A Theory of Socially Responsible Investment *

Martin Oehmke† Marcus Opp‡

September 2, 2019

Abstract

Based on a canonical model of corporate financing under agency frictions, we characterize how and when socially responsible investors can affect firm behavior and derive an investment criterion, the social profitability index (SPI), to guide scarce socially responsible capital. The SPI highlights the importance of counterfactual social costs that would arise in the absence of socially responsible investors. Accordingly, most existing ESG metrics are not suited to guide investment decisions. Our model also uncovers a complementarity between financial capital and socially responsible capital: The presence of financial investors without regard for externalities can raise welfare relative to a setting with only socially responsible investors.

Keywords: Socially responsible investing, ESG, SPI, capital allocation, sustainable investment, social ratings.

JEL Classification: G31 (Capital Budgeting; Fixed Investment and Inventory Studies; Capacity), G23 (Non-bank Financial Institutions; Financial Instruments; Institutional Investors).

*For helpful comments and suggestions, we thank Patrick Bolton, Ailsa Roëll, Joel Shapiro, Paul Woolley, and seminar participants at Bonn, LSE, OxFit, and, the Stockholm School of Economics.

†London School of Economics and CEPR, e-mail: m.oehmke@lse.ac.uk.

‡Stockholm School of Economics, e-mail: marcus.opp@hhs.se.

1 Introduction

Socially responsible investing can no longer be considered a niche. Over the past 20 years, assets under management of funds with explicit sustainability considerations have grown manyfold (see e.g., Tett (2019)). At the same time, a large number of investment criteria that reflect environmental, social, and governance (ESG) criteria have been developed to guide the capital allocation decision of socially responsible investors. But can socially responsible investors really change economic outcomes? Can they reduce the incidence of social costs, such as the carbon emitted by energy companies or the systemic externalities generated by large banks? If yes, how should scarce socially responsible capital be allocated and how does it interact with competitive financial capital?

To shed light on these questions, this paper develops an equilibrium model of socially responsible investment and analyzes the role of impact capital in encouraging the adoption of sustainable production technologies. We show that socially responsible investors most effectively achieve impact by enabling a scale increase of clean production above and beyond the scale that competitive profit-motivated investors are willing to finance. When socially responsible capital is scarce, socially responsible investors should rank the allocation of impact capital across firms according to a social profitability index (SPI). This theoretically founded ESG metric summarizes the interaction of environmental, social and governance (agency) aspects. Importantly, the SPI not only accounts for the social return generated by the reformed firm, but also for the counterfactual pollution that would have occurred in the absence of impact investment. Our model also highlights a complementarity between financial and socially responsible capital in an economy: the presence of purely profit-oriented investors can raise clean investment and social welfare above and beyond what can be achieved by social responsible capital alone (even if the latter is in ample supply).

Our model builds on the canonical Holmström and Tirole (1997) setup, in which a firm's production scale is limited by a moral hazard friction. The main innovation is that firms can choose between two constant-returns-to-scale production technologies, which we label *dirty* and *clean*. The dirty production has a higher per-unit financial return, but clean production is socially preferable. To finance their operations, firms can raise financing from (up to) two types of investors. Financial investors, as their name suggests, care solely about financial returns. Socially responsible investors, on the other hand, care not only about financial returns but also about external social costs caused by the firm's production activities, such as carbon emissions.

We first show that, in the absence of socially responsible investors, even an ethical entrepreneur, who partially internalizes social costs of production, may choose the dirty production technology. This is because investors who care only about financial returns allow the entrepreneur to operate the dirty technology at a larger scale. If this effect is sufficiently strong, even an ethical entrepreneur chooses the dirty production technology.

We then characterize when and how socially responsible investors can guide firm investment towards the socially preferable clean technology. Because socially responsible investors internalize social costs of production, they are willing to provide more capital for the clean investment technology than financial investors, while still breaking even with respect to their broader mandate (but not on financial terms). If the additional capital that is available for the clean technology in the presence of socially responsible investors is sufficiently large, the entrepreneur adopts the clean technology.

Our model shows that the reason why socially responsible investors can effect an increase in the scale under the clean production technology is that the counterfactual social cost that would arise from dirty firm production in the absence of socially responsible investors' engagement relaxes their participation constraint. Counterfactual pollution (enabled by financial investors' investments) thus generates additional financing capacity from socially responsible investors and, hence, acts like a "quasi asset" to the entrepreneur. This mechanism points to a complementarity between financial and socially responsible capital and implies that welfare is generally higher in an economy in which there is a balance between financial and socially responsible capital. Intuitively, the presence of profit-motivated financial capital alleviates underinvestment for a given production technology, precisely because financial investors do not internalize the negative externalities of production. However, this disregard for externalities can come at the cost of a socially inefficient technology choice. The role of socially responsible investors is then to guide technology choice via a sufficiently large co-investment.

Based on this single-firm analysis, we then develop an investment criterion to optimally guide scarce socially responsible capital in an economy with many heterogeneous firms, the social profitability index (SPI). The SPI is similar to the classic profitability index in that it measures "bang for buck" invested. However, unlike the profitability index, the SPI not only reflects the (social) return of the project that is being funded, but also the social costs or externalities that would have occurred in the absence of engagement by socially responsible investors. Accordingly, it can be optimal to invest in firms that generate relatively low social returns (e.g., a gun manufacturer or an oil firm), provided that the potential increase in social costs, if only financially driven investors were to invest, is sufficiently large. This contrasts with many common ESG metrics that focus on firms' "social status quo", without explicitly taking into account these counterfactual social costs.

Related Literature. This paper contributes to a small but growing literature on socially responsible investing. In a pioneering contribution, Heinkel et al. (2001) show that firms that are blacklisted by socially responsible investors suffer an increase in their cost of capital due to a reduction in risk-sharing among their investor base. This "capital-cost" effect can induce firms to clean up their activities.[1] More recently, Hart and Zingales (2017) characterize the objective of a firm with "prosocial" investors and argue that firms should not maximize shareholder value, but instead shareholder *welfare*. Our socially re-

[1] On a slightly more skeptical note, Davies and Van Wesep (2018) show that divestment can have unintended consequences by inducing firms to prioritize short-term profit at the expense of long-term value.

sponsible investors are similar to prosocial investors with the important difference that ours care about externalities regardless of whether they invest in the firm or not. Morgan and Tumlinson (forthcoming) characterize how a firm's investment in public goods can resolve a free-rider problem among their investor base. All of these papers take the firm's ownership structure as given. In contrast, we endogenize the assignment of socially responsible investors to firms. Another key difference is that our paper features a moral hazard problem, and, therefore, underinvestment, which is a key ingredient for the complementarity between financial and social capital.

Chowdhry et al. (2018) focus on a commitment problem with respect to technology adoption, in the spirit of Glaeser and Shleifer (2001). A common theme with our paper is that the firm can monetize the socially-minded investors' social preference. However, while their analysis focuses on how socially-minded investors' ability to blunt a firm's profit motive, thereby allowing the firm to commit to emphasize social goals, we focus on the ability of socially responsible investors to impact firms by expanding a firm's maximum scale under clean production. Another important difference is that we develop an investment criterion, the SPI, to guide scarce socially responsible capital in a multi-firm setting.

Empirically, Hong and Kacperczyk (2009) show that "sin" stocks that are shunned by some investors have higher expected returns. Chava (2014) documents higher implied costs of capital for firms with significant environmental concerns. Barber et al. (2018) show that impact funds earn lower financial returns. All of these empirical results are consistent with our prediction that it is necessary to sacrifice financial returns to achieve impact.

2 Model Setup

Our model builds on the canonical model of corporate financing in the presence of agency frictions laid out in Holmström and Tirole (1997) and Tirole (2006). The main innovation is that the firm has access to two different production technologies, one of them "clean" (i.e., associated with low social costs) and the other "dirty" (i.e., associated with larger social costs).

