A Theory of Socially Responsible Investment *

Martin Oehmke[†] Marcus Opp[‡]

March 21, 2023

Abstract

We characterize the conditions under which a socially responsible (SR) fund induces firms to reduce externalities, even when profit-seeking capital is in perfectly elastic supply. Such impact requires that the SR fund's mandate permits the fund to trade off financial performance against reductions in social costs—relative to the counterfactual in which the fund does not invest in a given firm. Based on such an impact mandate, we derive a micro-founded investment criterion, the social profitability index (SPI), which characterizes the optimal ranking of impact investments when SR capital is scarce. If firms face binding financial constraints, the optimal way to achieve impact is by enabling a scale increase for clean production. In this case, SR and profit-seeking capital are complementary: Surplus is higher when both investor types are present.

Keywords: Socially responsible investing, ESG, Social Profitability Index (SPI), fiduciary duty, capital allocation, sustainable investment, sustainability ratings.

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

*For comments and suggestions, we thank Veronica Guerrieri (the editor), three anonymous referees, Franklin Allen, Ulf Axelson, Paul Beaumont, Roland Bénabou, Bruno Biais, Markus Brunnermeier, Patrick Bolton, Shaun Davies, Peter DeMarzo, Diego Garcia, Deeksha Gupta, Bengt Holmström, Anastasia Kartasheva, Lukasz Pomorski, Ailsa Roëll, Joel Shapiro, Lin Shen, Jan Starmans, Paul Woolley, and seminar participants at ASU Sonoran, BIS, Bonn, Cambridge, Copenhagen Business School, CU Boulder, Deutsche Bundesbank, EFA, École Polytechnique, Erasmus University, EWFC, Frankfurt School of Finance, the FTZ Seminar, HEC, Humboldt Universität zu Berlin, IDC Financial Economics Conference, IMF Macro-Financial Research Conference, INSEAD Finance Symposium, LSE, MIT OrgEcon lunch, OxFIT, Princeton, Schroders, Stanford, Stockholm School of Economics, UCLA, USC, Warwick Business School, and WFA. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement 715467) as well as the Marianne and Marcus Wallenberg (MMW) foundation.

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

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

1 Introduction

In recent years, the question of the social responsibility of business, famously raised by Friedman (1970), has re-emerged in the context of the spectacular rise of socially responsible (SR) investment. Assets under management in SR funds have grown manifold,[1] and many investors seek to augment their asset allocation with environmental, social, and governance (ESG) scores (Pastor, Stambaugh and Taylor, 2021; Pedersen, Fitzgibbons and Pomorski, 2021). While the financial performance of such investments has been explored (see, e.g., Hong and Kacperczyk, 2009; Chava, 2014; Barber, Morse and Yasuda, 2021), it is less clear whether the presence of SR funds has any real consequences for firm behavior. After all, firms have access to an (approximately) abundant supply of purely profit-seeking capital willing to finance activities irrespective of the associated externalities (Welch, 2014).

Understanding the real effects of SR investments requires taking a corporate finance view. To this end, we incorporate a SR fund and the choice between clean and dirty production into a standard model of corporate financing with abundant profit-seeking financial capital, building on Holmström and Tirole (1997). The model's main results are driven by the interaction of negative production externalities (which can lead to overinvestment in socially undesirable dirty production) and financing constraints (leading to underinvestment in socially desirable clean production). Such financing frictions are not only empirically relevant for young firms (an important source of clean innovation), but they also matter for mature firms that seek to replace profitable dirty production with more expensive clean production technologies.

We find that impact is possible even in the presence of abundant profit-seeking capital. However, to achieve impact a SR fund needs to sacrifice financial returns. Therefore, rather than following a traditional notion of fiduciary duty, a SR fund with an impact

[1] For example, the Global Sustainable Investment Alliance (2018) reports sustainable investing assets of \$30.7tn at the beginning of 2018, an increase of 34% relative to two years prior.

mandate needs to explicitly specify its desired trade-off between impact and financial performance. Given this desired trade-off, we derive an investment criterion for SR impact funds, the social profitability index (SPI), which characterizes the optimal allocation of scarce SR capital across heterogeneous firms. Because impact is about avoided pollution (as opposed to its level), investments in "sin" industries are not necessarily inconsistent with such a mandate. When financial constraints are binding, impact is optimally achieved by raising a firm's financing capacity under clean production beyond the amount that purely profit-motivated investors would provide. The increase in clean production (and, hence, total surplus) is larger when both investor types are present, reflecting a complementarity between profit-seeking and SR capital.

We develop these results in a parsimonious model, initially focusing on the investment decision of a single firm. The firm is owned by an entrepreneur with limited wealth, who has access to two production technologies, dirty and clean, both with constant returns to scale up to a threshold and zero returns thereafter (yielding a particularly simple form of decreasing returns to scale). Dirty production has a higher per-unit financial return, but clean production is socially preferable because it generates lower social costs. Production under either technology requires the entrepreneur to exert unobservable effort, so that not all cash flows are pledgeable to outside investors. The firm can raise funding from (up to) two types of outside investors. Financial investors have abundant capital and behave competitively. As their name suggests, they care exclusively about financial returns. In addition to financial returns, a SR fund's mandate accounts for the social costs generated by firms. We distinguish between two mandates. Under a *narrow mandate*, the SR fund's mandate incorporates the absolute level of social costs produced by the firms in its portfolio. Under a broader *impact mandate*, the SR fund's mandate incorporates social costs relative to a counterfactual scenario in which the SR fund does not invest in a given firm.

We first develop two benchmark cases. In the first, we consider a setting in which only financial investors are present. Because financial investors care about monetary payoffs only, the entrepreneur is more likely to be financially constrained under clean production and, conditional on being financially constrained, the maximum scale that the entrepreneur can obtain is larger under dirty production. As a result, the entrepreneur may adopt the socially inefficient dirty production technology, even if she partially internalizes the associated externalities so that she would choose the clean technology under self-financing. The second benchmark characterizes the planner's solution. When the firm is not financially constrained, the planner can implement the first-best allocation via a Pigouvian tax. In contrast, if the firm is financially constrained, a Pigouvian tax alone does not achieve first best and must be complemented with an upfront investment subsidy. This result reflects that regulation targeting only one source of inefficiency (externalities) without addressing the other (financing constraints) has limited effectiveness.

In practice, informational frictions and political economy constraints make it difficult for governments to implement the planner's solution (see Tirole, 2012). This motivates the main part of our analysis, which investigates whether and how a SR fund addresses these inefficiencies. Our model demonstrates that the SR fund has impact (i.e., changes the firm's technology choice) if and only if the fund's mandate places sufficient weight on the reduction in social costs that arises from the fund's investment. Under such an impact mandate, the SR fund internalizes the counterfactual social cost that would arise if a firm chose dirty production when seeking financing from financial investors only. This implies that the SR fund is willing to make a financial loss on its investment, which is necessary to achieve impact. In contrast, if the SR fund were to follow a narrow mandate that only incorporates the absolute level of social costs generated by firms in its portfolio, the fund would simply invest in firms that are clean anyway. In this case, dirty firms remain dirty and obtain financing from financial investors, so that the equilibrium allocation is unchanged relative to the benchmark case in which only financial investors are present.

The optimal financing agreement in the presence of the SR fund with an impact mandate can be implemented by issuing two bonds, a green bond purchased by the SR fund and a regular bond purchased by financial investors. In this implementation, the green bond is issued at a premium in the primary market, consistent with evidence in Baker, Bergstresser, Serafeim and Wurgler (2022) and Zerbib (2019). Alternatively, the optimal financing arrangement can be implemented with two share classes. In this case, the share class controlling the technology choice is issued at a premium. In both cases, the presence of the fairly priced security allows financial investors to break even, economizing on the capital contribution the SR fund needs to make.

If the firm is financially constrained under the clean technology and the SR fund has an impact mandate, the optimal way for the SR fund to achieve impact is to facilitate an increase in the scale of clean production. In this case, there is a complementarity between financial and SR capital: Total surplus (which, in our model, is determined by the total scale of clean production) is generally higher if both investor types are present. The complementarity arises because of financial investors' disregard for externalities, which allows dirty production at a larger scale than the entrepreneur could achieve under self-financing. The resulting threat of dirty production relaxes the participation constraint for the SR fund and, thereby, generates additional financing capacity. Since binding financial constraints imply that clean production is below the socially optimal scale, this additional financing capacity enables a surplus-enhancing increase in the scale of (socially valuable) clean production.

While SR capital has seen substantial growth over the last few years, it is likely that such capital remains scarce relative to financial capital that only chases financial returns. This raises the question of how scarce SR capital is invested most efficiently. A multi-firm extension of our model yields a micro-founded investment criterion from the perspective of a SR impact fund, the *Social Profitability Index* (SPI).

Similar to the (standard) profitability index, the SPI measures "bang for buck"— in this case, the payoff the SR fund generates under its mandate per unit of SR capital.

Unlike the conventional profitability index, the SPI not only reflects the (social) return of the project that is being funded, but also the counterfactual social costs that a firm would have generated in the absence of investment by the SR fund. Therefore, investment criteria for SR funds should include estimates of, say, carbon emissions that are avoided if the firm adopts a cleaner production technology. Because avoided externalities matter, it can be efficient for the SR fund to invest in firms that generate substantial social costs, as long as the SR fund's investment generates a sufficient reduction in those costs. Conversely, it is efficient for the SR fund not to invest in firms that are clean anyway. Such investments would use up scarce SR capital but generate no impact.

