - \theta_1 y_1^2 + \theta_2 y_2^2 + \frac{\sigma_x^2}{\sigma_x^2+\sigma_z^2}\left(z - \hat{e} + \theta_1 y_1^2 + \theta_2 y_2^2\right) \quad \text{(A.5)} \\
\mathbb{E}[\tilde{v}|y,z,m_1,m_2] &= \mathbb{E}\left[\tilde{\eta}_1\tilde{y}_1 + \tilde{\eta}_2\tilde{y}_2 + \tilde{\epsilon}_y|y,z,m_1,m_2\right] \\
&= y_1\mathbb{E}\left[\tilde{\eta}_1|m_1,y_1\right] + y_2\mathbb{E}\left[\tilde{\eta}_2|m_2,y_2\right] \quad \text{(A.6)}
\end{aligned}$$

where:

$$\mathbb{E}\left[\tilde{\eta}_i|m_i,y_i\right] = \mathbb{E}\left[\tilde{\eta}_i|\gamma_i\right] = \mathbb{E}\left[\tilde{\eta}_i\right] + \frac{cov\left(\tilde{\eta}_i,\tilde{\gamma}_i\right)}{var\left(\tilde{\gamma}_i\right)}\left(\gamma_i - \mathbb{E}[\tilde{\gamma}_i]\right) = \bar{\eta} + \frac{\sigma_\eta^2}{\sigma_\eta^2+\sigma_\gamma^2}\left(\gamma_i - \bar{\eta}\right) \quad \text{(A.7)}$$

Substituting into equation (A.6):

$$\begin{aligned}
\mathbb{E}[\tilde{v}|y,z,m_1,m_2] &= y_1\left(\bar{\eta} + \frac{\sigma_\eta^2}{\sigma_\eta^2+\sigma_\gamma^2}\left(\gamma_1-\bar{\eta}\right)\right) + y_2\left(\bar{\eta} + \frac{\sigma_\eta^2}{\sigma_\eta^2+\sigma_\gamma^2}\left(\gamma_2-\bar{\eta}\right)\right) \\
&= \frac{\sigma_\eta^2}{\sigma_\eta^2+\sigma_\gamma^2}m_1 + \frac{\sigma_\gamma^2}{\sigma_\eta^2+\sigma_\gamma^2}y_1\bar{\eta} + \frac{\sigma_\eta^2}{\sigma_\eta^2+\sigma_\gamma^2}m_2 + \frac{\sigma_\gamma^2}{\sigma_\eta^2+\sigma_\gamma^2}y_2\bar{\eta} \quad \text{(A.8)}
\end{aligned}$$

■

Proof of Proposition 5:

The objective function of the board now rewrites as:

$$\mathbb{E}[V(y,e)] = \mathbb{E}\left[\tilde{x} + \sum_{i=1,2}\alpha_B^i\tilde{\eta}_i\tilde{y}_i - \left(w + \beta_x\tilde{x} + \beta_p\tilde{p} + \sum_{i=1,2}\beta_i\tilde{m}_i\right)\right] \quad \text{(A.9)}$$

The first part of the proof involves straightforward notational adjustments to the proof of Proposition 1. Using the equivalent of equation (41), the loss function rewrites as:

$$\begin{aligned}
\min_{\beta_x,\beta_p,\beta_i}\frac{1}{\theta_i}\sum_{i=1,2}\int_{\eta_i}\int_{\gamma_i}\Bigg(&\left(\frac{\beta_i}{\beta_x+\beta_p} + \left(\frac{\beta_p\alpha_I^i}{\beta_x+\beta_p} - \alpha_B^i\right)\frac{\sigma_\eta^2}{\sigma_\eta^2+\sigma_\gamma^2}\right)\gamma_i \\
&+ \left(\frac{\beta_p\alpha_I^i}{\beta_x+\beta_p} - \alpha_B^i\right)\frac{\sigma_\gamma^2}{\sigma_\eta^2+\sigma_\gamma^2}\bar{\eta}\Bigg)^2\varphi(\gamma_i|\eta_i)\phi(\eta_i)d\gamma_i d\eta_i \quad \text{(A.10)}
\end{aligned}$$

This expression is globally convex with respect to β_i. As above, the FOC w.r.t. β_i is:

$$\int_{\eta_i}\int_{\gamma_i} \frac{2}{\beta_x+\beta_p}\gamma_i\Bigg(\left(\frac{\beta_i}{\beta_x+\beta_p}+\left(\frac{\beta_p\alpha_I^i}{\beta_x+\beta_p}-\alpha_B^i\right)\frac{\sigma_\eta^2}{\sigma_\eta^2+\sigma_\gamma^2}\right)\gamma_i$$
$$+\left(\frac{\beta_p\alpha_I^i}{\beta_x+\beta_p}-\alpha_B^i\right)\frac{\sigma_\gamma^2}{\sigma_\eta^2+\sigma_\gamma^2}\bar{\eta}\Bigg)\varphi(\gamma_i|\eta_i)\phi(\eta_i)d\gamma_i d\eta_i = 0$$
$$\Leftrightarrow \quad \frac{\beta_i}{\beta_x+\beta_p} = \left(\frac{\sigma_\gamma^2}{\sigma_\eta^2+\sigma_\gamma^2}\frac{\bar{\eta}^2}{\sigma_\gamma^2+\sigma_\eta^2+\bar{\eta}^2}+\frac{\sigma_\eta^2}{\sigma_\eta^2+\sigma_\gamma^2}\right)\left(\alpha_B^i-\frac{\beta_p\alpha_I^i}{\beta_x+\beta_p}\right) \tag{A.11}$$

Given the nonnegativity constraint and the global convexity of the objective function with respect to β_i, there are two cases. If equation (A.11) gives a positive β_i given the optimal values of β_x and β_p derived below, then the optimum is $\beta_i^* = \beta_i$ as in equation (A.11). If equation (A.11) gives a nonpositive β_i, then the optimum is $\beta_i^* = 0$.