The entrepreneur, production, and moral hazard. Our setting considers a risk-neutral entrepreneur who is protected by limited liability and endowed with initial liquid assets of A. The entrepreneur has access to two mutually exclusive production technologies $\tau \in \{C, D\}$, each with constant returns to scale. The technologies are identical in terms of revenue generation. Denoting firm scale (capital) by K, the firm generates positive cash flow of RK with probability p (conditional on effort by the entrepreneur, see below) and zero otherwise. Where the technologies differ is with respect to their production cost and the social costs they generate. In particular, the dirty technology D generates a non-pecuniary negative externality of $\phi_D > 0$ per unit of scale and requires an upfront investment k_D per unit of scale. The clean technology, on the other hand, results in lower per-unit social costs, $\phi_C < \phi_D$, but comes at higher variable production cost $k_C > k_D$. The entrepreneur internalizes a fraction $\gamma^E \in [0, 1)$ of social costs, capturing (potential) intrinsic motives not to cause social harm. Since we do not model government intervention, the two technologies can be interpreted as those available to the firm after potential government intervention or regulation has taken place.[2] Alternatively, our analysis can be interpreted as establishing what market forces (in the form of socially responsible investors) can achieve before government intervention takes place.

To generate a meaningful trade-off in the choice of technologies, we assume that the

[2] Because regulation or intervention is usually subject to informational or political economy constraints, it seems reasonable that the social costs of production cannot be dealt with by the government alone, creating a potential role for socially responsible investors.

ranking of the two technologies differs depending on whether it is based on financial or social value. Denoting the per-unit financial value by $\pi_\tau := pR - k_\tau$ and the per-unit social value (welfare) by $v_\tau := \pi_\tau - \phi_\tau$, we posit that the dirty technology has a higher financial return, $\pi_D > \pi_C$, but clean production generates higher social welfare, $v_C > 0 > v_D$. The final inequality implies that the social return of the dirty production technologies is negative, meaning that the externalities caused by dirty production outweigh the financial value. The assumption that the dirty production technology has a negative social return is not necessary for our results, but it simplifies the exposition.

As in Holmström and Tirole (1997), the entrepreneur is subject to an agency problem. In particular, while the investment pays off with probability p if the entrepreneur exerts effort ($a = 1$), this probability is reduced to $p - \Delta p$ when the entrepreneur shirks ($a = 0$), where $\Delta p > 0$. Shirking yields a per-unit non-pecuniary benefit of B to the entrepreneur, for a total private benefit of BK. A standard result (which we will show below) is that this agency friction reduces the firms unit pledgeable income by $\xi := p\frac{B}{\Delta p}$, the per-unit agency cost. A high value of ξ can be interpreted as an indicator of poor governance, such as large private benefits or weak performance measurement. We make the following assumption on the per-unit agency cost:

Assumption 1 *For each technology τ, the agency cost per unit of capital $\xi := p\frac{B}{\Delta p}$ satisfies*

$$\pi_\tau < \xi < pR - \frac{p}{\Delta p}\pi_\tau. \tag{1}$$

This assumption states that the moral hazard problem, as characterized by the agency cost per unit of capital ξ, is neither too weak nor too severe. The first inequality implies a finite production scale. The second inequality is a sufficient condition that rules out equilibrium shirking and ensures feasibility of outside financing. To streamline notation, our definitions of π and v are defined conditional on the relevant case, in which the entrepreneur exerts effort.

Outside investors and securities. The entrepreneur can raise financing from (up to) two types of risk-neutral outside investors $i \in \{F, SR\}$, *financial investors* and *socially responsible* investors. Both investor types care about expected cash flows, but only socially responsible investors internalize social costs of production. Regardless of whether the entrepreneur raises financing from both investor types or just one, it is without loss of generality to restrict attention to financing arrangements in which the entrepreneur issues securities that pay out a total repayment amount of $X := X^F + X^{SR}$ upon project success and 0 otherwise, where X^F and X^{SR} denote the payments promised to financial and socially responsible investors, respectively. As usual, given that the firm has no resources in the low state, this security can be interpreted as debt or equity. In our baseline specification, we assume that the entrepreneur's technology choice is contractible.

Then, the entrepreneur's (net) utility as a function of the investment scale K, total repayment X, effort decision a, upfront consumption by the entrepreneur c, and technology choice $\tau \in \{C, D\}$ is given by

$$\begin{aligned} U^E(K, X, \tau, c, a) =& p(RK - X) - (A - c) - \gamma^E \phi_\tau K \\ &+ \mathbb{1}_{a=0}\left[BK - \Delta p(RK - X)\right]. \end{aligned} \tag{U^E}$$

The first term of this expression, $p(RK - X) - (A - c)$, represents the net financial payoff of the project under high effort, where $A - c$ can be interpreted as the upfront coinvestment made by the entrepreneur. The second term, $\gamma^E \phi_\tau K$, measures the social cost internalized by the entrepreneur. The third term, $BK - \Delta p(RK - X)$, captures the incremental payoff conditional on shirking. Exerting effort is incentive compatible if and only if $U^E(K, X, \tau, c, 1) \geq U^E(K, X, \tau, c, 0)$, which limits the total amount X that the entrepreneur can promise to repay to outside investors to

$$X \leq \left(R - \frac{B}{\Delta p}\right) K, \tag{IC}$$

so that the entrepreneur's unit pledgeable income is given by $pR - \xi$. The resource constraint at date 0 implies that the capital expenditures, Kk_τ, must equal the total investments made by the entrepreneur, $A - c$, financial investors, I^F, and socially responsible investors, I^{SR}, so that

$$Kk_\tau = A - c + I^F + I^{SR}. \tag{2}$$

The respective (net) utility functions of outside investors, given an incentive-compatible financing arrangement, are given by:

$$U^F = pX^F - I^F, \tag{U^F}$$

$$U^{SR} = pX^{SR} - I^{SR} - \gamma^{SR}\phi_\tau K. \tag{U^{SR}}$$

Here, $\gamma^{SR} \leq 1 - \gamma^E$ captures the degree to which socially responsible investors internalize externalities.[3] Their payoff function highlights two important features. First, socially responsible investors are affected by externalities $\gamma^{SR}\phi_\tau K$ as determined by the scale K and technology choice τ regardless of whether they invest in the firm or not (we discuss this in more detail in Section 3.3). Second, their payoff function is distinct from maximizing the social value of the project, $v_\tau K$, even if they fully account for the externalities $\left(\gamma^{SR} = 1\right)$. The reason is that socially responsible investors internalize neither the value of the cash flows that accrue to the entrepreneur, $p\left(RI - X^{SR} - X^F\right) - (A - c)$, nor those accruing to financial investors, $pX^F - K^F$.

We are interested in a setting in which deep-pocketed financial investors behave competitively. However, to abstract from free-rider issues, we assume that socially responsible investors allocate their capital in a coordinated fashion.[4] One interpretation of this as-

[3] The sum $\gamma^E + \gamma^{SR}$ represents the fraction of externalities that are taken into account by agents in the model. When $\gamma^E + \gamma^{SR} < 1$, some externalities (e.g., those imposed on future generations) are not taken into account by anyone.

[4] Morgan and Tumlinson (forthcoming) provide a framework in which shareholders of a company value public good production but are subject to free-rider problems.

sumption is that socially responsible capital is directed by one large fund.[5] While for the partial equilibrium, single-firm analysis of Section 3 we assume that socially responsible capital is abundant relative to the capital needed by the firm, the subsequent multi-firm setting presented in Section 4 considers a general equilibrium analysis with limited social capital.

3 The Effect of Socially Responsible Investment

In this section, we investigate whether and how socially responsible investors can impact the firm's investment choice. To do so, in Section 3.1, we first solve a benchmark case without socially responsible investors. This benchmark shows that, in the absence of socially responsible investors, the dirty technology may be chosen even when the entrepreneur has some concern for the higher social cost generated by dirty production (i.e., $\gamma^E > 0$). In Section 3.2, we add socially responsible investors to the model and characterize conditions under which their presence has impact, in the sense that it leads to the adoption of the clean production technology.

3.1 Benchmark: Only Financial Investors

We initially consider the benchmark setting in which the entrepreneur can only borrow from competitive financial investors. This setting corresponds to the special case of $I^{SR} = X^{SR} = 0$.