Given that impact requires that the SR fund sacrifices financial returns, would small individual investors ever contribute to such a fund? Our analysis highlights that this requires overcoming a free-rider problem if individual investors act in their self interest: Even though all small investors are affected by externalities arising from production by firms, they each rely on others to sacrifice financial returns. Hence, in the limit with infinitesimally small and identical investors, the SR fund cannot attract resources. Based on this negative benchmark, we then characterize conditions under which a SR impact fund can emerge in equilibrium. This is the case when some individual investors are disproportionately affected by the externality or when small investors are able to coordinate. Moreover, if individual investors obtain an additional warm-glow utility boost from "having done their part," thereby departing from purely self-interested behavior, the emergence of a SR fund with an impact mandate is facilitated. Finally, when the free-rider problem cannot be overcome by individuals, our analysis rationalizes the existence of state-owned funds that invest on behalf of their citizens.

Related Literature. The theoretical literature on socially responsible investing consists of two main strands: *exclusion* and *impact investing*. Following the pioneering paper by Heinkel, Kraus and Zechner (2001), the literature on *exclusion* studies the effects of investor boycotts, divestment, and portfolio tilting away from dirty firms. Whether the threat of exclusion impacts a firm's production decisions depends on the cost imposed on the firm by not being able to (fully) access capital from socially responsible investors. In most of this literature, exclusion increases the firm's cost of capital because the remaining investors demand higher risk premia to absorb the divested shares.[2] Edmans, Levit and Schneemeier (2022) highlight that unconditional divestment (to shrink the scale of dirty firms) can be dominated by a conditional threat of divestment (which incentivizes dirty firms to change their production technology).[3] Landier and Lovo (2020) consider a risk-neutral environment, in which divestment does not affect risk premia. Instead, the threat of divestment raises the firm's effective cost of capital because of a matching friction between firms and investors. In this setting, they analyze how to optimally achieve impact via the threat of divestment, including accounting for the emissions of suppliers.

Our model shuts down the exclusion channel by considering a risk-neutral environment with a perfectly elastic supply of profit-motivated capital. This setting captures that the impact of divestment on the cost of capital is likely to be small in competitive financial markets (see, e.g., Heinkel et al. (2001), Welch (2014), Broccardo, Hart and Zingales (2022), and Berk and van Binsbergen (2021)). Our paper, therefore, belongs to the second strand of the literature, which studies how *impact investors* can change firm behavior. Like most of this literature (see, e.g., Gollier and Pouget (2014), Chowdhry, Davies and Waters (2018), and Biais and Landier (2022)), we study the ability of a large SR fund to impact firm behavior and reduce externalities.[4] Rather than imposing costs on dirty firms via the threat of divestment (a "stick"), the SR fund in our model effectively subsidizes firms to adopt clean technologies (a "carrot"). One attractive feature

[2] See, e.g., Pastor, Stambaugh and Taylor (2021), Pedersen, Fitzgibbons and Pomorski (2021), De Angelis, Tankov and Zerbib (forthcoming), Broccardo, Hart and Zingales (2022), and Zerbib (2022).

[3] Davies and Van Wesep (2018) point out that blanket divestment can have other unintended consequences, such as inducing firms to prioritize short-term profit at the expense of long-term value.

[4] In contrast, Broccardo et al. (2022) study a setting in which being infinitesimal is of advantage. In particular, if the median investor in a firm has pro-social preferences and firm policies are governed by majority voting, small shareholders can achieve first best via voting.

of our framework is that it does not restrict attention to ad-hoc tools but instead takes an optimal contracting approach to solve for optimal engagement. When financing constraints are binding for clean firms, optimal engagement by the SR fund enables the firm to expand clean production relative to what profit-motivated investors would fund, a key ingredient for the complementarity between profit-motivated investors and the SR fund.[5]

In addition to highlighting the role of financial constraints, our paper makes several broader contributions that hold independent of whether financial constraints are binding. First, we show that when profit-seeking capital is abundant, impact requires that investors in a SR fund make financial sacrifices. Because impact does not come for free, it is essential that the objective of achieving impact and the desired trade-off between impact and financial performance are incorporated explicitly in the fund's mandate. Second, given an explicit impact mandate, our framework provides a micro-founded decision metric for the optimal allocation of scarce SR capital across firms (the SPI). Absent an explicit impact mandate, the SR fund will simply invest in firms that would have been clean regardless of the SR fund's investment. The result that investors without an explicit impact mandate may end up simply replacing profit-driven investors is robust beyond our specific modeling framework. In subsequent work, Green and Roth (2021) confirm this prediction using an assignment matching model. While our model does not consider competition between SR funds, Green and Roth (2021) show that funds without an explicit impact mandate end up competing for investments with impact-driven funds, further reducing impact and profitability.[6]

[5] Chowdhry et al. (2018) show that subsidies optimally take the form of investment by socially-minded activists if firms cannot credibly commit to pursuing social goals. There is no such commitment problem in our setting. Roth (2019) compares impact investing with grants, highlighting the ability of investors to withdraw capital as an advantage of investment over grants.

[6] Gupta, Kopytov and Starmans (2022) demonstrate that, in a dynamic setting, competition among SR investors can lead to a delay of abatement investments by polluting firms.

2 Model Setup

We study the role of socially responsible investing in a setting in which *production externalities* interact with *financing constraints*. Our analysis builds on the canonical model of corporate financing in the presence of agency frictions laid out in Holmström and Tirole (1997) and Tirole (2006). One key innovation of our framework is that it endogenizes the choice of production technologies, one of them "clean" (i.e., associated with low social costs), the other "dirty" (i.e., associated with higher social costs).

The entrepreneur, production, and moral hazard. We consider 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 $\tau \in \{C, D\}$. The technologies generate identical cash flows. Denoting firm scale by K, the firm generates positive cash flow of $R \cdot \min\left(K, \bar{K}\right)$ with probability p (conditional on effort by the entrepreneur, as discussed below) and zero otherwise. Both technologies therefore exhibit constant returns to scale up to $\bar{K}$ and no returns thereafter. This formulation captures decreasing returns to scale in the simplest possible fashion, while still maintaining the tractability of the Holmström and Tirole (1997) setup.[7]

While cash flows are identical, the technologies differ with respect to the required investment and the social costs they generate. Per unit of scale, the dirty technology D generates a negative (non-pecuniary) externality $\phi_D > 0$ and requires an upfront investment of k_D (also per unit). The clean technology results in a lower per-unit social cost $0 \leq \phi_C < \phi_D$, but requires a higher per-unit upfront investment $k_C > k_D$.[8] The entrepreneur internalizes a fraction $\gamma^E \in [0, 1)$ of social costs, capturing potential intrinsic motives not to cause social harm. In the special case $\gamma^E = 0$, the entrepreneur is

[7] In Online Appendix B, we discuss standard specifications of decreasing-returns-to-scale production functions and demonstrate robustness of our results to $N > 2$ production technologies.

[8] The assumption that $0 \leq \phi_C < \phi_D$ reflects that our analysis focuses on the mitigation of negative production externalities by a SR fund. We discuss the case of positive production externalities in Online Appendix B.

motivated purely by financial payoffs.

To generate a meaningful trade-off in the choice of technologies, we assume that the ranking of the two technologies differs depending on whether it is based on financial or social value. In the relevant region with positive returns ($K \leq \bar{K}$), the per-unit financial value of technology τ is given by $\pi_\tau := pR - k_\tau$, while the per-unit social value (or surplus) is $v_\tau := \pi_\tau - \phi_\tau$. We assume that the dirty technology creates higher financial value, $\pi_D > \pi_C$, but that clean production generates higher social value, $v_C > v_D$. These assumptions capture the idea that there exists a technology, here technology D, that increases profits relative to the socially optimal choice (here technology C) at the expense of higher social costs.[9] For ease of exposition, we initially assume that the social value of the dirty production technology is negative, $v_D < 0$, meaning that the externalities caused by dirty production outweigh its financial value.

As in Holmström and Tirole (1997), the entrepreneur is subject to an agency problem. Whereas the choice of production technology is assumed to be observable (and, hence, contractible), effort is assumed to be unobservable (and, therefore, not contractible). Under each technology, the investment pays off with probability p only if the entrepreneur exerts effort ($a = 1$). The payoff probability is reduced to $p - \Delta p$ if the entrepreneur shirks ($a = 0$), where $p > \Delta p > 0$. 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 firm's 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 $\xi := p\frac{B}{\Delta p}$*

[9] Once we allow for N technologies (see Online Appendix B), the dirtiest technology may no longer be the profit-maximizing technology. In this case, technology D corresponds to the profit-maximizing technology. The case where the profit-maximizing technology is also the cleanest technology is uninteresting for our analysis of SR investment, since even purely profit-motivated capital would ensure clean production in this case.

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 that the moral hazard problem alone ensures a finite production scale (even in the limit of constant returns to scale, i.e., $\bar{K} \to \infty$). The second inequality is a sufficient condition that rules out equilibrium shirking and ensures feasibility of outside financing. To streamline notation, π and v are defined assuming that the entrepreneur exerts effort (as usual, shirking is an off-equilibrium action).

Outside investors and securities. We assume that the entrepreneur's assets are not sufficient to fund the scale $\bar{K}$ under either technology, i.e., $A < \bar{K}k_D$, generating demand for outside financing. The entrepreneur can raise financing from (up to) two types of risk-neutral outside investors $i \in \{F, SR\}$, where F denotes a mass of competitive *financial investors* and SR denotes a *socially responsible fund*. As their name suggests, financial investors care exclusively about financial returns. In contrast, the SR fund's mandate also accounts for the social costs generated by firms by the firm, $\phi_\tau K$, with intensity γ^{SR}. We normalize $\gamma^{SR} + \gamma^E \leq 1$, so that jointly the SR fund and the entrepreneur do not internalize more than 100% of social costs.

Our analysis distinguishes between two types of objective functions (**mandates**) for the SR fund.