Define:

$$\lambda_i \equiv \int_{\eta_i}\int_{\gamma_i}\left(\frac{\sigma_\eta^2}{\sigma_\eta^2+\sigma_\gamma^2}\gamma_i+\frac{\sigma_\gamma^2}{\sigma_\eta^2+\sigma_\gamma^2}\bar{\eta}\right)^2\varphi(\gamma_i|\eta_i)\phi(\eta_i)d\gamma_i d\eta_i \tag{A.12}$$
$$g_i \equiv \frac{\sigma_\gamma^4}{(\sigma_\eta^2+\sigma_\gamma^2)^2}\int_{\eta_i}\int_{\gamma_i}\left(\frac{\bar{\eta}^2}{\sigma_\gamma^2+\sigma_\eta^2+\bar{\eta}^2}\gamma_i-\bar{\eta}\right)^2\varphi(\gamma_i|\eta_i)\phi(\eta_i)d\gamma_i d\eta_i, \tag{A.13}$$

Both λ_i and g_i are strictly positive. Since the distribution of $\tilde{\eta}_i$ is independent from i, and the distribution of $\tilde{\gamma}_i$ conditional on η_i is independent from i, the variables λ_i and g_i are independent from i and independent from the contract. Let $\Gamma_i = \lambda_i$ if the nonnegativity constraint on β_i is binding, and $\Gamma_i = g_i$ if the nonnegativity constraint on β_i is nonbinding.

There are three cases.

Case 1. We conjecture that $\beta_i > 0$ for $i = 1, 2$, so that $\Gamma_i = g_i$ for $i = 1, 2$. Substituting for β_i in equation (A.10), the optimization problem is:

$$\min_{\beta_x,\beta_p}\sum_{i=1,2}\Gamma_i\left(\frac{\beta_p}{\beta_x+\beta_p}\alpha_I^i-\alpha_B^i\right)^2 \tag{A.14}$$

where $\Gamma_i = g_i$ in case 1 (we henceforth keep the Γ_i notation because we will refer to equations below for cases other than case 1). This is a sum weighted by Γ_i of the quadratic distance between the board's preference for dimension i of SEP, and a fraction $\frac{\beta_p}{\beta_x+\beta_p}$ of investors' preference for same dimension of SEP. For a given β_p, this is equivalent to choosing:

$$\min_{\beta} \sum_{i=1,2} \Gamma_i \left(\alpha_B^i - \beta \alpha_I^i\right)^2 \quad \text{s.t.} \quad \beta \leq 1 \tag{A.15}$$

Denote by $\delta \geq 0$ the Lagrange multiplier associated with the constraint. The Lagrangian is:

$$\mathcal{L} = \sum_{i=1,2} \Gamma_i \left(\alpha_B^i - \beta \alpha_I^i\right)^2 + \delta(\beta - 1) \tag{A.16}$$

The FOC with respect to β is:

$$-2 \sum_{i=1,2} \Gamma_i \alpha_I^i \left(\alpha_B^i - \beta \alpha_I^i\right) + \delta = 0 \quad \Leftrightarrow \quad \beta = \frac{\sum_{i=1,2} \Gamma_i \alpha_I^i \alpha_B^i - \delta}{\sum_{i=1,2} \Gamma_i {\alpha_I^i}^2} \tag{A.17}$$

where $\delta > 0$ if and only if:

$$\sum_{i=1,2} \Gamma_i \alpha_I^i \left(\alpha_B^i - \alpha_I^i\right) > 0 \tag{A.18}$$

For $\delta > 0$, because of the complementary slackness condition we have $\beta = 1 \;\Leftrightarrow\; \frac{\beta_p}{\beta_x + \beta_p} = 1 \;\Leftrightarrow\; \beta_x = 0$. The value of β_p is determined according to incentive compatibility with respect to the manager's effort. Substituting in equation (30):

$$\beta_p \frac{\sigma_x^2}{\sigma_x^2 + \sigma_z^2} = \frac{c_e}{\overline{e} - \underline{e}} \quad \Leftrightarrow \quad \beta_p = \frac{c_e}{\overline{e} - \underline{e}} \left(1 + \frac{\sigma_z^2}{\sigma_x^2}\right) \tag{A.19}$$

For $\delta = 0$, the FOC rewrites as:

$$\sum_{i=1,2} \Gamma_i \alpha_I^i \left(\alpha_B^i - \beta \alpha_I^i\right) = 0 \quad \Leftrightarrow \quad \beta = \frac{\sum_{i=1,2} \Gamma_i \alpha_I^i \alpha_B^i}{\sum_{i=1,2} \Gamma_i {\alpha_I^i}^2} \quad \Leftrightarrow \quad \beta_x = \beta_p \left(\frac{\sum_{i=1,2} \Gamma_i {\alpha_I^i}^2}{\sum_{i=1,2} \Gamma_i \alpha_I^i \alpha_B^i} - 1\right) \tag{A.20}$$