The entrepreneur's objective is to choose a financing arrangement (consisting of scale $K \geq 0$, repayment $X^F \in [0, R]$, upfront consumption by the entrepreneur $c \geq 0$, and, technology choice $\tau \in \{C, D\}$) that maximizes the entrepreneur's utility U^E subject to the entrepreneur's IC constraint and financial investors' IR constraint, $U^F \geq 0$.

[5] Coordinated behavior is a natural assumption when socially responsible capital is deployed by large agents, such as sovereign wealth funds. In addition, there is also increasingly evidence for coordination among smaller players. One such example is the establishment of the Poseidon Principles, an initiative by eleven major to promote green shipping, see Nauman (2019).

As a preliminary step, it is useful analyze the financing arrangement that maximizes the scale for a given technology τ. Following standard arguments (see Tirole (2006)), this agreement requires the entrepreneur to co-invest all her wealth (i.e., $c = 0$) and that the entrepreneur's IC constraint as well as the financial investors' IR constraint bind. The binding IC constraint ensures that the firm optimally leverages its initial resources A, whereas the binding IR constraint is a consequence of competition among financial investors.

When all outside financing is raised from financial investors, the maximum firm scale under production technology τ is then given by

$$K_\tau^F = \frac{A}{\xi - \pi_\tau}. \tag{3}$$

This expression shows that the entrepreneur can scale his initial assets A by a factor that depends on the agency cost per unit of investment, $\xi = p\frac{B}{\Delta p}$, and the financial return under technology τ. As $\xi > \pi_D$ (see Assumption 1), the maximum investment scale is finite under either technology.

The key observation from Equation (3) is that the maximum scale that the entrepreneur can finance from financial investors is larger under dirty than under clean production,

$$K_D^F > K_C^F. \tag{4}$$

The reason for this difference in scale is that dirty production has a higher financial value than clean production, $\pi_D > \pi_C$ and that financial investors only care about financial returns.

The following Lemma 1 highlights that the technology choice of the entrepreneur is driven by a trade-off between scale and her concern for externalities. Of course, if the entrepreneur completely disregards externalities $\left(\gamma^E = 0\right)$, there is no trade-off and she will always choose dirty production, given that $K_D^F > K_C^F$.

Lemma 1 (Benchmark: Financial Investors Only) *When only financial investors are present, the entrepreneur chooses*

$$\bar{\tau} = \arg\max_{\tau}(\xi - \gamma^E \phi_\tau)K_\tau^F. \tag{5}$$

The firm operates at maximum scale that allows financial investors to break even $K_{\bar{\tau}}^F$. The entrepreneur's net utility is given by

$$\bar{U}^E = (\xi - \gamma^E \phi_{\bar{\tau}})K_{\bar{\tau}}^F - A. \tag{6}$$

Maximum scale is optimal because, under the equilibrium technology $\bar{\tau}$, the project generates positive surplus for the entrepreneur and financial investors. Moreover, because the maximum scale is greater under the dirty technology, the entrepreneur may prefer the dirty technology even if he internalizes some of the social costs of production (i.e., $\gamma^E > 0$). From equation (6), we see that this is the case when

$$(\xi - \gamma^E \phi_D)K_D^F > (\xi - \gamma^E \phi_C)K_C^F, \tag{7}$$

which yields

Corollary 1 (Financial Investors Can Corrupt Ethical Entrepreneur) *When only financial investors are present, the entrepreneur chooses the dirty technology whenever $\gamma^E < \bar{\gamma}^E$, where $\bar{\gamma}^E := \frac{\xi(\pi_D - \pi_C)}{\phi_D(\xi - \pi_C) - \phi_C(\xi - \pi_D)} > 0$.*

Therefore, if the entrepreneur's concern for social costs lies below the cutoff $\bar{\gamma}^E$, the entrepreneur adopts the dirty production technology because it can be run at a larger scale. As a result, outside investors that are driven purely by financial returns can induce even an ethical entrepreneur ($\gamma^E > 0$) to abandon principle and adopt dirty production.

3.2 Equilibrium with Socially Responsible Investors

We now analyze whether and how the financing arrangement is altered when socially responsible investors are also present. Because the entrepreneur could still raise financing exclusively from financial investors, the utility under the financing arrangement with financial investors only, $\bar{U}^E$, now takes the role of an outside option to the entrepreneur.

3.2.1 Optimal Financing with Socially Responsible Investors

Due to their unconditional concern for externalities, socially responsible investors are affected by the social costs of production regardless of whether they have a financial stake in the firm or not. In particular, if socially responsible investors do not engage with the firm, their (reservation) utility is given by

$$\bar{U}^{SR} = -\gamma^{SR}\phi_{\bar{\tau}}K^F_{\bar{\tau}} < 0, \tag{8}$$

which reflects the social costs generated when the entrepreneur raises financing from financial investors only.

To improve their payoff above and beyond this status quo, socially responsible investors can engage with the entrepreneur. Because socially responsible investors act in a coordinated fashion, they make a take-it-or-leave-it contract offer that specifies the technology τ, scale K as well as the required financial investments and payoffs for all investors and the entrepreneur. This contract solves the following maximization problem:

Problem 1 (Maximization problem faced by socially responsible investors)

$$\max_{I^F, I^{SR}, X^{SR}, X^F, K, c, \tau} pX^{SR} - I^{SR} - \gamma^{SR}\phi_\tau K \tag{9}$$

subject to IR of the entrepreneur:

$$U^E\left(K, X^{SR} + X^F, \tau, c, 1\right) \geq \bar{U}^E \tag{IR^E}$$

as well as the entrepreneur's IC constraint, the resource constraint (2), the financial investors' IR constraint $U^F \geq 0$*, and the non-negativity constraints* $K \geq 0, c \geq 0$.

The key difference relative to the previous section is that the financing agreement is now chosen to maximize the socially responsible investor's utility subject to the constraint that the entrepreneur is weakly better off than her outside option of raising financing exclusively from financial investors (IR^E). We note that this formulation permits the possibility of compensating the entrepreneur with sufficiently high upfront consumption ($c > 0$) in return for smaller scale K, possibly even shutting down production completely. However, (at a minimum) the clean production technology generates positive joint surplus for the entrepreneur and socially responsible investors, the optimal financing arrangement rewards the entrepreneur with (weakly) larger scale than what could be funded by financial investors alone for the chosen technology, as shown in Theorem 1.

Theorem 1 (Financing with Financial and Socially Responsible Investors) *Let* $\hat{v}_\tau := \pi_\tau - \left(\gamma^E + \gamma^{SR}\right)\phi_\tau \geq v_\tau$ *denote the joint surplus, per unit of scale, accruing to all investors and the entrepreneur. Then, in any optimal financing arrangement, production is characterized by*

$$\hat{\tau} = \arg\max_{\tau} \frac{\hat{v}_\tau}{\xi - \gamma^E \phi_\tau}, \tag{10}$$

$$\hat{K} = \frac{\xi - \gamma^E \phi_{\bar{\tau}}}{\xi - \gamma^E \phi_{\hat{\tau}}} K^F_{\bar{\tau}} \geq K^F_{\hat{\tau}}. \tag{11}$$

The entrepreneur consumes no resources upfront, $\hat{c} = 0$*. The total date-*0 *investment by both investors is* $\hat{I} = \hat{K} k_{\hat{\tau}} - A$ *and the total payout to both investors satisfies* $\hat{X} = \left(R - \frac{B}{\Delta p}\right)\hat{K}$*. The set of optimal co-investment arrangements, can be obtained by tracing* *out* $x^F \in \left[0, \hat{X}\right]$ *and setting* $\hat{X}^F = x^F$, $\hat{X}^{SR} = \hat{X} - \hat{X}^F$ *as well as* $\hat{I}^F = p\hat{X}^F$ *and* $\hat{I}^{SR} = \hat{I} - \hat{I}^F$. *The utility of socially responsible investors satisfies:*

$$\hat{U}^{SR} = \left(\pi_{\hat{\tau}} - \xi\right)\hat{K} + A - \gamma^{SR}\phi_{\hat{\tau}}\hat{K}. \tag{12}$$

The optimal choice of technology maximizes total joint surplus, which is governed by the joint surplus that is created per unit of capital, $\hat{v}_\tau$, and a term, $\frac{1}{\xi - \gamma^E \phi_{\hat{\tau}}}$, that reflects the optimal scale $\hat{K}$ (see Equation (11)). An immediate implication is that if the entrepreneur and the socially responsible investors jointly internalize all externalities, $\gamma^E + \gamma^{SR} = 1$, production will always be clean, since in this case $\hat{v}_\tau$ coincides with social welfare v_τ (and dirty production generates negative social welfare). Another implication of Theorem 1 is that the optimal financing arrangement rewards the entrepreneur entirely with scale, in the sense that the optimal capital stock $\hat{K}$ is chosen so that the entrepreneur obtains the same utility as in her outside option $\bar{U}^E$. Intuitively, any upfront consumption by the entrepreneur is suboptimal in the presence of a moral hazard problem that gives rise to capital rationing and, consequently, underinvestment.