Definition 1 (The SR Fund's Mandate) *A SR fund has a **narrow mandate** if it accounts for the absolute level of social costs produced by firms in its portfolio. A SR fund has an **impact mandate** if it accounts for social costs relative to a counterfactual scenario in which the SR fund does not invest in a given firm. Under both mandates, we refer to the weight given to social costs, γ^{SR}, as the **social responsibility parameter**.*

The SR fund's mandate can be linked to different moral criteria. The impact mandate is essentially consequentialist with respect to social costs. In contrast, the narrow mandate is closer to a notion of direct responsibility that only arises if the fund has invested in the firm that produces the social cost.[10]

Regardless of the entrepreneur's source of financing, it is without loss of generality to restrict attention to financing arrangements in which the entrepreneur issues securities that pay a total 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 investors and the SR fund, respectively. Given that the firm has no resources in the low state, this security can be interpreted as debt or equity. The entrepreneur's utility can then be written as a function of the investment scale $K \leq \bar{K}$,[11] the total promised repayment X, the effort decision a, upfront consumption by the entrepreneur c, and the technology choice $\tau \in \{C, D\}$,

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

The first two terms of this expression, $p(RK - X) - (A - c)$, represent the project's net financial payoff to the entrepreneur under high effort, where $A - c$ can be interpreted as the upfront co-investment made by the entrepreneur. The third term, $\gamma^E \phi_\tau K$, measures the social cost internalized by the entrepreneur. The final term, $BK - \Delta p(RK - X)$, captures the incremental payoff conditional on shirking ($a = 0$). 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

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

[10] See Moisson (2020) and Dangl, Halling, Yu and Zechner (2023) for an analysis of how different moral criteria affect social preferences and outcomes.

[11] It is without loss of generality to restrict the equilibrium scale to $K \leq \bar{K}$. Given zero returns above $\bar{K}$, it is never optimal to pick a scale $K > \bar{K}$.

Per unit of scale, the entrepreneur's pledgeable income is therefore given by $pR - \xi$. The resource constraint at date 0 implies that capital expenditures, Kk_τ, must equal the total investments made by the entrepreneur and outside investors,

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

where I^F and I^{SR} represent the amounts invested by financial investors and the SR fund, respectively.

3 Benchmark Analysis

Our benchmark analysis consists of two parts. In Section 3.1, we show that if investors care exclusively about financial returns, the dirty technology may be chosen even if the entrepreneur has some concern for the higher social cost generated by dirty production (i.e., $\gamma^E > 0$). In Section 3.2, we analyze how a benevolent planner would address this inefficiency.

3.1 Financing from Financial Investors Only

The setting in which the entrepreneur can borrow exclusively from competitive financial investors corresponds to the special case $I^{SR} = X^{SR} = 0$. The entrepreneur's objective is then to choose a financing arrangement (consisting of scale $K \in \left[0, \bar{K}\right]$, promised repayment $X^F \in [0, R]$, upfront consumption $c \geq 0$, 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

$$U^F := pX^F - I^F \geq 0 \tag{IR}$$

As a preliminary step, it is useful analyze the financing arrangement that maximizes scale for a given technology τ absent technological limits (i.e., $\bar{K} \to \infty$). 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 $\frac{A}{\xi - \pi_\tau}$. This expression shows that the entrepreneur can scale her initial assets A by a factor that depends on the agency cost per unit of investment, $\xi := p\frac{B}{\Delta p}$, and the per-unit financial value under technology τ, π_τ. Because $\xi > \pi_D$ (see Assumption 1), the moral hazard problem alone ensures a finite scale of $\frac{A}{\xi - \pi_\tau}$ under either technology.

The comparison between this agency-induced scale limit $\frac{A}{\xi - \pi_\tau}$ and the technological limit $\bar{K}$ then determines whether a firm is financially constrained.

Definition 2 (Financing Constraints) *A firm is financially constrained for technology τ if and only if the entrepreneur's assets A are sufficiently low, $A < \bar{K}(\xi - \pi_\tau)$.*

The amount of liquid assets A required to eliminate financing constraints is higher for technology C, which is financially less profitable ($\pi_D > \pi_C$). Moreover, conditional on being financially constrained, $A < \bar{K}(\xi - \pi_C)$, the maximum scale that the entrepreneur can obtain from financial investors is larger under dirty production. In our continuous-scale framework, financing constraints therefore manifest themselves via a reduction in scale. We note that the loss of value due to suboptimal scale is economically equivalent to complete rationing of capital that would arise in a fixed-scale model with a binary investment decision.

The following lemma highlights that the entrepreneur's technology choice τ_F is then driven by a trade-off between achieving larger production scale and her concern for externalities. Of course, if the entrepreneur completely disregards externalities $\left(\gamma^E = 0\right)$, no trade-off arises and the entrepreneur always chooses the more profitable dirty production technology.

Lemma 1 (Benchmark: Financial Investors Only) *If only financial investors are present, the entrepreneur chooses technology τ that maximizes her utility*

$$\underline{U}^E = \max_{\tau}(\pi_\tau - \gamma^E \phi_\tau) K_\tau^F. \tag{3}$$

where

$$K_\tau^F := \min\left\{\frac{A}{\xi - \pi_\tau}, \bar{K}\right\}. \tag{4}$$

According to Lemma 1, if financing is raised from financial investors only, the entrepreneur chooses the technology τ_F that maximizes her payoff, which is given by the product of the per-unit payoff to the entrepreneur (financial NPV net off internalized social costs) and K_τ^F. Maximum scale (up to $\bar{K}$) is optimal because, under the equilibrium technology τ_F, the project generates positive surplus for the entrepreneur and financial investors. It follows that the entrepreneur adopts the dirty technology whenever

$$(\pi_D - \gamma^E \phi_D) K_D^F > (\pi_C - \gamma^E \phi_C) K_C^F. \tag{5}$$

Given that the dirty technology is financially more profitable, $\pi_D > \pi_C$, and the scale is larger under the dirty technology, $K_D^F \geq K_C^F$, this condition is satisfied whenever the entrepreneur's concern for externalities γ^E lies below a strictly positive cutoff $\bar{\gamma}^E$.

Corollary 1 (Benchmark: Conditions for Dirty Production) *If only financial investors are present, the entrepreneur adopts the dirty production technology if and only if $\gamma^E < \bar{\gamma}^E := \frac{\pi_D K_D^F - \pi_C K_C^F}{\phi_D K_D^F - \phi_C K_C^F}$.*

Corollary 1 implies that the entrepreneur may choose the dirty technology when financing from financial investors is available, even if she were to choose the clean technology under self-financing.[12]

3.2 The Planner's Problem

As a second benchmark, we characterize the solution to the planner's problem. In our setting, welfare is defined as the total surplus created by production (including social costs),

$$\Omega := \min\left\{K, \bar{K}\right\} \cdot v_\tau. \tag{6}$$

First-best welfare is achieved by choosing the socially optimal technology C and producing at the socially optimal scale $K = \bar{K}$ (given that $v_C = \pi_\tau - \phi_\tau > 0$).

Going forward, we focus on the interesting case in which the laissez-faire equilibrium with financial investors only (see Lemma 1) does not achieve first-best welfare. For ease of exposition, we also set $\phi_C = 0$ for the remainder of this section.

Proposition 1 (Planner's Solution) *The solution to the planner's problem is as follows.*

1. *If the firm is financially unconstrained under the clean technology, $A < \bar{K}\left(\xi - \pi_C\right)$, first-best welfare can be achieved by a Pigouvian tax of ϕ_τ per unit of scale.*

2. *Otherwise, a Pigouvian tax alone cannot achieve first best, but needs to be complemented with an investment subsidy of $\bar{K}\left(\xi - \pi_C\right) - A$.*

If financial constraints do not bind, the planner's only concern is to ensure the correct technology choice. A Pigouvian tax is then sufficient to render dirty production less profitable than clean production. The entrepreneur then responds by adopting the clean technology and, because financial constraints do not bind, can raise sufficient funds from

[12] Because the entrepreneur is constrained under self-financing, $A < k_D\bar{K}$, she prefers the clean technology if and only if $\frac{A}{k_C}\left(\pi_C - \gamma^E\phi_C\right) \geq \frac{A}{k_D}\left(\pi_D - \gamma^E\phi_D\right)$. Hence, the entrepreneur is "corrupted" by financial markets when $\gamma^E \in \left(\tilde{\gamma}^E, \bar{\gamma}^E\right)$ where $\tilde{\gamma}^E := \frac{k_C\pi_D - k_D\pi_C}{k_C\phi_D - k_D\phi_C}$.

capital markets to achieve the socially efficient scale $\bar{K}$. Note that, in this setting, banning the dirty technology would be equivalent to a Pigouvian tax.[13]

If, instead, financial constraints are binding for the clean technology, a Pigouvian tax of ϕ_τ (or banning technology D) would achieve the correct technology choice, but would fail to address the underinvestment problem that arises due to financial constraints. To achieve first best, the regulator now needs to additionally subsidize clean production by an amount of $\bar{K}\,(\xi - \pi_C) - A$. This investment subsidy could be provided through an equity injection (which the firm uses to raise additional funds from financial investors) or via a subsidized loan.

For simplicity, we have ignored the potential social costs of subsidies, which could arise, for example, from the deadweight costs of taxes required to finance the subsidy. In the presence of such costs, it would be necessary to trade off the costs of the subsidy against the social benefits of increased clean production. Even in our simple setting, the information required to calibrate such a subsidy would demand expertise that is typically associated with private investors, such as understanding of agency rents, profitability, and efficient production scales.[14]

4 Investment by a Socially Responsible Fund

We now turn to our main question: whether and how a SR fund impacts the firm's investment decision (in the absence of optimal government policies). Section 4.1 develops our main results in a single-firm setting, assuming that socially responsible capital is abundant relative to the funding needs of the firm. In Section 4.2, we consider a multi-firm setting to investigate how scarce socially responsible capital should be allocated across firms.

[13] If $\phi_C > 0$, a Pigouvian tax is no longer equivalent to banning the dirty technology because, in addition to reducing the profitability of the dirty technology, the tax would also tighten financial constraints (see proof of Proposition 1).

[14] These information requirements make it difficult to implement the optimal policy, even if there is no lack of political willpower (see, e.g., Tirole, 2012).