This equation gives the optimal value of β_x as long as this value is nonnegative, so that the nonnegativity constraint is nonbinding (i.e. $\delta = 0$). The value of β_p is determined according to incentive compatibility with respect to the manager's effort. Substituting in equation (30):

$$\beta_p\left(\frac{\sum_{i=1,2}\Gamma_i{\alpha_I^i}^2}{\sum_{i=1,2}\Gamma_i\alpha_I^i\alpha_B^i}-1\right)+\beta_p\frac{\sigma_x^2}{\sigma_x^2+\sigma_z^2}=\frac{c_e}{\overline{e}-\underline{e}}$$

$$\Leftrightarrow\quad \beta_p=\frac{c_e}{\overline{e}-\underline{e}}\Bigg/\left(\frac{\sum_{i=1,2}\Gamma_i{\alpha_I^i}^2}{\sum_{i=1,2}\Gamma_i\alpha_I^i\alpha_B^i}-1+\frac{\sigma_x^2}{\sigma_x^2+\sigma_z^2}\right) \tag{A.21}$$

Finally, we verify the conjecture that $\beta_i > 0$ for $i = 1, 2$, by plugging β_x and β_p thus derived into equation (A.11). If the conjecture is verified, the algorithm stops. Otherwise we move to case 2 below.

Case 2. We conjecture that $\beta_1 > 0$ and $\beta_2 = 0$ if $\frac{\alpha_B^1}{\alpha_I^1} > \frac{\alpha_B^2}{\alpha_I^2}$, in which case $\Gamma_1 = g_1$ and $\Gamma_2 = \lambda_2$, and $\beta_1 = 0$ and $\beta_2 > 0$ otherwise, in which case $\Gamma_1 = \lambda_1$ and $\Gamma_2 = g_2$. The proof follows the same steps as in case 1 above except for the different values of Γ_i. If the conjecture is verified, the algorithm stops. Otherwise we move to case 3 below.

Case 3. We conjecture that $\beta_i = 0$ for $i = 1, 2$, so that $\Gamma_i = \lambda_i$ for $i = 1, 2$. The proof follows the same steps as in case 1 above except for the different values of Γ_i.

We now rely on this algorithm to establish the four points of Proposition 5.

(i) If $\alpha_B^i \geq \alpha_I^i$ for $i = 1, 2$, then β_i is as in equation (A.11) such that $\beta_i \geq 0$ for $i = 1, 2$. As a result, we have $\Gamma_i = g_i$ for $i = 1, 2$ and case 1 as described above is relevant. Moreover, with $\alpha_B^i \geq \alpha_I^i$ for $i = 1, 2$, the inequality in equation (A.18) holds, so that $\beta_x = 0$, and β_p is as in equation (A.19).

(ii) If $\alpha_B^i < \alpha_I^i$ for $i = 1, 2$, the inequality in equation (A.18) does not hold, so that $\beta_x > 0$, and β_p is as in equation (A.21). Given these values of β_x and β_p, β_i is as in equation (A.11) if the expression is positive, and zero otherwise.

(iii) If $\alpha_I^i = \alpha_I$ for $i = 1, 2$, $\alpha_B^1 < \alpha_I \leq \alpha_B^2$ with $\frac{\alpha_I}{\alpha_B^1}$ sufficiently large and $\frac{\alpha_B^2}{\alpha_I}$ sufficiently close to 1, the inequality in equation (A.18) does not hold:

$$\sum_{k=1,2}\Gamma_k\alpha_I^k\left(\alpha_B^k-\alpha_I^k\right)=\alpha_I\sum_{k=1,2}\Gamma_k\left(\alpha_B^k-\alpha_I\right)<0,$$

so that $\beta_x > 0$, and β_p is as in equation (A.21). With $\alpha_B^1 < \alpha_I \leq \alpha_B^2$ such that $\frac{\alpha_I}{\alpha_B^1}$ is sufficiently large and $\frac{\alpha_B^2}{\alpha_I}$ is sufficiently close to 1, we have $\frac{\sum_{i=1,2}\Gamma_i{\alpha_I^i}^2}{\sum_{i=1,2}\Gamma_i\alpha_I^i\alpha_B^i} > 1$, i.e. $\beta_p > 0$. With $\alpha_B^2 \geq \alpha_I$ and

$\beta_x > 0$, which implies $\beta < 1$, we have $\beta_2 > 0$. Using equation (A.20):

$$\alpha_B^i - \frac{\beta_p}{\beta_x + \beta_p}\alpha_I^i = \alpha_B^i - \beta\alpha_I = \alpha_B^i - \frac{\sum_{k=1,2}\Gamma_k \alpha_I^2 \alpha_B^k}{\sum_{k=1,2}\Gamma_k \alpha_I^2} = \alpha_B^i - \frac{\sum_{k=1,2}\Gamma_k \alpha_B^k}{\sum_{k=1,2}\Gamma_k} \tag{A.22}$$

By contradiction, suppose that $\beta_1 > 0$ so that $\Gamma_1 = \Gamma_2 \equiv \Gamma$. Then, substituting in equation (A.22):

$$\alpha_B^1 - \frac{\beta_p}{\beta_x + \beta_p}\alpha_I^1 = \alpha_B^1 - \frac{\sum_{k=1,2}\Gamma \alpha_B^k}{\sum_{k=1,2}\Gamma} = \alpha_B^1 - \frac{\alpha_B^1 + \alpha_B^2}{2} = \frac{\alpha_B^1 - \alpha_B^2}{2} < 0,$$

so that from equation (A.11) we would have $\beta_1 < 0$, a contradiction. Thus, $\beta_1 = 0$.