While the optimal financing arrangement uniquely pins down the production side (i.e., technology choice and scale), there exists a continuum of feasible co-investment arrangements between financial and socially responsible investors that implements this outcome. Intuitively speaking, this is the case because any increase in the cash flow stake of financial investors $\hat{X}^F$ is reflected competitively in a higher upfront investment $\hat{I}^F$. Because also the entrepreneur remains at her reservation utility, the payoff to socially responsible investors as well as aggregate surplus remains unchanged.

We now compare the optimal arrangement to the benchmark case presented in Lemma 1. Of course, if the engagement by socially responsible investors does not result in a change in production technology compared to the benchmark case occurs (i.e., $\hat{\tau} = \bar{\tau}$), we obtain the same level of the capital stock and same utility for all agents in the economy.

This occurs either if the entrepreneur adopts the clean production technology even in the absence of investment by socially responsible investors, or if the entrepreneur adopts the dirty technology irrespective of whether socially responsible investors provide funding.

The interesting case is the one in which the optimal financing arrangement described in Theorem 1 induces a change in the production technology from dirty to clean. In this case, engagement by socially responsible investors has real impact. Specifically, socially responsible investors facilitate additional scale under the clean technology (relative to the case with only financial investors) to induce the entrepreneur to adopt the clean technology. When the entrepreneur does not internalize any of the social costs ($\gamma^E = 0$), this requires that the production scale under the clean technology is the same as when financial investors fund the dirty technology (i.e., $\hat{K} = K_D^F > K_C^F$). Intuitively, when the entrepreneur does not care about social costs of production, socially responsible investors have to completely make up for "lost scale" that results from the switch to the clean technology. When the entrepreneur internalizes some of the social costs of production ($\gamma^E > 0$), partially making up for lost scale is sufficient, because the entrepreneur is compensated in part for the switch to clean production by an increase in intrinsic utility (i.e., $K_D^F > \hat{K} > K_C^F$) .

By engaging with the firm, socially responsible investors increase their utility relative to the case in which they remain passive,

$$\Delta U^{SR} := \hat{U}^{SR} - \bar{U}^{SR} = \hat{v}_C \hat{K} - \hat{v}_D K_D^F > 0. \tag{13}$$

However, even though socially responsible investors increase their (overall) payoff by engaging with the company, they never break even when looking purely at financial returns, as stated in the following corollary.

Corollary 2 (Socially Responsible Investors Make a Financial Loss) *Any induced switch in the production technology from $\bar{\tau} = D$ to $\hat{\tau} = C$ requires that socially respon-*

sible investors make a financial loss. That is, in any optimal financing arrangement, as characterized in Theorem 1,

$$p\hat{X}^{SR} < \hat{I}^{SR}. \tag{14}$$

Intuitively, to induce a change from dirty to clean production, socially responsible investors need to enable a scale for the clean technology that is greater than that offered by financial investors in isolation. Because competitive financial investors just break even at the scale of production they are willing to finance in isolation, it must be the case that socially responsible investors make a financial loss when they finance an expansion in scale of the clean technology above and beyond what is offered by financial investors. Nevertheless, socially responsible investors are willing to provide financing because this financial loss, $p\hat{X}^{SR} - \hat{I}^{SR}$, is outweighed by the utility gain from reduced social costs, $\gamma^{SR}\left(\phi_D K_D^F - \phi_C \hat{K}\right)$, which generates the net gain in utility in Equation (13). It is important to note that our model predicts that this financial loss, $p\hat{X}^{SR} - \hat{I}^{SR}$, occurs when the firm seeks financing in the primary market. That is, if socially responsible investors were to sell their cash flow stake $\hat{X}^{SR}$ to financial investors after the firm has financed the clean technology, our model does not predict a price premium in the secondary market (i.e., in the secondary market we would observe $p\hat{X}^{SR} = \hat{I}^{SR}$).

3.2.2 Complementarity between Financial and Social Capital

To highlight the economic mechanism behind Theorem 1, this section provides a more detailed investigation of the relevant special case, in which socially responsible investors have impact (i.e., the entrepreneur would have chosen the dirty technology at scale $K_D^F = \frac{A}{\xi - \pi_D}$ in the absence of socially responsible investors). The key insight of this section is that the counterfactual pollution under the dirty technology (which is enabled by financial investors) acts like a quasi asset to the firm, thereby raising the financing capacity from socially responsible investors. This quasi asset, in turn, is instrumental in generating a complementarity between financial and social capital, which we highlight in Proposition

1 below: social surplus is higher when both financial and socially responsible investors deploy capital, relative to cases where all capital is allocated by either financial or socially responsible investors.

To illustrate this complementarity, it is instructive to first consider a setting in which only the clean technology is available and to compare the maximum feasible scale of operation that can be sustained with either type of capital. While the maximum clean scale under financial capital is given by Equation (3) the maximum feasible clean scale under socially responsible capital, $\overline{K_C^{SR}}$, is obtained analogously from binding IC of the entrepreneur and binding IR of socially responsible investors.[6] As is immediate from the following equation, financial capital can sustain a strictly higher clean scale:

$$K_C^F = \frac{A}{\xi - \pi_C} > \overline{K_C^{SR}} = \frac{A}{\xi - \pi_C + \gamma^{SR}\phi_C}. \tag{15}$$

Intuitively, financial investors alleviate capital rationing that results from the entrepreneur's agency problem. They do so precisely because they do not internalize negative externalities and, hence, perceive each unit of the project as more valuable (by $\gamma^{SR}\phi_C$). Since clean production suffers from an underinvestment problem, higher scale is socially valuable so that, conditional on the clean technology being adopted, welfare in an economy with only financial capital is strictly higher than in an economy with only social responsible capital,

$$v_C \left(K_C^F - \overline{K_C^{SR}} \right) > 0. \tag{16}$$

Hence, for a given technology with positive social value, financial investors are more efficient at funding the firm than socially responsible investors.

However, this "aggressive" investment style of financial investors can have negative social implications with regards to technology adoption. In particular, if the dirty pro-

[6] $\overline{K_C^{SR}}$ is the maximum scale at which socially responsible investors would just break even. Note, however, that, as discussed below, coordinated socially responsible investors will generally not fund up to this scale.

duction technology is also available, then financing from financial investors only can lead to dirty rather than clean production (recall from Corollary 1 that this happens when $\gamma^E < \bar{\gamma}^E$). Therefore, financial investors can induce overinvestment (from a social perspective) in the dirty production technology with negative social value.