4.1 Single-Firm Analysis

In contrast to financial investors, the SR fund's mandate incorporates not only financial payoffs X^{SR} but also social costs $\phi_\tau K$. The extent to which social costs are incorporated depends both on the fund's mandate $M \in \{N, I\}$ (see Definition 1) and the associated social responsibility parameter γ^{SR},

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

$$U_N^{SR} = pX^{SR} - I^{SR} - \gamma^{SR}\phi_\tau K \cdot \mathbb{1}_{I^{SR}>0}. \tag{U_N^{SR}}$$

Accordingly, a SR fund with an impact mandate the fund internalizes social costs independent of whether the fund has invested in the company. As a result, the fund accounts for incremental social costs relative to the counterfactual scenario of not investing in the firm. In contrast, under a narrow mandate the fund internalizes the absolute level of social costs, but only if it has invested in the firm. It is useful to note that even under an impact mandate with full internalization of social costs ($\gamma^E + \gamma^{SR} = 1$), the SR fund's objective does not coincide with the planner's objective. The reason is that the SR fund does not internalize rents that accrue to the entrepreneur. We view this as a realistic restriction on the SR fund's objective, consistent with plausible preferences for the fund's investors (see Section 5).[15]

4.1.1 Optimal Financing Arrangement with a SR Fund

We now analyze whether and how the financing arrangement and the resultant technology choice are altered when a SR fund is present. Because the entrepreneur could still raise financing exclusively from financial investors, the utility she receives under the financing arrangement with financial investors only, $\underline{U}^E$ given in Equation (3), now becomes the entrepreneur's outside option. If the SR fund remains passive, $I^{SR} = 0$, its payoff under

[15] If the SR fund's objective accounted for those rents, its objective would be equivalent to the planner's problem discussed in Section 3.2.

an impact mandate is given by

$$\underline{U}_I^{SR} = -\gamma^{SR}\phi_{\tau_F} K_{\tau_F}^F < 0. \tag{7}$$

This expression, which acts as the SR fund's reservation utility under an impact mandate, accounts for the social costs generated when the entrepreneur raises financing exclusively from financial investors and chooses technology τ_F and scale $K_{\tau_F}^F$ (see Lemma 1). In contrast, under a narrow mandate, the SR fund's reservation payoff is unaffected by the social costs generated if the SR fund does not invest, so that $\underline{U}_N^{SR} = 0$. The dependence of the SR fund's outside option on its mandate plays a key role for our results.

To generate Pareto improvements relative to their respective outside options $\underline{U}_M^{SR}$ and $\underline{U}^E$, the SR fund can engage with the entrepreneur and agree on a financing contract that specifies the technology τ, scale K, as well as the required financial investments and cash flow rights for all investors and the entrepreneur. For ease of exposition, we give all the bargaining power to the SR fund, so that the optimal bilateral agreement maximizes the payoff to the socially responsible fund subject to the entrepreneur's outside option. In the appendix, we show that all of our main results are unaffected by the specific assumption regarding who has the bargaining power.

Problem 1 (Optimal Bilateral Agreements) *Given a mandate M, the SR fund's objective is*

$$\max_{I^F, I^{SR}, X^{SR}, X^F, K, c, \tau} U_M^{SR} \tag{8}$$

subject to the entrepreneur's IR constraint:

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

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

Constraint IR^E ensures that the entrepreneur receives at least as much as she would under her outside option of raising financing exclusively from financial investors, $\underline{U}^E$. Note that the above 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 (as suggested by Harstad, 2012).

Proposition 2 (Technology and Scale with a SR Fund) *The equilibrium technology choice and scale depend on the SR fund's mandate:*

1. *If the SR fund has a narrow mandate, the equilibrium technology choice and scale are identical to the benchmark equilibrium described in Lemma 1.*

2. *Let* $\hat{v}_\tau := \pi_\tau - \left(\gamma^E + \gamma^{SR}\right)\phi_\tau \geq v_\tau := \pi_\tau - \phi_\tau$ *denote bilateral surplus (per unit of scale) for the SR fund and the entrepreneur. If the SR fund has an impact mandate, the equilibrium technology choice is given by*

$$\hat{\tau} = \arg\max_{\tau} \hat{v}_\tau \hat{K}_\tau, \tag{9}$$

where the scale given technology τ *satisfies*

$$\hat{K}_\tau = \min\left\{\frac{A + \underline{U}^E}{\xi - \gamma^E \phi_\tau}, \bar{K}\right\}. \tag{10}$$

Proposition 2 contains the main theoretical result of the paper. First, it shows that a SR fund with a narrow mandate has no impact. The reason that, under a narrow mandate, the SR fund can avoid "responsibility for pollution" simply by not investing. Moreover, since financial investors provide financing at competitive terms under both technologies, there is no way for the SR fund to extract financial rents. Hence, under a narrow mandate, it is strictly optimal for the SR fund not to invest in firms that generate social costs ($\phi > 0$). As a result, the firm obtains the same financing terms as in the benchmark case, in which the SR fund is not present.

Because the outcome under a narrow mandate is the same as under the benchmark model without a SR fund, in what follows we focus on a SR fund with an impact mandate. Under an impact mandate, the equilibrium technology choice $\hat{\tau}$ maximizes total bilateral surplus accruing to the SR fund and the entrepreneur, which is given by the product of the per-unit surplus $\hat{v}_\tau$ and the production scale $\hat{K}_\tau$. As long as the entrepreneur is financially constrained under the financing arrangement with a SR fund, the offered scale, $\frac{A+\underline{U}^E}{\xi-\gamma^E\phi_\tau}$, ensures that the entrepreneur earns the same utility as her outside option $\underline{U}^E$. (In the absence of binding financial constraints, the equilibrium scale is equal to the unconstrained scale $\bar{K}$).

While the optimal financing arrangement uniquely pins down the production side (i.e., technology choice and scale), there exists a continuum of co-investment arrangements between financial investors and the SR fund that solve Problem 1. This indifference arises because any increase in cash flows accruing to financial investors, $\hat{X}^F$, translates at competitive terms into higher upfront investment by financial investors, $\hat{I}^F$.

Corollary 2 (Optimal Co-investment Arrangements) *For any total payout to investors $\hat{X}$, the set of optimal co-investment arrangements between financial investors and the SR fund can be obtained by tracing out the cash-flow share accruing to the SR fund $\lambda \in [0,1]$ and setting $\hat{X}^{SR} = \lambda\hat{X}$, $\hat{X}^F = (1-\lambda)\hat{X}$, $\hat{I}^F = p\hat{X}^F$ and $\hat{I}^{SR} = \hat{I} - \hat{I}^F$. Pledged income and upfront consumption satisfy:*

$$p\hat{X} = \max\left\{\left(pR - \gamma^E\phi_{\hat{\tau}}\right)\hat{K}_{\hat{\tau}} - \left(A + \underline{U}^E\right), 0\right\}, \tag{11}$$

$$\hat{c} = A + \underline{U}^E - \left(pR - \gamma^E\phi_{\hat{\tau}}\right)\hat{K}_{\hat{\tau}} + p\hat{X} \geq 0. \tag{12}$$

If the firm is financially constrained, so that $\hat{K}_{\hat{\tau}} = \frac{A+\underline{U}^E}{\xi-\gamma^E\phi_{\hat{\tau}}}$, the entrepreneur optimally co-invests all her wealth, $\hat{c} = 0$, and the financing arrangement exhausts the entrepreneur's pledgeable income, $p\hat{X} = (pR - \xi)\hat{K}_{\hat{\tau}}$. The only indeterminacy in this case is the cash flow share accruing to the SR fund and financial investors, respectively. If the firm is not financially constrained, $\hat{K}_{\hat{\tau}} = \bar{K}$, the entrepreneur can raise more financing than needed to finance scale $\bar{K}$. Since pledgeable income is no longer a constraining factor, either the income pledged to investors $p\hat{X}$ lies below the incentive-compatible maximum or the entrepreneur consumes upfront. In Equation (11), we make the assumption that the entrepreneur initially co-invests all her wealth, $\hat{c} = 0$, in return for a reduction in pledged income. Only once entrepreneurial assets are sufficiently high, such that $p\hat{X} = 0$, the reduction in pledged income must be supplemented with strictly positive upfront consumption, $\hat{c} > 0$. This extreme outcome can be interpreted as a pure grant (without cash flow rights).

There are two particularly intuitive ways in which the optimal financing arrangement characterized in Proposition 2 and Corollary 2 can be implemented.[16]

Corollary 3 (Implementation) *The following securities implement the optimal financing agreement under an impact mandate:*

1. Green bond and regular bond: *The entrepreneur issues two bonds with respective face values $\hat{X}^F$ and $\hat{X}^{SR}$ at prices $\hat{I}^F$ and $\hat{I}^{SR}$. The green bond contains a technology-choice covenant specifying technology $\hat{\tau}$.*

2. Dual-class share structure: *The entrepreneur issues voting and non-voting shares, where shares with voting rights yield an issuance amount of $\hat{I}^{SR}$ in return for control rights and a fraction λ of dividends. The remaining proceeds $\hat{I}^F$ are obtained in return for non-voting shares with a claim on a fraction $1 - \lambda$ of dividends.*

[16] Under both implementations, the security targeted at the SR fund is issued at a premium in the primary market (see Corollary 5 below), ensuring that only the SR fund has an incentive to purchase this security. If the technology choice cannot be contracted upon (due to incomplete contracts), the green bond implementation may be dominated by a dual-class share structure.

4.1.2 Impact

To shed light on the economic mechanism behind Proposition 2, this section provides a more detailed investigation of the case in which the SR fund has impact, which we define as an induced change in the firm's production decision, through a switch in technology from $\tau_F = D$ to $\hat{\tau} = C$ and/or a change in production scale.[17] Based on Proposition 2, the following corollary summarizes the conditions for impact.