(iv) If α_B^i is sufficiently large and $\alpha_I^i > 0$, then β_i is as in equation (A.11) such that $\beta_i > 0$ even without the nonnegativity constraint, i.e. this constraint does not bind for β_i, and we have $\Gamma_i = g_i$. Moreover, when α_B^i is sufficiently large and $\alpha_I^i > 0$, the inequality in equation (A.18) holds:

$$\sum_{k=1,2}\Gamma_k \alpha_I^k \left(\alpha_B^k - \alpha_I^k\right) > g_i \alpha_I^i \left(\alpha_B^i - \alpha_I^i\right) - \Gamma_j {\alpha_I^j}^2 > 0$$

Thus, $\beta_x = 0$, and β_p is as in equation (A.19). ■

Proof of Proposition 6:

With N ESG scores on each ESG dimension i whose noise terms are i.i.d., the average score is a sufficient statistic for the mean (this is a standard application of the factorization theorem). Define:

$$\bar{m}_i \equiv \frac{1}{N}\sum_{j=1}^{N} m_i^j \tag{A.23}$$

We have:

$$\frac{1}{N}\sum_{j=1}^{N}\tilde{m}_i^j = \frac{1}{N}\sum_{j=1}^{N} y_i\left(\eta_i + \tilde{\zeta}_i^j\right) = \eta_i y_i + \frac{y_i}{N}\sum_{j=1}^{N}\tilde{\zeta}_i^j \tag{A.24}$$

where the random variable $\frac{1}{N}\sum_{j=1}^{N}\tilde{\zeta}_i^j$ is normally distributed with the following mean and variance:

$$\mathbb{E}\left[\frac{1}{N}\sum_{j=1}^{N}\tilde{\zeta}_i^j\right] = \frac{1}{N}\sum_{j=1}^{N}\mathbb{E}\left[\tilde{\zeta}_i^j\right] = 0 \tag{A.25}$$

$$var\left(\frac{1}{N}\sum_{j=1}^{N}\tilde{\zeta}_i^j\right) = \frac{1}{N^2}\sum_{j=1}^{N}var\left(\tilde{\zeta}_i^j\right) = \frac{1}{N}\sigma_\gamma^2 \tag{A.26}$$

where we used the i.i.d. assumption about the noise terms. Therefore, the unconditional variance of $\tilde{\bar{m}}_i/y_i$ is:

$$var\left(\tilde{\bar{m}}_i/y_i\right) = \sigma_\eta + \frac{1}{N}\sigma_\gamma^2 \tag{A.27}$$

Belief updating about the firm's technology for social output after observing ESG scores is as follows:

$$\begin{aligned}\mathbb{E}[\tilde{x}|y,z,\{m_i^j\}] &= \mathbb{E}[\tilde{x}] + \frac{cov(\tilde{z},\tilde{x})}{var(\tilde{z})}\left(z-\mathbb{E}[\tilde{z}]\right)\\ &= \hat{e}-\theta_1 y_1^2+\theta_2 y_2^2+\frac{\sigma_x^2}{\sigma_x^2+\sigma_z^2}\left(z-\hat{e}+\theta_1 y_1^2+\theta_2 y_2^2\right) \qquad \text{(A.28)}\\ \mathbb{E}[\tilde{v}|y,z,\{m_i^j\}] &= \mathbb{E}\left[\tilde{\eta}_1\tilde{y}_1+\tilde{\eta}_2\tilde{y}_2+\tilde{\epsilon}_y|z,\bar{m}\right] = y_1\mathbb{E}\left[\tilde{\eta}_1|\bar{m}_1,y_1\right]+y_2\mathbb{E}\left[\tilde{\eta}_2|\bar{m}_2,y_2\right] \qquad \text{(A.29)}\end{aligned}$$

Moreover, for any $i\in\{1,2\}$:

$$\mathbb{E}\left[\tilde{\eta}_i|\bar{m}_i\right]=\mathbb{E}\left[\tilde{\eta}_i\right]+\frac{cov\left(\tilde{\eta}_i,\tilde{\bar{m}}_i/y_i\right)}{var\left(\tilde{\bar{m}}_i/y_i\right)}\left(\bar{m}_i/y_i-\mathbb{E}[\tilde{\bar{m}}_i/y_i]\right)=\bar{\eta}+\frac{\sigma_\eta^2}{\sigma_\eta+\sigma_\gamma^2/N}\left(\bar{m}_i/y_i-\bar{\eta}\right) \tag{A.30}$$

Substituting into equation (A.29):

$$\begin{aligned}\mathbb{E}[\tilde{v}|y,z,\{m_i^j\}] &= y_1\left(\bar{\eta}+\frac{\sigma_\eta^2}{\sigma_\eta^2+\sigma_\gamma^2/N}\left(\bar{m}_1/y_1-\bar{\eta}\right)\right)+y_2\left(\bar{\eta}+\frac{\sigma_\eta^2}{\sigma_\eta^2+\sigma_\gamma^2/N}\left(\bar{m}_2/y_2-\bar{\eta}\right)\right)\\ &= \frac{\sigma_\eta^2}{\sigma_\eta^2+\sigma_\gamma^2/N}\bar{m}_1+\frac{\sigma_\gamma^2/N}{\sigma_\eta^2+\sigma_\gamma^2/N}y_1\bar{\eta}+\frac{\sigma_\eta^2}{\sigma_\eta^2+\sigma_\gamma^2/N}\bar{m}_2+\frac{\sigma_\gamma^2/N}{\sigma_\eta^2+\sigma_\gamma^2/N}y_2\bar{\eta} \qquad \text{(A.31)}\end{aligned}$$

The stock price p is set accordingly.