Finally, consider an economy with both production technologies and both financial and socially responsible capital in ample supply and consider the case $\gamma^E < \bar{\gamma}^E$, so that dirty production would occur in the absence of socially responsible capital. Now the presence of financial investors implies that the entrepreneur has the outside option of adopting the dirty production technology. Because this outside option generates negative externalities, the socially responsible investors' participation constraint is relaxed, i.e., their payoff in (U^{SR}) must now exceed $-\gamma^{SR}\phi_D K_D^F$. This relaxation of the participation constraint, in turn, raises the financing capacity from socially responsible investors and enables a scale increase relative to $\overline{K_C^{SR}}$. The maximal feasible total scale, $\overline{K_C^{F+SR}}$, with both types of investors satisfies

$$\overline{K_C^{F+SR}} = \frac{A + \widetilde{A}}{\xi - \pi_C + \gamma^{SR}\phi_C}. \tag{17}$$

Equation (17) highlights that the counterfactual social cost $\tilde{A} := \gamma^{SR}\phi_D K_D^F > 0$ enters the maximum scale in the same way as the entrepreneur's financial assets A. Hence, it acts like a quasi asset to the entrepreneur. Since the privately efficient arrangement with competitive financial investors maximizes scale (see Lemma 1) it, therefore, maximizes the value of this quasi asset. The following Proposition shows that this effect unlocks sufficient additional investment from socially responsible investors so that this maximum feasible scale not only exceeds $\overline{K_C^{SR}}$, but also K_C^F.

Proposition 1 (Financial and Social Capital Are Complementary) *Suppose that $\gamma^E < \bar{\gamma}^E$ and $\hat{\tau} = C$, then financial capital and socially responsible capital are complements: The maximum feasible scale under the clean technology in the presence of both* *financial and socially responsible capital, $\overline{K_C^{F+SR}}$, is larger than the maximum scales attainable with only one investor type,*

$$\overline{K_C^{F+SR}} > \hat{K} > K_C^F > \overline{K_C^{SR}}. \tag{18}$$

Under the maximum feasible scale, $\overline{K_C^{F+SR}}$, socially responsible investors would just break even. This implies that socially responsible investors will generally not provide financing up to this scale. Because socially responsible investors make a financial loss on each unit they finance and because they act in a coordinated fashion, the equilibrium scale $\hat{K}$ is just sufficient to induce the entrepreneur to switch to the clean technology, thereby keeping the entrepreneur at her outside option of running the dirty technology at scale K_D^F. By revealed preference, this equilibrium scale $\hat{K}$ also has to exceed K_C^F: The entrepreneur could have always chosen to run the firm in clean mode at scale K_C^F when relying on financial investors only but, given $\gamma^E < \bar{\gamma}^E$ chose not to do so. Perhaps surprisingly, in the presence of both financial and socially responsible investors, equilibrium welfare $v_C\hat{K}$ is therefore larger than in an economy in which the dirty technology is not available (e.g., due to government regulation).

Of course, an important assumption underlying the scale and concomitant welfare increase is that socially responsible capital is available in sufficient amounts to ensure adoption of the clean production technology. When this is not the case (i.e., when socially responsible capital is scarce) the presence of financial capital can move firms toward the adoption of dirty production technologies, leading to a social loss. In Section 4, we consider an economy with multiple firms and limited aggregate socially responsible capital. This analysis will shed further light on how the composition of investor capital (and not just the aggregate amount of capital) matters.

3.3 Broad vs. Narrow Socially Responsible Investment

A key assumption in our analysis is that socially responsible investors care unconditionally about external social costs, irrespective of whether they are complicit in the generation of these costs through investment in the company that is responsible for them. To illustrate the importance of this assumption, we briefly consider an alternative setting in which socially responsible investors follow a narrow mandate that is determined only by social costs that are a direct consequence their own investments. Under such a narrow mandate, socially responsible investors continue to internalize the social cost generated by their own investments, but not the social costs that are generated when they not investors themselves. In this case, the participation constraint for socially responsible investors becomes

$$U^{SR} = pX^{SR} - I^{SR} - \gamma^{SR}\phi_\tau K \geq 0. \tag{19}$$

In this case, there is no quasi asset ($\widetilde{A} = 0$) and socially responsible investors cannot increase the scale of the clean production technology above and beyond what financial investors are willing to offer. Therefore, a narrow mandate that focuses only on the direct consequences of the funds invested by socially responsible investors themselves does not allow for an increase in scale of the clean technology beyond what financial investors are willing to fund, and is therefore not sufficient for effective impact investment.

4 The Social Profitability Index

Based on the results presented in Section 3, we now extend the model to a multi-firm setting in order to derive a micro-founded investment criterion to guide scarce socially responsible capital. We denote by κ the aggregate amount of socially responsible capital (and continue to assume that financial capital is abundant).

We consider an economy with a continuum of infinitesimal firms grouped into distinct firm types. Firms that belong to the same firm type j are identical in terms of all relevant fundamentals of the model, whereas firms belonging to different types differ according to at least one dimension, with Assumption 1 satisfied for all types. Let $\mu(j)$ denote the distribution function of firm types, then the aggregate social cost in the absence of socially responsible investors is given by

$$\int_{\gamma_j^E < \bar{\gamma}_j^E} \phi_{D,j} K_{D,j}^F d\mu(j) + \int_{\gamma_j^E \geq \bar{\gamma}_j^E} \phi_{C,j} K_{C,j}^F d\mu(j). \tag{20}$$

The first term of this expression captures the social cost generated by firms that, in the absence of socially responsible investors, choose the dirty technology ($\gamma_j^E < \bar{\gamma}_j^E$), whereas the second term captures firm types run by entrepreneurs that have enough concern for external social costs that they choose the clean technology even in absence of socially responsible investors ($\gamma_j^E \geq \bar{\gamma}_j^E$).

Given this aggregate social cost, how should socially responsible investors allocate their limited capital? One direct implication of Theorem 1 is that any investment in firm types with $\gamma_j^E \geq \bar{\gamma}_j^E$ cannot be optimal. For these firms, socially responsible investors cannot induce a change in the adopted technology, such that any social capital used on those firms would be wasted from a social perspective. For the remaining "reformable" firm types, the payoff to socially responsible investors from reforming a firm of a given type j is given by:

$$\Delta U_j^{SR} = \left(\pi_{C,j} - \xi_j\right) \hat{K}_j + A_j + \gamma^{SR} \left[\phi_{D,j} K_{D,j}^F - \phi_{C,j} \hat{K}_j\right]. \tag{21}$$

The first term captures the project's financial return, net of the agency cost that is necessary to incentivize the entrepreneur, scaled by the (optimal) scale $\hat{K}_j$. The second term captures the (internalized) change in social cost that results from inducing a firm of type j to adopt the clean production technology.

Given that socially responsible investors with limited capital may not be able to reform all firms, they should prioritize investments in firm types that maximize the *impact per* *dollar invested.* This is achieved by ranking firms according to a variation on the classic profitability index, the *social profitability index*, which divides the change in payoffs to socially responsible investors, ΔU_j^{SR}, by the amount socially responsible investors need to co-invest to impact the firm's behavior, I^{SR}. By Theorem 1, the required co-investment depends on the fraction of cash-flow rights X that socially responsible investors receive. In the polar cases of a zero cash-flow stake ($X^{SR} = 0$) and a full cash-flow stake ($X^{SR} = \hat{X}$), their required co-investment is given by

$$I_{\min,j}^{SR} = (\xi_j - \pi_{C,j})\,\hat{K}_j - A_j, \text{ and} \tag{22}$$

$$I_{\max,j}^{SR} = \hat{K}_j k_{C,j} - A_j, \tag{23}$$

respectively.

As shown by Chowdhry et al. (2018), if the technology change is fully contractible, then the minimum co-investment, $I_{\min,j}^{SR}$, is optimal when socially responsible capital is scarce. However, in characterizing the SPI we want to allow for the realistic situation, in which socially responsible investors do receive cash flow rights. This could be the case because the entrepreneur cannot commit to the adoption of the clean technology. In this case, a cash-flow stake for socially responsible investors and blunt the entrepreneur's profit motive or may allow socially responsible investors to enforce appropriate technology adoption, for example via voting rights. Alternatively, socially responsible investors may be subject to an institutional constraint that requires them to deliver a certain fraction of their returns in terms of financial rather than nonpecuniary form.

To capture these considerations in a simple fashion, we introduce the parameter $\lambda_j \in [0, 1]$ which denotes the fraction of cash flow rights that socially responsible investors require in order to be willing to invest. We can then write the SPI as follows.