Corollary 4 (Impact) *Suppose $\gamma^E < \bar{\gamma}^E$, so that the firm chooses the dirty technology when raising financing from financial investors only. Then, the SR fund has impact if and only if it follows an impact mandate with a sufficiently high social responsibility parameter, $\gamma^{SR} \geq \bar{\gamma}^{SR}$, where the threshold $\bar{\gamma}^{SR}$ is decreasing in γ^E.*

Impact therefore requires that the SR fund follows and explicit impact mandate and places sufficient weight on the reduction in social costs that arises from the fund's investment ($\gamma^{SR} \geq \bar{\gamma}^{SR}$). If the entrepreneur and the SR fund jointly internalize all externalities, $\gamma^E + \gamma^{SR} = 1$, production will always be clean, because bilateral surplus coincides with total surplus (i.e., $\hat{v}_C = v_C > 0 > v_D = \hat{v}_D$).

Complementarity between financial and SR capital. If the conditions for impact are satisfied, the equilibrium of our model features a complementarity between financial investors and the SR fund. This complementarity results not from co-investment by both types of investors but by the presence of both types of capital.

Proposition 3 (Complementarity) *Suppose the conditions for impact are satisfied:*

1. *If assets are below a cutoff so that both K_C^F and K_C^{SR} are below $\bar{K}$, financial and SR capital act as complements: The equilibrium clean scale with both investor types,*

[17] If investment by the SR fund does not result in a change in production technology compared to the benchmark case (i.e., $\hat{\tau} = \tau_F$), there is no impact. In this case, we obtain the same scale, $\hat{K}_{\hat{\tau}} = K_{\tau_F}^F$, and utility for all agents in the economy as in the benchmark case. This less interesting situation occurs if the entrepreneur adopts the clean production technology even in the absence of investment by the SR fund, or if the entrepreneur adopts the dirty technology irrespective of whether the SR fund provides funding.

$\hat{K}_C$, is larger than the clean scale that can be financed in an economy with only one of the two investor types,

$$\hat{K}_C > \max\left\{K_C^F, K_C^{SR}\right\}. \tag{13}$$

2. Otherwise, there is no complementarity and $\hat{K}_C = \bar{K} = \max\left\{K_C^F, K_C^{SR}\right\}$.

Intuitively, if the clean technology is not subject to financial constraints, the only relevant inefficiency is the wrong technology choice. Impact is then achieved via a Coasian transfer (e.g., upfront consumption) to induce the entrepreneur to switch the technology. Equilibrium scale is not affected and there is no complementarity. In contrast, if the clean technology is subject to financial constraints, the presence of the SR fund leads to both a change in the production technology and an increase in scale. In this case, the equilibrium clean scale in the presence of both investor types strictly exceeds the scale that is attainable with only one investor type.

Consider first why the equilibrium clean scale with both investors exceeds the maximum clean scale that can be funded by financial investors, $\hat{K}_C > K_C^F$. If $\gamma^E < \bar{\gamma}^E$, a clean scale of K_C^F is not large enough to induce clean production if only financial investors are present. As shown in Corollary 1, in this case the entrepreneur prefers dirty production at scale K_D^F. Therefore, to induce the entrepreneur to switch to the clean production technology, the SR fund needs to inject additional resources into the firm. Due to the moral hazard friction and the resultant underinvestment problem, this capital injection is optimally used to increase the scale of clean production above and beyond what financial investors are willing to offer, so that $\hat{K}_C > K_C^F$.

Perhaps more surprisingly, $\hat{K}_C$ also exceeds the scale that could be financed if only the SR fund were present. The reason is that financial investors' disregard for externalities allows dirty production at a larger scale than the entrepreneur could achieve under self-financing (i.e., if no financial investors are around). The resulting pollution threat relaxes the participation constraint for the SR fund, through its effect on their reservation utility, $\underline{U}^{SR} = -\gamma^{SR}\phi_D K_D^F$. This unlocks additional financing capacity, so that $\hat{K}_C > K_C^{SR}$. Because clean production is socially valuable, Proposition 3 implies that total surplus, $v_C \hat{K}_C$, is strictly higher if both financial investors and the SR fund deploy capital, relative to the case in which all capital is allocated one investor type.

Abstracting from specific modeling details, *two basic ingredients* are necessary for the complementarity between the two investor types to arise. First, there must be underinvestment in the clean technology. Second, the SR fund needs to internalize social costs relative to the counterfactual of not investing in the firm (the impact mandate). Because the SR fund internalizes this counterfactual, the threat of dirty production (enabled by financial investors) acts as a *quasi asset* to the firm, generating additional financing capacity from the SR fund. Because of underinvestment (the first ingredient), the additional financing from the SR fund results in an increase in clean scale, which is socially valuable.

Whether this complementarity is present matters for (additional) government intervention. In particular, when the complementarity arises—binding financing constraints and a SR fund with impact mandate—the introduction of a Pigouvian tax would strictly reduce welfare.[18] By eliminating the threat of dirty production, the key ingredient for additional clean financing capacity from the SR fund is lost. Of course, if the planner were to choose the optimal policy in the presence of financial constraints, a Pigouvian tax accompanied with an investment subsidy, see Proposition 1, first-best could be achieved regardless of whether a SR fund is present or not.

[18] The result that Pigouvian taxes generally do not achieve first best in the presence of financial constraints echoes the findings of Hoffmann, Inderst and Moslener (2017) and Inderst and Heider (2022). Most closely related, Inderst and Heider (2022) show that in an industry equilibrium building on Holmström and Tirole (1997), optimal regulation depends on whether financial constraints bind in aggregate.

The cost of impact. Even though the SR fund only invests if doing so increases its utility relative to the case in which it remains passive,

$$\Delta U^{SR} := \hat{v}_C \hat{K}_C - \hat{v}_D K_D^F > 0, \tag{14}$$

the SR fund does not break even in financial terms.

Corollary 5 (Impact Requires a Financial Loss) *Impact (a switch from $\tau_F = D$ to $\hat{\tau} = C$) requires that a SR fund makes a financial loss. That is, in any optimal financing arrangement as characterized in Proposition 2,*

$$p\hat{X}^{SR} - \hat{I}^{SR} = \left(\pi_C - \gamma^E \phi_C\right) \hat{K}_C - \left(\pi_D - \gamma^E \phi_D\right) K_D^F < 0. \tag{15}$$

A SR fund with a narrow mandate breaks even financially but has no impact.

Intuitively, to induce a change from dirty to clean production, the SR fund must offer an agreement consisting of scale for the clean technology and upfront consumption that would not be offered by competitive financial investors. Because financial investors just break even, the SR fund must make a financial loss. The financial loss to the SR fund reflects the loss in bilateral surplus for financial investors and the entrepreneur relative to their preferred agreement, which yields a joint payoff of $\left(\pi_D - \gamma^E \phi_D\right) K_D^F$. If the entrepreneur is purely profit-motivated ($\gamma^E = 0$) she needs to be compensated for the loss of total profits, $\pi_C \hat{K}_C - \pi_D K_D^F$.

Empirically, Corollary 5 predicts that SR funds with impact must have a negative alpha and, conversely, that SR funds that generate weakly positive alpha do not generate impact. Our model also predicts that the financial loss for the SR fund, $p\hat{X}^{SR} - \hat{I}^{SR}$, occurs at the time when the firm seeks financing in the primary market, consistent with evidence on the at-issue pricing of green bonds in Baker et al. (2022) and Zerbib (2019). However, if the SR fund were to sell its cash flow stake $\hat{X}^{SR}$ after the firm has financed the clean technology, our model does not predict a price premium for the green security in the secondary market (i.e., in the secondary market, the security would be fairly priced at $p\hat{X}^{SR}$).[19]

4.2 The Social Profitability Index

We now derive a micro-founded investment criterion for allocation of scarce socially responsible capital from the perspective of a SR fund with impact mandate. To do so, we extend the single-firm analysis presented in Section 4 to a multi-firm setting with limited socially responsible capital, denoted by κ^{SR}. We endogenize the capitalization of the SR fund in Section 5. We initially continue to assume that financial capital is abundant.

The economy consists of a continuum of infinitesimal firms grouped into distinct firm types.[20] Firms that belong to the same type j are identical in terms of all relevant parameters of the model, whereas firms belonging to distinct 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 the SR fund is given by

$$\int_{\gamma_j^E < \bar{\gamma}_j^E} \phi_{D,j} K_{D,j}^F d\mu(j) + \int_{\gamma_j^E \geq \bar{\gamma}_j^E} \phi_{C,j} K_{C,j}^F d\mu(j). \tag{16}$$

The first term of this expression captures the social cost generated by firms that, in the absence of the SR fund, 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 the SR fund ($\gamma_j^E \geq \bar{\gamma}_j^E$).

[19] In our static model, control (or a technology covenant) matters only once, at the time of the initial investment. In a dynamic setting, control could matter multiple times (whenever investment technologies are chosen).

[20] The assumption that firms are infinitesimally small rules out well-known difficulties that arise when ranking investment opportunities of discrete size.

Given this aggregate social cost, how should a SR fund with impact mandate allocate its limited capital? One direct implication of Proposition 2 is that any investment in firm types with $\gamma_j^E \geq \bar{\gamma}_j^E$ cannot be optimal for the SR fund, because these firms adopt the clean technology even when raising financing exclusively from competitive financial investors. For the remaining firm types, an impact mandate with social responsibility parameter γ^{SR} implies that the SR fund receives the following payoff from reforming a firm of type j:

$$\Delta U_j^{SR} = \left(\pi_{C,j} - \gamma_j^E\right) \hat{K}_{C,j} - \left(\pi_{D,j} - \gamma_j^E\right) K_{D,j}^F + \gamma^{SR} \left(\phi_{D,j} K_{D,j}^F - \phi_{C,j} \hat{K}_{C,j}\right) \tag{17}$$

Here, $\left(\pi_{C,j} - \gamma_j^E\right) \hat{K}_{C,j} - \left(\pi_{D,j} - \gamma_j^E\right) K_{D,j}^F < 0$, captures the financial loss required to induce a firm of type j to adopt the clean production technology. The remaining term, $\gamma^{SR} \left(\phi_{D,j} K_{D,j}^F - \phi_{C,j} \hat{K}_{C,j}\right) > 0$, captures the mandate-implied benefit associated with the resulting reduction in social costs.