Since the average ESG score is a sufficient statistic, considering contracts based on the average score $\bar{m}_i$ on each dimension i of SEP rather than on each individual score is WLOG. The average score is defined as: $\bar{m}_i = y_i\bar{\gamma}_i$ with $\bar{\gamma}_i = \frac{1}{N}\sum_{j=1}^{N}\gamma_i^j$. Rewriting the manager's objective function accordingly gives:

$$\max_{e,y_1,y_2} \mathbb{E}\left[w + \beta_x \left(e - \theta_1 y_1^2 - \theta_2 y_2^2\right) + \beta_p \tilde{p} + \sum_{i=1,2} \beta_i \bar{m}_i \middle| \bar{\gamma}_1, \bar{\gamma}_2 \right] - C(e) \tag{A.32}$$

For $\beta_x + \beta_p > 0$, the manager's objective function is concave in y_i, for $i = 1, 2$. The FOC with respect to y_i is:

$$\beta_x \left(-2\theta_i y_i\right) + \beta_p \left(-2\theta_i y_i + \alpha_I \left(\frac{\sigma_\eta^2}{\sigma_\eta^2 + \sigma_\gamma^2/N} \bar{\gamma}_i + \frac{\sigma_\gamma^2/N}{\sigma_\eta^2 + \sigma_\gamma^2/N} \bar{\eta} \right) \right) + \beta_i \bar{\gamma}_i = 0$$

$$\begin{aligned} \Leftrightarrow \quad y_i(\bar{\gamma}_i) &= \frac{\beta_i \bar{\gamma}_i + \beta_p \alpha_I \left(\frac{\sigma_\eta^2}{\sigma_\eta^2+\sigma_\gamma^2/N} \bar{\gamma}_i + \frac{\sigma_\gamma^2/N}{\sigma_\eta^2+\sigma_\gamma^2/N} \bar{\eta} \right)}{\beta_x + \beta_p} \frac{1}{2\theta_i} \\ &= \frac{\left(\beta_i + \beta_p \alpha_I \frac{\sigma_\eta^2}{\sigma_\eta^2+\sigma_\gamma^2/N} \right) \bar{\gamma}_i + \beta_p \alpha_I \frac{\sigma_\gamma^2/N}{\sigma_\eta^2+\sigma_\gamma^2/N} \bar{\eta}}{\beta_x + \beta_p} \frac{1}{2\theta_i} \end{aligned} \tag{A.33}$$

Given his contract, the manager will optimally exert high effort ($e = \bar{e}$) if and only if the inequality in equation (30) holds. The fixed component of pay, w, is set to guarantee the manager's participation given $e = \bar{e}$:

$$\begin{aligned} \mathbb{E}[u(y,e)] &= \bar{W} \\ \Leftrightarrow \quad w + \beta_x \mathbb{E}\left[e - \theta_1 y_1^2 - \theta_2 y_2^2 + \tilde{\epsilon}_x\right] + \beta_p \mathbb{E}[\tilde{p}] + \sum_{i=1,2} \beta_i \mathbb{E}[\tilde{\bar{m}}_i] &= \bar{W} + C(e) \end{aligned} \tag{A.34}$$

Thus, the board's objective function can be rewritten as:

$$\begin{aligned} \mathbb{E}[V(y,e)] &= \mathbb{E}\left[(1-\beta_x)\left(e - \theta_1 y_1^2 - \theta_2 y_2^2 + \tilde{\epsilon}_x\right) + \alpha_B \left(\tilde{\eta}_1 y_1 + \tilde{\eta}_2 y_2 + \tilde{\epsilon}_y\right) - \left(w + \beta_p \tilde{p} + \sum_{i=1,2} \beta_i \tilde{\bar{m}}_i \right) \right] \\ &= \mathbb{E}\left[e - \theta_1 y_1^2 - \theta_2 y_2^2 + \tilde{\epsilon}_x + \alpha_B \left(\tilde{\eta}_1 y_1 + \tilde{\eta}_2 y_2 + \tilde{\epsilon}_y\right) - \left(\bar{W} + C(e)\right) \right] \end{aligned} \tag{A.35}$$

Following the same steps as in the proof of Lemma 3, maximizing the board's objective function is equivalent to minimizing the expected quadratic distance between $y_i(\bar{\gamma}_i)$ and $y_i^*(\bar{\gamma}_i)$ for $i = 1, 2$, where:

$$\mathbb{E}\left[\tilde{\eta}_i | \bar{\gamma}_i\right] = \frac{\sigma_\eta^2}{\sigma_\eta^2 + \sigma_\gamma^2/N} \bar{\gamma}_i + \frac{\sigma_\gamma^2/N}{\sigma_\eta^2 + \sigma_\gamma^2/N} \bar{\eta} \tag{A.36}$$

$$y_i^*(\bar{\gamma}_i) = \frac{\alpha_B}{2} \frac{1}{\theta_i} \left(\frac{\sigma_\eta^2}{\sigma_\eta^2 + \sigma_\gamma^2/N} \bar{\gamma}_i + \frac{\sigma_\gamma^2/N}{\sigma_\eta^2 + \sigma_\gamma^2/N} \bar{\eta} \right) \tag{A.37}$$