Proposition 2 (The Social Profitability Index) *Socially responsible investment should be guided by the social profitability index SPI_j, which for any firm type j is given by the* *harmonic mean of polar SPIs,*

$$SPI_j := \mathbb{1}_{\gamma_j^E < \bar{\gamma}_j^E} \frac{1}{\frac{\lambda_j}{SPI_{\min,j}} + \frac{1-\lambda_j}{SPI_{\max,j}}}. \tag{24}$$

where $SPI_{\max,j} := \frac{\Delta U_j^{SR}}{I_{\min,j}^{SR}}$ *and* $SPI_{\min,j} := \frac{\Delta U_j^{SR}}{I_{\max,j}^{SR}}$. *There exists a threshold level* $SPI^*(\kappa) \geq 0$ *such that socially responsible investors with scarce capital* κ *should invest in all firms for which* $SPI_j \geq SPI^*(\kappa)$.

According to Proposition 2, the optimal investment strategy for socially responsible investors is to first rank firms according to the social profitability index and then invest into these ranked firms until no funds are left, which will occur at the cutoff $SPI^*(\kappa)$. Social capital is scarce if and only if the amount κ is not sufficient to fund all firm types with $SPI_j > 0$.

The welfare change relative to the counterfactual case without socially responsible investors, $\Delta\Omega$, results purely from the set of reformed firms, i.e., firms for which $\gamma_j^E < \bar{\gamma}_j^E$ and $SPI_j \geq SPI^*(\kappa)$. That is,

$$\Delta\Omega = \int_{j:\gamma_j^E < \bar{\gamma}_j^E \ \&\ SPI_j \geq SPI^*(\kappa)} \left(v_{C,j}\hat{K}_j - v_{D,j}K_{D,j}^F \right) d\mu(j). \tag{25}$$

Clearly, if social capital is abundant and $\gamma^E + \gamma^{SR} = 1$, then the partial equilibrium results of Proposition 1 still apply. Welfare is strictly higher than in an economy where all capital is held exclusively by either financial or socially responsible investors. However, when social capital is scarce, there is a trade-off. On the one hand, the set of reformed firms contributes towards higher welfare as before. On the other hand, the set of unreformed "dirty" firms may exhibit overinvestment in the dirty technology due to the presence of competitive financial capital without regard for externalities. This trade-off leads to the following Proposition, which highlights the importance of a balance between social and financial capital.

Proposition 3 (Balanced Capital) *Fix the aggregate amount of capital in the economy. Welfare is higher compared to the case in which no capital is held by financial investors if and only if the amount of social capital exceeds a threshold.*

To examine how the SPI guides capital allocation by socially responsible investors, let us consider which types of firms rank highest according to this metric. In performing these comparative statics it is instructive to first consider the special case in which $\gamma^E = 0$ and $\gamma^{SR} = 1$. Denoting the difference in social costs by $\Delta\phi := \phi_D - \phi_C$ and the difference in (financial) profitability by $\Delta\pi := \pi_D - \pi_C$, we obtain that:

$$\text{SPI}_j = \mathbb{1}_{\gamma_j^E < \bar{\gamma}_j^E} \frac{\Delta\phi_j - \Delta\pi_j}{\Delta\pi_j + \lambda_j \left(p_j R_j - \xi_j\right)}. \tag{26}$$

This expression illustrates an important feature of the SPI: it reflects not only the social costs ϕ_C produced by the firm under the clean technology (i.e., conditional on impact investing), but also the counterfactual social costs that would have occurred in the absence of engagement from impact investors, ϕ_D. This means that optimal capital allocation by socially responsible investors can include investments in firms that generate significant social costs (e.g., because of heavy reliance on fossil fuels) if these firms would have generated much larger social costs in the absence of engagement by socially responsible investors. Of course, the reform potential, as summarized by the relevant difference in $\Delta\phi_j$, has to be traded off against the costs, as measured by the resulting reduction in financial profits $\Delta\pi_j$. Moreover, intuitively, firm types that require a higher cash-flow stake λ_j to ensure a technology change rank lower.

In the general case (allowing for $\gamma^E > 0$), we obtain the following comparative statics:

Proposition 4 *As long as $\gamma_j^E < \bar{\gamma}_j^E$, the SPI is increasing in $\Delta\phi$, ξ, and γ^E and decreasing in $\Delta\pi$ and λ.*

Thus, as long as $\gamma_j^E < \bar{\gamma}_j^E$ firm types with more socially minded entrepreneurs are cheaper to invest in, as a smaller scale is needed to convince the entrepreneur to reform.

However, as soon as the entrepreneur internalizes enough of the externalities so that she chooses the clean technology even if financed by financial investors ($\gamma_j^E > \bar{\gamma}_j^E$), the SPI drops discontinuously to zero. That is, socially responsible investors should not invest in these types of firms (or should divest in the case of pre-existing ownership). Finally, the SPI is (perhaps surprisingly) increasing in the agency cost ξ. On the one hand, higher agency costs imply that, per unit of scale, a larger fraction of cash flows needs to go to the entrepreneur. On the other hand, higher agency costs reduce the counterfactual scale the entrepreneur can finance from financial investors under the dirty technology. This reduces the entrepreneur's outside option, making it cheaper for socially responsible investors to reform the firm. The latter effect dominates in our setup.

5 Conclusion

One of the major trends facing the investment management industry is a growing demand for "socially responsible" investing. How should the investment industry respond to this demand? Does this trend represent meaningless certification that allows investors to feel better about their investments? Or does it capture an actual demand for impact, such as changes in corporate policies that reduce carbon emissions, systemic risk, and other social costs?

This paper develops a parsimonious model to answer these questions. Conceptually, our analysis shows that co-investment by socially responsible investors can indeed have real impact, in the sense that it can induce firms to adopt cleaner production technologies, even when profit-motivated (financial) capital is abundant. Based on this conceptual analysis, our main practical contribution is the development of an investment criterion to optimally guide scarce socially responsible capital in an economy, the social profitability index (SPI).

The SPI summarizes the interaction of environmental, social and governance (ESG) aspects. Importantly, the SPI not only reflects the (social) return of the project that is being funded, but also the social costs or externalities that would have occurred in the absence of engagement by socially responsible investors. Accordingly, it can be optimal to invest in firms that generate relatively low social returns (e.g., a firm with significant carbon emissions), provided that the potential increase in social costs, if only financially-driven investors were to invest, is sufficiently large. This contrasts with many common ESG metrics that focus on firms' "social status quo" (i.e., on how green the company is at the moment). Most current ESG ratings are therefore not suited to achieve maximum impact.

The importance of counterfactual pollution in inducing changes in corporate policies implies that socially responsible fund need to follow a broader mandate than the maximization of returns subject to excluding polluting firms. As long as there is a large supply of competitive, profit-motivated capital, such a narrow mandate implies zero real impact on excluded firms (and poorer diversification for fund investors). The flip side of this insight is that, if socially responsible funds follow a broad mandate that unconditionally accounts for externalities, they must make a loss in financial terms (negative alpha). If this were not the case, competitive profit-motivated investors would have already funded these operational changes.

From a macro perspective, we uncover a complementarity between financial and socially responsible capital. Welfare is generally highest in an economy in which there is a balance between financial and socially responsible capital. The presence of profit-motivated financial capital alleviates underinvestment for a given production technology, precisely because financial investors do not internalize the negative externalities of production. However, this disregard for externalities can come at the cost of socially inefficient technology choice. The role of socially responsible investors is then to guide technology choice via co-investment. As a result, the composition of investor capital, not just its aggregate amount, matters in our setting.

A Proofs

Proof of Lemma 1. The Proof of Lemma 1 follows immediately from the proof of Theorem 1 given below. First, set $\gamma^{SR} = 0$ (so that socially responsible investors have the same preferences as financial investors). Second, to obtain the competitive financing arrangement (i.e., the agreement that maximizes the utility of the entrepreneur subject to the investors' participation constraint) one needs to choose the utility level of the entrepreneur u in (A.10) such that $\hat{v}_\tau K_\tau(u) - u = 0$.[7]

Proof of Theorem 1. The Proof of Theorem 1 will make use of Lemmas A.1 to A.5.