Given limited capital κ^{SR}, the SR fund is generally not able to reform all firms. To optimally fulfill its mandate, it should therefore prioritize investments in firm types that maximize the mandate-implied payoff per dollar invested. This is achieved by ranking firms according to a variation on the classic profitability index, the *social profitability index* (SPI).[21] The SPI is the ratio of the incremental payoff, as defined by its mandate, the SR fund generates by reforming a firm, ΔU_j^{SR}, and the amount of capital the SR fund needs to invest to reform the firm, I_j^{SR},[22]

$$\mathrm{SPI}_j = \mathbb{1}_{\gamma_j^E < \bar{\gamma}_j^E} \frac{\Delta U_j^{SR}}{I_j^{SR}}. \tag{18}$$

[21] The profitability index yields a consistent ranking of investments if there is a single resource constraint and if the scarce resource is completely exhausted (see Berk and DeMarzo, 2020). In our setting, the single resource constraint is the total amount of SR capital κ^{SR}. SR capital is fully exhausted because firms are of infinitesimal size.

[22] The change in the payoff to the SR fund ΔU_j^{SR} is the same across all financing agreements characterized in Proposition 2. Absent other constraints, it is therefore optimal for the SR fund to choose the minimum co-investment that implements clean production.

Proposition 4 (The Social Profitability Index (SPI)) *A SR fund with an impact mandate ranks firms according to the social profitability index, SPI_j. There exists a threshold $SPI^* \left(\kappa^{SR}\right) \geq 0$ such that a SR fund with scarce capital κ^{SR} invests in all firms for which $SPI_j \geq SPI^* \left(\kappa^{SR}\right)$.*

According to Proposition 4, it is optimal to invest in firms with the highest SPI until no funds are left, which happens at the cutoff $\text{SPI}^* \left(\kappa^{SR}\right)$. SR capital is scarce if and only if the amount κ^{SR} is not sufficient to reform all firm types with $\text{SPI}_j > 0$.

The SPI links the attractiveness of an investment for the SR fund to the underlying model parameters, thereby shedding light on the types of investments that the SR fund should prioritize.

Proposition 5 (SPI Comparative Statics) *As long as $\gamma_j^E < \bar{\gamma}_j^E$, the SPI is increasing in the avoided social cost, $\Delta\phi_j := \phi_{D_j} - \phi_{C_j}$, and the entrepreneur's concern for social cost, γ_j^E, and decreasing in the financial cost associated with switching to the clean technology, $\Delta\pi_j := k_{C,j} - k_{D,j}$.*

Proposition 5 states that SR funds with an impact mandate should prioritize firms for which avoided social cost $\Delta\phi_j$ is high. Note that, because the SPI reflects difference in social costs, it can be optimal for the SR fund to invest in firms that generate significant social costs, provided that these firms would have caused even larger social costs in the absence of engagement by the SR fund. The avoided social cost $\Delta\phi_j$ has to be traded off against the associated financial costs, as measured by the reduction in financial profits $\Delta\pi_j$.

The ranking of investments implied by SPI also has implications for the assortative matching between the social-mindedness of entrepreneurs and SR capital (see also Green and Roth, 2021).[23] As long as the SR fund is needed to generate impact, $\gamma_j^E < \bar{\gamma}_j^E$, there

[23] Our analysis assumes that the entrepreneur's social preference is observable (e.g., inferred from past decisions). In future work, it could be interesting to analyze the effects of unobservable social preferences on the optimal financing agreement, so as to ensure truth-telling.

is a form of positive assortative matching: The SR fund optimally prioritizes firms with more socially-minded entrepreneurs because they generate larger bilateral surplus and require a smaller investment from the SR fund to become clean. However, as soon as the entrepreneur internalizes enough of the externalities so that she chooses the clean technology even if financed by financial investors (i.e., $\gamma_j^E \geq \bar{\gamma}_j^E$), the SPI drops discontinuously to zero. It is inefficient for the SR fund to invest in these firms.

To obtain a closed-form expression for the SPI, it is useful to consider the special case $\gamma^E = 0$ and $\gamma^{SR} = 1$. Moreover, while strictly speaking it is optimal to minimize the SR fund's investment by assigning all cash-flow rights to financial investors, suppose that the SR fund needs to receive a fraction λ_j of a firm's cash flow rights. This minimum cash-flow stake then pins down I_j^{SR}.[24] Given these assumptions, the SPI is given by,

$$\mathrm{SPI}_j = \frac{\Delta\phi_j - \Delta\pi_j}{\Delta\pi_j + \lambda_j \min\left\{p_j R_j - \xi_j, k_{D,j} - \frac{A_j}{\bar{K}_j}\right\}}. \tag{19}$$

This expression reveals the intuitive trade-off between the two main ingredients of the SPI, avoided pollution $\Delta\phi_j$ and foregone profits $\Delta\pi_j$. If financial constraints bind, then $\mathrm{SPI}_j = \frac{\Delta\phi_j - \Delta\pi_j}{\Delta\pi_j + \lambda_j(p_j R_j - \xi_j)}$. In this case, the SPI implies that firms with tighter financial constraints should be prioritized (where, following Tirole (2006), financial constraints are measured by lower unit-pledgeable income $p_j R_j - \xi_j$). If firms are not financially constrained, $\mathrm{SPI}_j = \frac{\Delta\phi_j - \Delta\pi_j}{\Delta\pi_j + \lambda_j\left(k_{D,j} - A_j/\bar{K}_j\right)}$. In this case, firms with more liquid assets (higher A) should be prioritized. This happens because these firms can contribute more of their own resources, whereas their pollution threat is capped at $\phi_D \bar{K}$ (and therefore independent of A).

To conclude this section, we analyze how the composition of investor capital (and

[24] The assumption of a required cash-flow stake for the SR fund can be justified on two grounds. First, it is natural that investors in the SR fund cannot rely purely on utility derived from the non-pecuniary benefits of reducing social costs, but require a certain amount of financial payoffs alongside non-pecuniary payoffs. Second, the minimum cash flow share λ_j can be interpreted as a reduced form representation of the control rights that are necessary to implement ensure that firm j implements the clean technology.

not simply its aggregate amount) matters for total surplus, motivated by the recent growth in ESG investing. Increasing the amount of capital deployed by the SR fund does not automatically translate into higher welfare. The reason is that the ranking of investments implied by the SPI does not necessarily coincide with the planner's ranking, even if $\gamma_j^E + \gamma^{SR} = 1$. Even though the SR fund's payoff from reforming a firm, ΔU_j^{SR}, coincides with the associated welfare change, $v_C \hat{K}_C - v_D K_D^F$, the planner would increase scale up to the efficient scale $\bar{K}$, which is strictly larger than the scale funded by the SR fund if the firm is financially constrained post reform, $\hat{K}_{C,j} < \bar{K}$. This wedge arises because the SR fund does not internalize rents that accrue to the entrepreneur. Therefore, the allocation implemented by the SR fund coincides with the planner's solution only if the firm is financially unconstrained post reform, $\hat{K}_C = \bar{K}$. Binding financial constraints introduce a wedge between the planner's solution and the allocation implemented by the SR fund.[25]

The change in total surplus relative to the case without the SR fund, $\Delta\Omega$, results from the set of reformed firms (i.e., firms with $\gamma_j^E < \bar{\gamma}_j^E$ and $SPI_j \geq SPI^*\left(\kappa^{SR}\right)$). We can therefore write the change in total surplus as

$$\Delta\Omega = \int_{j:\gamma_j^E < \bar{\gamma}_j^E \;\&\; SPI_j \geq SPI^*(\kappa^{SR})} \left(v_{C,j} \hat{K}_{C,j} - v_{D,j} K_{D,j}^F \right) d\mu(j). \tag{20}$$

We then immediately obtain

Lemma 2 *Assume that financial capital is fixed and abundant. Aggregate welfare is increasing in the amount of SR capital* κ^{SR}.

Intuitively, increasing the level of SR capital has strictly positive welfare effects if it reduces externalities (that would have been financed by financial investors) and increases the scale of clean production for the set of reformed, financially constrained firms. Be-

[25] A corollary of this statement is that, if all firms are financially unconstrained, the planner's ranking of investments coincides with the ranking implied by the SPI.

cause financial capital is abundant, this positive effect is not driven by the (trivial) reason that there is more capital in the economy.

We now fix the total amount of capital in the economy and investigate the conjecture that increasing the *fraction* of SR capital, denoted by x^{SR}, is always welfare enhancing. Perhaps surprisingly, even if all externalities are accounted for (i.e., $\gamma^{SR} + \gamma^{E} = 1$) this conjecture is not generally true.

Proposition 6 (Composition of Capital) *Assume that aggregate capital is fixed and abundant. If financial constraints are absent, $K^{SR}_{C,j} = K^{F}_{C,j} = \bar{K}_j$ for all firm types j, welfare is maximized for $x^{SR} = 1$. Otherwise, it may be optimal from a welfare perspective that a strictly positive fraction of capital is deployed by financial investors, $x^{SR} < 1$.*

Recall that first-best welfare requires that both the correct technology C and the efficient scale $\bar{K}_j$ be chosen. If financial constraints do not bind, the only concern is whether the correct technology C is chosen, which the SR fund will ensure for all firms when $x^{SR} = 1$ (since $\gamma^{SR} + \gamma^{E} = 1$). Because clean production is already at the efficient scale, the only effect of an increase in the fraction of financial capital is that, eventually, this will induce some firms to switch to dirty production. This happens once the fraction of socially responsible capital is too low to ensure that all firms adopt the clean technology.