This problem is:

$$
\begin{aligned}
&\min_{\beta_x,\beta_p,\beta_i} \sum_{i=1,2} \int_{\eta_i}\int_{\bar{\gamma}_i} \left(\frac{1}{2}\frac{1}{\theta_i} \frac{\left(\beta_i + \beta_p\alpha_I \frac{\sigma_\eta^2}{\sigma_\eta^2+\sigma_\gamma^2/N}\right)\bar{\gamma}_i + \beta_p\alpha_I \frac{\sigma_\gamma^2/N}{\sigma_\eta^2+\sigma_\gamma^2/N}\bar{\eta}}{\beta_x+\beta_p} \right.\\
&\left. -\frac{\alpha_B}{2}\frac{1}{\theta_i}\left(\frac{\sigma_\eta^2}{\sigma_\eta^2+\sigma_\gamma^2/N}\bar{\gamma}_i + \frac{\sigma_\gamma^2/N}{\sigma_\eta^2+\sigma_\gamma^2/N}\bar{\eta}\right)\right)^2 \varphi(\bar{\gamma}_i|\eta_i)\phi(\eta_i)d\bar{\gamma}_i d\eta_i \\
\Leftrightarrow \quad &\min_{\beta_x,\beta_p,\beta_i} \sum_{i=1,2} \int_{\eta_i}\int_{\bar{\gamma}_i} \left(\left(\frac{\beta_i}{\beta_x+\beta_p} + \left(\frac{\beta_p\alpha_I}{\beta_x+\beta_p} - \alpha_B\right)\frac{\sigma_\eta^2}{\sigma_\eta^2+\sigma_\gamma^2/N}\right)\bar{\gamma}_i \right.\\
&\left. + \left(\frac{\beta_p\alpha_I}{\beta_x+\beta_p} - \alpha_B\right)\frac{\sigma_\gamma^2/N}{\sigma_\eta^2+\sigma_\gamma^2/N}\bar{\eta}\right)^2 \varphi(\bar{\gamma}_i|\eta_i)\phi(\eta_i)d\bar{\gamma}_i d\eta_i \qquad \text{(A.38)}
\end{aligned}
$$

We have:

$$
\int_{\eta_i}\int_{\bar{\gamma}_i} \bar{\gamma}_i\varphi(\bar{\gamma}_i|\eta_i)\phi(\eta_i)d\bar{\gamma}_i d\eta_i = \int_{\eta_i}\left(\int_{\bar{\gamma}_i}\bar{\gamma}_i\varphi(\bar{\gamma}_i|\eta_i)d\bar{\gamma}_i\right)\phi(\eta_i)d\eta_i = \int_{\eta_i}\eta_i\phi(\eta_i)d\eta_i = \bar{\eta}
$$

Moreover, $\int_{\bar{\gamma}_i}\bar{\gamma}_i^2\varphi(\bar{\gamma}_i|\eta_i)d\bar{\gamma}_i = \mathbb{E}[\tilde{\bar{\gamma}}_i^2|\eta_i] = var(\tilde{\bar{\gamma}}_i|\eta_i) + (\mathbb{E}[\tilde{\bar{\gamma}}_i|\eta_i])^2 = \sigma_\gamma^2/N + \eta_i^2$. This implies:

$$
\int_{\eta_i}\int_{\bar{\gamma}_i} \bar{\gamma}_i^2\varphi(\bar{\gamma}_i|\eta_i)\phi(\eta_i)d\bar{\gamma}_i d\eta_i = \int_{\eta_i}\left(\sigma_\gamma^2/N + \eta_i^2\right)\phi(\eta_i)d\eta_i = \sigma_\gamma^2/N + \int_{\eta_i}\eta_i^2\phi(\eta_i)d\eta_i = \sigma_\gamma^2/N + \sigma_\eta^2 + \bar{\eta}^2
$$

The expression in equation (A.38) is globally convex with respect to β_i. The FOC w.r.t. β_i is:

$$
\begin{aligned}
&\left(\frac{\beta_i}{\beta_x+\beta_p} + \left(\frac{\beta_p\alpha_I}{\beta_x+\beta_p} - \alpha_B\right)\frac{\sigma_\eta^2}{\sigma_\eta^2+\sigma_\gamma^2/N}\right)\int_{\eta_i}\int_{\bar{\gamma}_i}\bar{\gamma}_i^2\varphi(\bar{\gamma}_i|\eta_i)\phi(\eta_i)d\bar{\gamma}_i d\eta_i \\
&= \left(\alpha_B - \frac{\beta_p\alpha_I}{\beta_x+\beta_p}\right)\frac{\sigma_\gamma^2/N}{\sigma_\eta^2+\sigma_\gamma^2/N}\bar{\eta}\int_{\eta_i}\int_{\bar{\gamma}_i}\bar{\gamma}_i\varphi(\bar{\gamma}_i|\eta_i)\phi(\eta_i)d\bar{\gamma}_i d\eta_i \\
\Leftrightarrow \quad &\frac{\beta_i}{\beta_x+\beta_p} = \left(\alpha_B - \frac{\beta_p\alpha_I}{\beta_x+\beta_p}\right)\left(\frac{\sigma_\gamma^2/N}{\sigma_\eta^2+\sigma_\gamma^2/N}\frac{\bar{\eta}^2}{\sigma_\gamma^2/N+\sigma_\eta^2+\bar{\eta}^2} + \frac{\sigma_\eta^2}{\sigma_\eta^2+\sigma_\gamma^2/N}\right) \qquad \text{(A.39)}
\end{aligned}
$$