Lemma A.1 *In any solution to Problem 1, the IR constraint of financial investors, $pX^F - I^F \geq 0$ must bind,*

$$pX^F - I^F = 0. \tag{A.1}$$

Proof: The proof is by contradiction. Suppose there was an optimal contract for which $pX^F - I^F > 0$. Then, one could increase X^{SR} while lowering X^F by the same amount (until (A.1) holds). This perturbation strictly increases the objective function of socially responsible investors in (9), satisfies by construction the IR constraint of financial investors, whereas all other constraints are unaffected since $X = X^{SR} + X^F$ is unchanged. Hence, we found a feasible contract that increases the utility of socially responsible investors, which contradicts that the original contract was optimal. ■

Lemma A.2 *There exists an optimal financing arrangement with $I^F = X^F = 0$.*

Proof: Take an optimal contract $\left(I^F, I^{SR}, X^{SR}, X^F, K, c, \tau\right)$ with $I^F \neq 0$. Now consider the following "tilde" perturbation of the contract (leaving K, c and τ unchanged). Set $\tilde{X}^F$ and $\tilde{I}^F$ to 0 and set $\tilde{I}^{SR} = I^{SR} + I^F$ and $\tilde{X}^{SR} = X^{SR} + X^F$. The objective of socially

[7] Note that $\hat{v}_\tau = \pi_\tau - \gamma^E \phi_\tau$ in the special case when $\gamma^{SR} = 0$.

responsible investors in (9) is unaffected since

$$p\tilde{X}^{SR} - \tilde{I}^{SR} - \gamma^{SR}\phi_\tau K = pX^{SR} - I^{SR} + \underbrace{pX^F - I^F}_{0} - \gamma^{SR}\phi_\tau K \tag{A.2}$$

$$= pX^{SR} - I^{SR} - \gamma^{SR}\phi_\tau K, \tag{A.3}$$

where the second line follows from Lemma A.1. All other constraints are unaffected since $\tilde{X}^F + \tilde{X}^{SR} = X^F + X^{SR}$ and $\tilde{I}^F + \tilde{I}^{SR} = I^F + I^{SR}$ ∎

Lemma A.2 implies that we can phrase Problem 1 in terms of total investment I and total repayment to investors X in order to determine the optimal consumption c, technology τ, and scale K. However, to make the proof most instructive, it is useful to replace X and I as control variables and instead use the expected repayment to investors Ξ and expected utility provided to the entrepreneur u, which satisfy

$$\Xi := pX, \tag{A.4}$$

$$u := \left(pR - k_\tau - \gamma^E\phi_\tau\right) K + I - pX. \tag{A.5}$$

Then, using the definition $\hat{v}_\tau := \pi_\tau - \left(\gamma^E + \gamma^{SR}\right)\phi_\tau \geq v_\tau$, we can write Problem 1 as:

Problem 1^*

$$\max_{\tau} \max_{u \geq \bar{U}^E} \max_{K,\Xi} \hat{v}_\tau K - u \tag{A.6}$$

subject to

$$K \geq 0 \tag{A.7}$$

$$\Xi \leq (pR - \xi) K \tag{IC}$$

$$\Xi \geq -(A + u) + \left(pR - \gamma^E\phi_\tau\right) K \tag{LL}$$

Here, the last constraint (LL) can be interpreted as a limited liability constraint, since it refers to the constraint that upfront consumption is greater than zero (using the aggregate resource constraint in (2)). As the problem formulation suggests, it is useful to sequentially solve the optimization in 3 steps to exploit the fact that Ξ only enters the linear program via the constraints (IC) and (LL), but not the objective (A.6).

As is obvious from Problem 1*, only a technology that delivers positive surplus to investors and the entrepreneur (i.e., $\hat{v}_\tau > 0$) is a relevant candidate for the equilibrium technology.[8] Now consider the inner problem, i.e., for a fixed technology τ with $\hat{v}_\tau > 0$ and a fixed utility $u \geq \bar{U}^E$ we solve for the optimal vector (K, Ξ) as a function of τ and u.

Lemma A.3 *For any τ with $\hat{v}_\tau > 0$ and $u \geq \bar{U}^E$, the solution to the inner problem, i.e., $\max_{K,\Xi} \hat{v}_\tau K - u$ subject to (A.7), (IC) and (LL), implies maximal scale, i.e.,*

$$K_\tau (u) = \frac{A + u}{\xi - \gamma^E \phi_\tau} > 0. \tag{A.8}$$

The expected payment to investors is:

$$\Xi_\tau (u) = (pR - \xi) K_\tau (u) . \tag{A.9}$$

Proof: The feasible set for (K, Ξ) as implied by the three constraints (A.7), (IC) and (LL) forms a polygon (see orange region in Figure 1). The upper bound on Ξ in (IC) is an affine function of K through the origin (i.e., linear in K) whereas the lower bound in Equation (LL) is an affine function of K (with negative intercept $-(A+u)$). The slope of the lower bound in Equation (LL) is strictly greater than the slope of the upper

[8] Note that $\hat{v}_C$ is unambiguously positive whereas $\hat{v}_D$ could be negative or positive depending on whether the sum $\gamma^E + \gamma^{SR}$ is sufficiently close to 1.

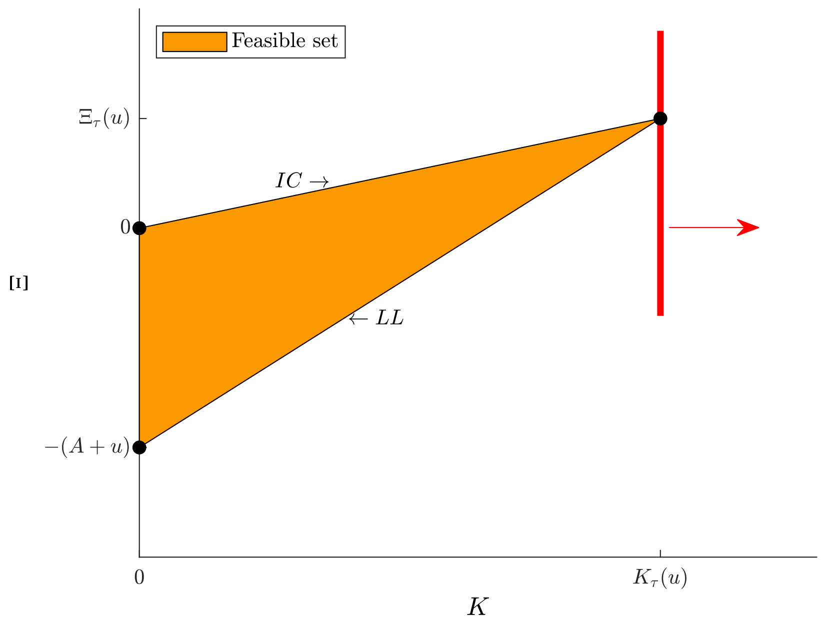

Figure 1. Feasible set of the inner problem: The set of feasible solutions is depicted in orange and forms a polygon. The objective function is represented by the red line and the arrow: The red line is a level set of the objective function of socially responsible investors, and the arrow indicates the direction in which we are optimizing.

bound in Equation (IC) since

$$\begin{aligned}\left(pR-\gamma^{E}\phi_{\tau}\right)-(pR-\xi) &= \xi-\gamma^{E}\phi_{\tau}\\ &> \pi_{\tau}-\gamma^{E}\phi_{\tau}\\ &> \pi_{\tau}-\left(\gamma^{E}+\gamma^{SR}\right)\phi_{\tau}=\hat{v}_{\tau}>0,\end{aligned}$$

where the second line follows from the finite scale that is implied by Assumption 1 (i.e., $\xi > \pi_\tau$). Therefore, the intersection of the upper bound (IC) and the lower bound in (LL) defines the maximal feasible scale of K. Choosing the maximal scale $K_\tau(u)$ is optimal, since for any given τ with $\hat{v}_\tau > 0$ and any fixed $u \geq \bar{U}^E$, the objective function $\hat{v}_\tau K - u$ is strictly increasing in K and independent of Ξ. The expression for $K_\tau(u)$ in Equation (A.8) is obtained from $(pR-\xi)K = -(A+u) + \left(pR-\gamma^E\phi_\tau\right)K$. ■

Given the solution to the inner problem, $(K_\tau(u), \Xi_\tau(u))$, we now turn to the optimal choice of u which maximizes $\hat{v}_\tau K_\tau(u) - u$ subject to $u \geq \bar{U}^E$.