In contrast, if a sufficient fraction of firms operates below the optimal scale $\bar{K}_j$ when all capital is held by the SR fund, we essentially obtain an aggregate version of the complementarity result given in Proposition 3. An increase in financial capital provides firms with the outside option of producing dirty at larger scale. This threat, in turn, unlocks additional financing capacity by the SR fund, enabling a welfare-improving scale increase of clean production. Thus, in the presence of binding financial constraints, the right balance between socially responsible and financial capital is important.

5 Delegation to a SR Fund

So far we have focused on the decisions of a (large) SR fund with a given capital endowment, highlighting the importance of the fund's mandate for generating impact (Corollary 4). However, given that an impact mandate entails a financial loss (Corollary 5), the question arises whether a SR fund with an impact mandate can obtain capital from (small) individual investors.

For ease of exposition, we investigate this question using a special case of the setup in Section 4.2, with a continuum of identical firms of mass 1, owned by profit-motivated entrepreneurs (i.e., $\gamma^E = 0$). There are many investors who only care about firm cash flows, so that, as before, financial capital is abundant. Given $\gamma^E = 0$, all firms adopt the socially inefficient dirty technology in the absence of a SR fund with impact mandate.

Rather than taking the endowment of the SR impact fund with social responsibility parameter γ^{SR} as given, we now assume that there are n small investors. Small investors are self-interested and only care about the externality to the extent that it affects them personally. We assume that each small investor bears a fraction γ^i of that aggregate externality, so that

$$\sum_{i=1}^{n} \gamma^i = 1. \tag{21}$$

Each investor i has total funds κ_i, which can be allocated to (i) a SR fund with impact mandate and a social responsibility parameter $\gamma^{SR} \geq \bar{\gamma}^{SR}$, (ii) competitive profit-maximizing funds, and (iii) a storage technology offering zero net return. As long as $\gamma^{SR} \geq \bar{\gamma}^{SR}$, the SR fund wants to reform all firms provided that it has sufficient capital. In contrast, a SR fund with social responsibility parameter $\gamma^{SR} < \bar{\gamma}^{SR}$ would not want to reform firms and, therefore, would behave like a profit-maximizing fund.

Let $\kappa_i^{SR} \in [0, \kappa_i]$ denote the total amount that investor i contributes to the SR fund. The total endowment of the SR fund is then given by $\kappa^{SR} = \sum_{i=1}^{n} \kappa_i^{SR}$. If the SR fund reforms a fraction ω of firms, the aggregate externality is given by $\omega \phi_C \hat{K}_C +$

$(1-\omega)\,\phi_D K_D^F$. If $\kappa^{SR} > \pi_D K_D^F - \pi_C \hat{K}_C$, (see Equation (15) with $\gamma^E = 0$), the SR fund has sufficient capital to reform all firms, so that $\omega = 1$. Otherwise, only a fraction $\omega = \frac{\kappa^{SR}}{\pi_D K_D^F - \pi_C \hat{K}_C}$ of firms can be reformed. As a result, we can write the fraction of reformed firms as

$$\omega = \min\left\{\frac{\kappa^{SR}}{\pi_D K_D^F - \pi_C \hat{K}_C}, 1\right\}. \tag{22}$$

Given that both the storage technology and profit-maximizing funds offer zero net return in equilibrium, investor i's payoff is

$$U^i = \frac{\kappa_i^{SR}}{\kappa^{SR}}\omega\left[\pi_C \hat{K}_C - \pi_D K_D^F\right] - \gamma^i\left[\omega\phi_C \hat{K}_C + (1-\omega)\,\phi_D K_D^F\right]. \tag{23}$$

The first term captures investor i's share of the loss incurred by the SR fund to reform a fraction ω of firms. The second term captures the effect of the aggregate externality on individual i's utility.

We now determine investor i's optimal individual allocation in the SR fund, $\hat{\kappa}_i^{SR}$, given a total contribution by other investors of κ_{-i}^{SR}. It follows from (22) and (23) that, if $\kappa_{-i}^{SR} > \pi_C \hat{K}_C - \pi_D K_D^F$, the SR fund is sufficiently capitalized to reform all firms ($\omega = 1$) regardless of whether investor i contributes. In this case, it is optimal for investor i not to invest in the SR fund ($\hat{\kappa}_i^{SR} = 0$) because she would participate in the SR fund's financial loss without generating any additional reduction in the aggregate externality.[26] In contrast, when $\omega < 1$, investor i's contribution to the SR fund generates a reduction in the aggregate externality. Investor i then trades off the loss from contributing to the SR fund against the additional reduction in the externality, resulting in a total payoff of

$$U^i = -\kappa_i^{SR} - \gamma^i \phi_D K_D^F + \gamma^i \frac{\phi_D K_D^F - \phi_C \hat{K}_C}{\pi_D K_D^F - \pi_C \hat{K}_C}\left(\kappa_i^{SR} + \kappa_{-i}^{SR}\right), \tag{24}$$

where $-\kappa_i^{SR}$ represents the financial loss from investing in the SR fund and $\gamma^i \phi_D K_D^F$

[26] This follows directly from (23) evaluated at $\omega = 1$.

the aggregate externality absent reform. In the third term, $\gamma^i \frac{\phi_D K_D^F - \phi_C \hat{K}_C}{\pi_D K_D^F - \pi_C \hat{K}_C} \kappa_i^{SR}$ captures the reduction in the externality due to investor i's investment, whereas $\gamma^i \frac{\phi_D K_D^F - \phi_C \hat{K}_C}{\pi_D K_D^F - \pi_C \hat{K}_C} \kappa_{-i}^{SR}$ captures that agent i benefits reduction in the externality resulting from the contribution of other agents to the SR fund.

We first provide a negative benchmark result for the non-cooperative allocation of capital to the SR fund, which builds on the large literature on the private provision of public goods (see, e.g., Samuelson (1954) and Bergstrom, Blume and Varian (1986)).

Result 1 (Free-rider Problem) *Suppose investors are symmetric, $\gamma^i = \frac{1}{n}$. For n sufficiently large, there exists no Nash equilibrium in which investors allocate funds to the SR impact fund, i.e., $\kappa^{SR} = 0$.*

Intuitively, because the benefits of investment by the SR fund (reduced externalities) are non-rival and non-excludable, each individual investor only partly internalizes the social benefits. If this internalization is sufficiently small, $\gamma^i = \frac{1}{n} < \frac{\pi_D K_D^F - \pi_C \hat{K}_C}{\phi_D K_D^F - \phi_C \hat{K}_C}$, no individual investor contributes to the SR fund. Note that this condition is more likely to be satisfied for diffuse externalities that affect a large number of individuals.

Turned on its head, Result 1 also characterizes settings in which individual investors will provide sufficient capital to the SR fund. One such situation is when exposure to the externality is asymmetric.

Corollary 6 (Asymmetric Exposure) *Suppose one agent internalizes the externalities to a sufficient degree, $\gamma^i \geq \frac{\pi_D K_D^F - \pi_C \hat{K}_C}{\phi_D K_D^F - \phi_C \hat{K}_C}$. Then a SR impact fund has a positive equilibrium endowment $\kappa^{SR} > 0$.*

Clearly, the effect of asymmetric exposure is particularly relevant if the investor whose utility is most affected by externalities (high γ^i) is also wealthy (high κ_i). One example is the Breakthrough Energy Catalyst (BEC) fund by the Gates Foundation, which invests in climate-friendly technologies that would otherwise not be financially viable (see Financial Times, 2022).

Following a similar logic, suppose that a subset of agents $n_1 \leq n$ is able to coordinate. Even when individual investors are small, such coordination can ensure that at least part of the social cost is internalized via a SR fund.

Corollary 7 (Coordination) *Suppose a subset n_1 of agents coordinate and that* $\sum_{i=1}^{n_1} \gamma^i \geq \frac{\pi_D K_D^F - \pi_C \hat{K}_C}{\phi_D K_D^F - \phi_C \hat{K}_C}$. *Then a SR impact fund has a positive equilibrium endowment* $\kappa^{SR} > 0$.

Taken together, Corollaries 6 and 7 show that when the agents are consequentialist (i.e., their utility depends on aggregate impact via κ^{SR}), effective size (either via asymmetry or coordination) is a necessary condition for a SR fund with impact mandate to emerge. However, if individual investor decisions are, in addition, subject to a warm-glow utility boost from "having done their part,"(see, e.g., Andreoni, 1990)), even "effectively small" investors may choose to contribute.[27]

Corollary 8 (Warm Glow) *Suppose investors experience an additional warm-glow utility boost of* $w^i \kappa_i^{SR}$ *from their own investment in a SR impact fund. Then an investor contributes to the fund if and only if* $w^i \geq 1 - \gamma^i \frac{\phi_D K_D^F - \phi_C \hat{K}_C}{\pi_D K_D^F - \pi_C \hat{K}_C}$.

Corollary 8 states that the free-rider problem is mitigated if, in addition to the impact generated by the SR fund, individual investors care directly about how much they have contributed to the fund. One interesting implication of this result is that, in order to ensure sufficient capitalization of a SR fund that acts consequentialist when achieving impact, it helps if individual investors are non-consequentialist. This finding is consistent with results obtained by Landier and Lovo (2020) in a different framework.[28]

In some cases, individual investors may not be able to overcome the free-rider problem. In such situations, our model rationalizes the existence of state-owned funds that invest on

[27] Riedl and Smeets (2017), Hartzmark and Sussman (2019), and Bonnefon, Landier, Sastry and Thesmar (2019) provide evidence that individual investor behavior is consistent with warm-glow utility arising from socially responsible investment decisions.

[28] Consistent with Broccardo et al. (2022) and Inderst and Opp (2022), we assume that the warm-glow utility component reflects "decisional utility" that affects individual decisions but does not enter welfare.

behalf of their citizens (like the Norwegian sovereign wealth fund). Direct investment by a sovereign fund circumvents the free-rider problem that would arise if governments paid out their resource income and, thereby, left investment decisions to individual citizens.[29] In fact, under some circumstances, citizens would vote for the establishment of a SR sovereign fund (see, e.g., Broccardo et al. (2022)) because it provides a commitment device not to free-ride on externality-reducing investments.