Substituting in equation (A.38) gives:

$$
\min_{\beta_x,\beta_p}\left(\alpha_B - \frac{\beta_p\alpha_I}{\beta_x+\beta_p}\right)^2\frac{\sigma_\gamma^4}{(\sigma_\eta^2+\sigma_\gamma^2)^2/N}\sum_{i=1,2}\int_{\eta_i}\int_{\gamma_i}\left(\frac{\bar{\eta}^2}{\sigma_\gamma^2/N+\sigma_\eta^2+\bar{\eta}^2}\gamma_i - \bar{\eta}\right)^2\varphi(\gamma_i|\eta_i)\phi(\eta_i)d\gamma_i d\eta_i \qquad \text{(A.40)}
$$

The expression under the integral sign is positive and independent from the contract. The expression in equation (A.40) is minimized by minimizing $\left(\alpha_B - \frac{\beta_p\alpha_I}{\beta_x+\beta_p}\right)^2$. With a nonnegativity constraint on β_x, this is achieved by setting $\beta_x = \max\{\frac{\beta_p\alpha_I}{\alpha_B} - \beta_p, 0\}$. As above, for $\alpha_B \leq \alpha_I$, we have $\beta_p > 0$, $\beta_x \geq 0$, β_i as defined as in equation (A.39) is equal to zero, and investment in dimension i of SEP is:

$$y_i(\bar{\gamma}_i) = \frac{\beta_p \alpha_I \left(\frac{\sigma_\eta^2}{\sigma_\eta^2 + \sigma_\gamma^2/N} \tilde{\bar{\gamma}}_i + \frac{\sigma_\gamma^2/N}{\sigma_\eta^2 + \sigma_\gamma^2/N} \bar{\eta} \right)}{\beta_x + \beta_p} \frac{1}{2\theta_i} = \frac{\alpha_B}{2} \frac{1}{\theta_i} \left(\frac{\sigma_\eta^2}{\sigma_\eta^2 + \sigma_\gamma^2/N} \tilde{\bar{\gamma}}_i + \frac{\sigma_\gamma^2/N}{\sigma_\eta^2 + \sigma_\gamma^2/N} \bar{\eta} \right),$$

which is the same as $y_i^*(\bar{\gamma}_i)$ as defined in equation (A.37). Moreover, expected social output when $\alpha_B \leq \alpha_I$ is:

$$\mathbb{E}[\tilde{\eta}_i \tilde{y}_i] = \frac{\beta_p \alpha_I \left(\frac{\sigma_\eta^2}{\sigma_\eta^2 + \sigma_\gamma^2/N} \mathbb{E}[\tilde{\eta}_i \tilde{\bar{\gamma}}_i] + \frac{\sigma_\gamma^2/N}{\sigma_\eta^2 + \sigma_\gamma^2/N} \bar{\eta} \mathbb{E}[\tilde{\eta}_i] \right)}{\beta_x + \beta_p} \frac{1}{2\theta_i} = \frac{\alpha_B}{2\theta_i} \left(\frac{\sigma_\eta^4}{\sigma_\eta^2 + \sigma_\gamma^2/N} + \bar{\eta}^2 \right) \tag{A.41}$$

In the limit:

$$\lim_{N \to \infty} \mathbb{E}[\tilde{\eta}_i \tilde{y}_i] = \frac{\alpha_B}{2\theta_i} \left(\bar{\eta}^2 + \sigma_\eta^2 \right)$$

For $\alpha_B > \alpha_I$, we have $\beta_x = 0$, and to elicit high effort, use equation (30) to set:

$$\beta_p \frac{\sigma_x^2}{\sigma_x^2 + \sigma_z^2} = \frac{c_e}{\bar{e} - \underline{e}} \quad \Leftrightarrow \quad \beta_p = \frac{c_e}{\bar{e} - \underline{e}} \left(1 + \frac{\sigma_z^2}{\sigma_x^2} \right) \tag{A.42}$$

In this case, $\beta_i > 0$ and is defined as in equation (A.39). Substituting for β_x and β_p in equation (A.39), in this case we have:

$$\beta_i = (\alpha_B - \alpha_I) \frac{c_e}{\bar{e} - \underline{e}} \left(1 + \frac{\sigma_z^2}{\sigma_x^2} \right) \left(\frac{\sigma_\gamma^2/N}{\sigma_\eta^2 + \sigma_\gamma^2/N} \frac{\bar{\eta}^2}{\sigma_\gamma^2/N + \sigma_\eta^2 + \bar{\eta}^2} + \frac{\sigma_\eta^2}{\sigma_\eta^2 + \sigma_\gamma^2/N} \right) \tag{A.43}$$