Lemma A.4 *In any solution to Problem 1*, the entrepreneur obtains her reservation utility* $u = \bar{U}^E$.

Proof: It suffices to show that the objective is strictly decreasing in u. Using $K_\tau(u) = \frac{A+u}{\xi - \gamma^E \phi_\tau}$ and $\hat{v}_\tau = \pi_\tau - \left(\gamma^E + \gamma^{SR}\right)\phi_\tau$, we obtain that:

$$\hat{v}_\tau K_\tau(u) - u = \frac{\hat{v}_\tau}{\xi - \gamma^E \phi_\tau} A - \frac{\xi + \gamma^{SR}\phi_\tau - \pi_\tau}{\xi - \gamma^E \phi_\tau} u \tag{A.10}$$

Since $\xi > \pi_\tau$ and $\xi > \gamma^E \phi_\tau$ (both by Assumption 1), both the numerator and the denominator of $\frac{\xi + \gamma^{SR}\phi_\tau - \pi_\tau}{\xi - \gamma^E \phi_\tau}$ are positive, so that Equation (A.10) is strictly decreasing in u. ■

Given that $u = \bar{U}^E$ the optimal payoff to socially responsible investors *for a given* τ is given by:

$$U^{SR} = \hat{v}_\tau K_\tau\left(\bar{U}^E\right) - \bar{U}^E. \tag{A.11}$$

We now turn to the final step, i.e., the optimal technology choice.

Lemma A.5 *The optimal technology choice is given by:*

$$\hat{\tau} = \arg\max_\tau \frac{\hat{v}_\tau}{\xi - \gamma^E \phi_\tau}. \tag{A.12}$$

Proof: In the relevant case where $\hat{v}_D > 0$, we need to compare payoffs in (A.11). The clean technology is chosen if and only if $\hat{v}_C K_C\left(\bar{U}^E\right) > \hat{v}_D K_D\left(\bar{U}^E\right)$, which simplifies to (A.12). If $\hat{v}_D \leq 0$, then A.12 trivially holds as only $\hat{v}_C > 0$. ■

Lemmas A.3 to A.5, thus, jointly characterize the solution to Problem 1*, which, in turn, allows us to retrieve the solution to the original Problem 1. That is, substituting the expression for $\bar{U}^E$ in Equation (6) into $\hat{K} = K_{\hat{\tau}}\left(\bar{U}^E\right)$ yields Equation (11). Moreover, since (*LL*) binds, we obtain that $\hat{c} = 0$. The aggregate resource constraint in (2) then implies that total investment by both investors must satisfy $\hat{I} = \hat{K}k_{\hat{\tau}} - A$, whereas (IC) implies that $\hat{X} = \left(R - \frac{B}{\Delta p}\right)\hat{K}$. Since any agreement must satisfy $X^F + X^{SR} = \hat{X}$ and

$I^F + I^{SR} = \hat{I}$, we can trace out all possible agreements using the fact that financial investors break even (Lemma A.1), meaning that $pX^F - I^F = 0$ and $X^F \in [0, R]$.

Proof of Proposition 1. See discussion in main text.

Proof of Proposition 2. The social profitability index is defined as:

$$\text{SPI} = \frac{\Delta U^{SR}}{I^{SR}} \tag{A.13}$$

Using Theorem 1, we obtain that the maximum investment by socially responsible investors is given by

$$I^{SR}_{\max} = \hat{K} k_C - A. \tag{A.14}$$

The minimum investment that is sufficient to induce a change in production technology is given by

$$I^{SR}_{\min} = I^{SR}_{\max} - p\hat{X} = (\xi - \pi_C)\,\hat{K} - A. \tag{A.15}$$

This implies that SPI$\in [\text{SPI}_{\min}, \text{SPI}_{\max}]$ where $\text{SPI}_{\min} = \frac{\Delta U^{SR}}{I^{SR}_{\max}}$ and $\text{SPI}_{\max} = \frac{\Delta U^{SR}}{I^{SR}_{\min}}$. Now suppose that socially responsible investors require a cash flow share of λ, then

$$I^{SR} = I^{SR}_{\max} - (1-\lambda)\,p\hat{X} = \lambda I^{SR}_{\max} + (1-\lambda)\,I^{SR}_{\min}. \tag{A.16}$$

This yields expression (24) in Proposition 2.

Proof of Proposition 3. See discussion in main text.

Proof of Proposition 4. As a first step, we calculate the polar cases $\text{SPI}_{\max}$ and $\text{SPI}_{\min}$. For $\text{SPI}_{\max} = \frac{\Delta U^{SR}}{I^{SR}_{\min}}$ we obtain that:

$$\text{SPI}_{\max} = \gamma^{SR}\frac{\Delta\phi}{\Delta\pi - \frac{\gamma^E}{\xi}\left(\Delta\phi\left(\xi - \pi_C\right) + \Delta\pi\phi_C\right)} - 1 \tag{A.17}$$

which is increasing in $\Delta\phi$, ξ, and γ^E and decreasing in $\Delta\pi$ and λ.

For $\text{SPI}_{\min} = \frac{\Delta U^{SR}}{I^{SR}_{\max}}$ we obtain

$$\text{SPI}_{\min} = \frac{\gamma^{SR}\Delta\phi - \Delta\pi + \frac{\gamma^E}{\xi}\left(\Delta\phi\left(\xi - \pi_C\right) + \Delta\pi\phi_C\right)}{\Delta\pi + \left(pR - \xi\right)\frac{\xi - \gamma^E\phi_C}{\xi} - \frac{\gamma^E}{\xi}\left(\Delta\pi\phi_C + \Delta\phi k_C\right)}. \tag{A.18}$$

which is increasing in $\Delta\phi$, ξ, and γ^E and decreasing in $\Delta\pi$ and λ.

Finally, from the definition $\text{SPI} = \frac{1}{\frac{\lambda}{\text{SPI}_{\min}} + \frac{1-\lambda}{\text{SPI}_{\max}}}$, it is immediate that the SPI is decreasing in λ.

References

Barber, Brad M, Adair Morse, and Ayako Yasuda, "Impact investing," 2018. Working paper.

Chava, Sudheer, "Environmental externalities and cost of capital," *Management Science*, 2014, *60* (9), 2223–2247.

Chowdhry, Bhagwan, Shaun William Davies, and Brian Waters, "Investing for impact," *Review of Financial Studies*, 2018, *32* (3), 864–904.

Davies, Shaun William and Edward Dickersin Van Wesep, "The unintended consequences of divestment," *Journal of Financial Economics*, 2018, *128* (3), 558–575.

Glaeser, Edward L and Andrei Shleifer, "Not-for-profit entrepreneurs," *Journal of Public Economics*, 2001, *81* (1), 99–115.

Hart, Oliver and Luigi Zingales, "Companies Should Maximize Shareholder Welfare Not Market Value," *Journal of Law, Finance, and Accounting*, 2017, *2* (2), 247–275.

Heinkel, Robert, Alan Kraus, and Josef Zechner, "The effect of green investment on corporate behavior," *Journal of financial and quantitative analysis*, 2001, *36* (4), 431–449.

Holmström, Bengt and Jean Tirole, "Financial intermediation, loanable funds, and the real sector," *Quarterly Journal of Economics*, 1997, *112* (3), 663–691.

Hong, Harrison and Marcin Kacperczyk, "The price of sin: The effects of social norms on markets," *Journal of Financial Economics*, 2009, *93* (1), 15–36.

Morgan, John and Justin Tumlinson, "Corporate provision of public goods," *Management Science*, forthcoming.

Nauman, Billy, "Global bank agree on framework to promote green shipping," *Financial Times*, 2019, *June 18.*

Tett, Gillian, "Ethical investing has reaced a tipping point," *Financial Times*, 2019, *June 18.*

Tirole, Jean, *The theory of corporate finance*, Princeton University Press, 2006.