In the case $\alpha_B > \alpha_I$, substituting for β_i in equation (A.33):

$$y_i(\bar{\gamma}_i) = \frac{1}{2\theta_i} \frac{(\alpha_B - \alpha_I) \frac{c_e}{\bar{e} - \underline{e}} \left(1 + \frac{\sigma_z^2}{\sigma_x^2} \right) \left(\frac{\sigma_\gamma^2/N}{\sigma_\eta^2 + \sigma_\gamma^2/N} \frac{\bar{\eta}^2}{\sigma_\gamma^2/N + \sigma_\eta^2 + \bar{\eta}^2} + \frac{\sigma_\eta^2}{\sigma_\eta^2 + \sigma_\gamma^2/N} \right) \bar{\gamma}_i + \beta_p \alpha_I \left(\frac{\sigma_\eta^2}{\sigma_\eta^2 + \sigma_\gamma^2/N} \bar{\gamma}_i + \frac{\sigma_\gamma^2/N}{\sigma_\eta^2 + \sigma_\gamma^2/N} \bar{\eta} \right)}{\beta_x + \beta_p}$$

Substituting for β_x and β_p, investment in dimension i of SEP is:

$$\begin{aligned} y_i(\bar{\gamma}_i) &= \frac{1}{2\theta_i} \Bigg((\alpha_B - \alpha_I) \left(\frac{\sigma_\gamma^2/N}{\sigma_\eta^2 + \sigma_\gamma^2/N} \frac{\bar{\eta}^2}{\sigma_\gamma^2/N + \sigma_\eta^2 + \bar{\eta}^2} + \frac{\sigma_\eta^2}{\sigma_\eta^2 + \sigma_\gamma^2/N} \right) \bar{\gamma}_i \\ &\quad + \alpha_I \left(\frac{\sigma_\eta^2}{\sigma_\eta^2 + \sigma_\gamma^2/N} \bar{\gamma}_i + \frac{\sigma_\gamma^2/N}{\sigma_\eta^2 + \sigma_\gamma^2/N} \bar{\eta} \right) \Bigg) \end{aligned}$$

In the limit: $\lim_{N\to\infty}(y_i(\bar{\gamma}_i) - y_i^*(\bar{\gamma}_i)) = 0$, as defined in equation (A.37). Moreover, expected social output when $\alpha_B > \alpha_I$ is:

$$\begin{aligned}
\mathbb{E}[\tilde{\eta}_i\tilde{y}_i] &= \frac{1}{2\theta_i}\Bigg((\alpha_B - \alpha_I)\left(\frac{\sigma_\gamma^2/N}{\sigma_\eta^2 + \sigma_\gamma^2/N}\frac{\bar{\eta}^2}{\sigma_\gamma^2/N + \sigma_\eta^2 + \bar{\eta}^2} + \frac{\sigma_\eta^2}{\sigma_\eta^2 + \sigma_\gamma^2/N}\right)\mathbb{E}[\tilde{\eta}_i\tilde{\bar{\gamma}}_i] \\
&\quad + \alpha_I\left(\frac{\sigma_\eta^2}{\sigma_\eta^2 + \sigma_\gamma^2/N}\mathbb{E}[\tilde{\eta}_i\tilde{\bar{\gamma}}_i] + \frac{\sigma_\gamma^2/N}{\sigma_\eta^2 + \sigma_\gamma^2/N}\bar{\eta}\mathbb{E}[\tilde{\eta}_i]\right)\Bigg) \\
&= \frac{1}{2\theta_i}\Bigg((\alpha_B - \alpha_I)\left(\frac{\sigma_\gamma^2/N}{\sigma_\eta^2 + \sigma_\gamma^2/N}\frac{\bar{\eta}^2}{\sigma_\gamma^2/N + \sigma_\eta^2 + \bar{\eta}^2} + \frac{\sigma_\eta^2}{\sigma_\eta^2 + \sigma_\gamma^2/N}\right)\left(\bar{\eta}^2 + \sigma_\eta^2\right) \\
&\quad + \alpha_I\left(\bar{\eta}^2 + \frac{\sigma_\eta^2}{\sigma_\eta^2 + \sigma_\gamma^2/N}\sigma_\eta^2\right)\Bigg)
\end{aligned}$$

In the limit:

$$\lim_{N\to\infty}\mathbb{E}[\tilde{\eta}_i\tilde{y}_i] = \frac{1}{2\theta_i}\left((\alpha_B - \alpha_I)\left(\bar{\eta}^2 + \sigma_\eta^2\right) + \alpha_I\left(\bar{\eta}^2 + \sigma_\eta^2\right)\right) = \frac{\alpha_B}{2\theta_i}\left(\bar{\eta}^2 + \sigma_\eta^2\right)$$

■

References

Albuquerque, A.M., Carter, M.E., Guo, Z.M., Lynch, L.J., 2024. Complexity of CEO compensation packages. *Journal of Accounting and Economics*, forthcoming.

Barbalau, A., Zeni, F., 2022. The optimal design of green securities. Working paper, HEC Paris.

Burkert, S., Oberpaul, T., Tichy, N., Weller, I., 2024. *Academy of Management Discoveries*, 10, 273-306.

Grossman, S.J., Hart, O.D., 1979. A theory of competitive equilibrium in stock market economies. *Econometrica*, 47, 293-329.

Huang, S., Hwang, B.H., Lou, D. Yin, C., 2020. Offsetting disagreement and security prices. *Management Science*, 66, 3444-3465.

Murphy, K.J., Sandino, T., 2020. Compensation consultants and the level, composition, and complexity of CEO pay. *The Accounting Review*, 95, 311-341.

Starks, L.T., 2023. Presidential address: Sustainable finance and ESG issues–value versus values. *Journal of Finance*, 78, 1837-1872.

Yu, J., 2011. Disagreement and return predictability of stock portfolios. *Journal of Financial Economics*, 99, 162-183.