Finally, we note that governments could help reduce the free-rider problem by taxing the returns of SR funds with impact mandate at a lower rate (see also Nguyen, Rivera and Zhang, 2021). Advantageous tax treatment would partially offset the lower pre-tax returns generated by impact funds. The resulting change in relative after-tax returns would have similar effects to warm-glow utility.

6 Conclusion

A key question in today's investment environment is to understand conditions under which socially responsible investment can achieve impact. To shed light on this question, this paper develops a parsimonious theoretical framework, based on the interaction of production externalities and corporate financing constraints.

Our analysis uncovers the importance of an explicit *impact mandate* for socially responsible funds. Given an abundant supply of profit-motivated capital, it is not enough for SR funds to simply invest in firms that generate a small absolute level of social costs. Rather, social costs must be accounted for relative to the counterfactual social costs that would arise when not investing in a given firm. The necessity of an impact mandate generates both normative and positive implications. From a positive perspective, our model implies that as most current ESG funds lack such a broad impact mandate, they do not have impact. From a normative perspective, it states that, if society wants SR funds to

[29] This idea is related to Morgan and Tumlinson (2019) who provide a model in which shareholders value public goods but are subject to free-rider problems. This free-rider problem can be overcome if, instead of paying dividends, the company invests on behalf of shareholders.

have impact, then their mandate needs to violate a traditional notion of fiduciary duty, because achieving impact requires sacrificing financial returns. Building on the idea of "what gets measured gets managed," our results further suggest that socially responsible funds need to be evaluated according to broader measures, explicitly accounting for real impact rather than focusing solely on financial metrics.

From a practical investment perspective, our model implies a micro-founded investment criterion for scarce socially responsible capital, the *social profitability index* (SPI). In line with the impact mandate, the SPI accounts for social costs that would have occurred in the absence of engagement by a socially responsible fund. 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. While conceptually intuitive, the implementation of the SPI requires relatively detailed knowledge of the production process within a given industry, in order to be able to estimate avoided social costs as well as the associated financial sacrifice. Estimating the SPI using increasingly detailed data available on emissions and production technologies is a potentially fruitful avenue for future research.

To highlight the key ideas in a transparent fashion, our model abstracts from a number of realistic features which could be analyzed in future work. First, our model considers a static framework. In a dynamic setting, a number of additional interesting questions would arise: How to account for dirty legacy assets? How to ensure the timely adoption of novel (and cleaner) production technologies as they arrive over time? Because the adoption of future green technologies may be hard to contract ex ante, a dynamic theory might yield interesting implications on the issue of control. Second, our model considers the natural benchmark case in which individual investors have the same directional social preferences (e.g., to lower carbon emissions). More challenging is the case in which socially responsible investors' objectives conflict or are multi-dimensional (e.g., there is agreement on the goal of lowering carbon emissions, but disagreement on the social costs imposed by nuclear energy). Finally, we excluded the possibility that firms interact as part of a supply chain or as competitors (as in Dewatripont and Tirole, 2020). For example, when the adoption of the clean technology by one firm crowds out dirty production by other firms, this generates additional benefits from the perspective of the SR fund, which, in turn, would increase the fund's willingness to finance clean production. It would be interesting to study such spillovers in future work.

A Proofs

Proof of Lemma 1: We present this proof as a special case of the proof of Proposition 2 given below. Set $\gamma^{SR} = 0$, so that the SR fund has the same preferences as financial investors and $\hat{v}_\tau = \pi_\tau - \gamma^E \phi_\tau$. To obtain the competitive financing arrangement (i.e., the agreement that maximizes the entrepreneur's utility u subject to the investors' participation constraint), set u such that $\hat{v}_\tau K^*_\tau(u) - u = 0$, using Equation (A.15). ■

Proof of Corollary 1: The result follows directly from a comparison of the net payoff to the entrepreneur, $\underline{U}^E$, in the presence of financial investors only under the clean and dirty technology, based on Equation (3) in Lemma 1. ■

Proof of Proposition 1: The proof of this proposition covers the general case $\phi_C \geq 0$. Therefore, the proof also applies to the special case $\phi_C = 0$ considered in the benchmark section. Consider a Pigouvian tax that is equal to the marginal social cost generated by technology τ per unit of capital, ϕ_τ. Then the after-tax profit for the dirty technology (per unit of capital) is strictly negative (i.e., $\pi_D - \phi_D < 0$) so that the dirty technology will not be adopted by the firm. We now distinguish two cases.
Case 1: If $A \geq \bar{K}(\xi - \pi_C + \phi_C)$, the firm can finance the efficient scale $\bar{K}$ for the clean technology by raising financing from financial investors, taking in to account the associated tax $\phi_C \bar{K}$. This follows from Equation (4) adjusted for "after-tax"assets $\tilde{A} = A - \bar{K}\phi_C$. This proves the first statement of Proposition 1.
Case 2: If $A < \bar{K}(\xi - \pi_C + \phi_C)$, Equation (4) implies that the efficient scale cannot be achieved when raising financing from financial investors. A subsidy of (at least) $s = \bar{K}(\xi - \pi_C + \phi_C) - A > 0$ is required for the entrepreneur to finance a scale of $\bar{K}$. This proves the second statement of Proposition 1. ■

Proof of Proposition 2: The proof of Proposition 2 proceeds separately for the two mandates $M \in \{N, I\}$ of the SR fund.

Narrow Mandate: If $M = N$, the objective function of the SR fund is given by

$$U_N^{SR} = pX^{SR} - I^{SR} - \gamma^{SR}\phi_\tau K \cdot \mathbb{1}_{I^{SR}>0} \leq 0. \tag{A.1}$$

The inequality follows from two ingredients. First, due to competitive pricing by financial investors, the net financial payoff for the SR fund, $pX^{SR} - I^{SR} \leq 0$ is bounded above by zero (for any technology τ). Second, the externality term satisfies $-\gamma^{SR}\phi_\tau K \cdot \mathbb{1}_{I^{SR}>0} \leq 0$ with strict equality if $I^{SR} = 0$ or $\phi_\tau = 0$ (or both). The maximum total payoff of $U_N^{SR} = 0$ is then achieved by setting $I^{SR} = 0$. Non-investment is strictly optimal for the SR fund if $\tau_F = D$ (in which case the entrepreneur needs to be subsidized financially to switch to the clean technology) or if the clean technology has a positive social cost, $\phi_C > 0$. If $\tau_F = C$ and $\phi_C = 0$, then the SR fund may co-invest at competitive terms and would get the same total payoff (zero) as under non-investment. In either case, the equilibrium scale and production technology is the same as in the benchmark equilibrium with financial investors only.

Impact Mandate: The proof makes use of Lemmas A.1 to A.5. As discussed in the main text, we prove our statements for a general bargaining procedure: With probability η, the entrepreneur gets to make a take-it-or-leave-it offer, giving her the maximum payoff, denoted by $\overline{U}^E$, while the SR fund remains at its reservation utility $\underline{U}^{SR}$. With probability $1-\eta$, the SR fund gets to make a take-it-or-leave-it offer, leading to the analogous respective payoffs of $\overline{U}_I^{SR}$ and $\underline{U}^E$ (these payoffs are derived in Equations (A.19) and (A.20), respectively.) The analysis in the main text considers the special case $\eta = 0$. Following Hart and Moore (1998), we augment this bargaining game by allowing the SR fund to make an offer before the above bargaining game starts. Then, for a given surplus division parameter η, we obtain

Problem 1* *Under an impact mandate, the SR fund's problem is*

$$\max_{I^F, I^{SR}, X^{SR}, X^F, K, c, \tau} pX^{SR} - I^{SR} - \gamma^{SR}\phi_\tau K, \tag{A.2}$$

subject to the entrepreneur's IR constraint given bargaining power η,

$$U^E\left(K, X^{SR} + X^F, \tau, c, 1\right) \geq (1-\eta)\,\underline{U}^E + \eta\bar{U}^E, \tag{A.3}$$

as well as the entrepreneur's IC constraint, the resource constraint (2), the financial investors' IR constraint, the non-negativity constraints $K \geq 0, c \geq 0$, and the technological constraint $K \leq \bar{K}$.

Lemma A.1 *In any solution to Problem 1*, the financial investors' IR constraint must bind,*

$$pX^F - I^F = 0. \tag{A.4}$$

Proof: The proof is by contradiction. Suppose there were 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 Equation (A.4) holds). This perturbation strictly increases the SR fund's objective function under an impact mandate (A.2) and satisfies (by construction) the financial investors' IR constraint. All other constraints are unaffected because $X = X^{SR} + X^F$ is unchanged. Hence, we have found a feasible contract that increases the utility of the SR fund, contradicting that the original contract was optimal. ■

Lemma A.2 *There exists an optimal financing arrangement without participation of financial investors, i.e.,* $I^F = X^F = 0$.

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 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 SR fund's objective (A.2) is unaffected since

$$p\tilde{X}^{SR} - \tilde{I}^{SR} - \gamma^{SR}\phi_\tau K = pX^{SR} - I^{SR} + \underbrace{pX^F - I^F}_{0} - \gamma^{SR}\phi_\tau K \tag{A.5}$$

$$= pX^{SR} - I^{SR} - \gamma^{SR}\phi_\tau K, \tag{A.6}$$

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 express Problem 1* in terms of total investment I and the total promised repayment to investors X in order to determine the optimal consumption c, technology choice τ, and scale K. To make the proof instructive, it is useful to replace X and I as control variables by the expected repayment to investors Ξ and the expected utility provided to the entrepreneur u, which satisfy

$$\Xi := pX, \tag{A.7}$$

$$u := \left(\pi_\tau - \gamma^E\phi_\tau\right) K + I - pX. \tag{A.8}$$

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 a sequential maximization problem:

Problem 1^{**}

$$\max_{\tau} \max_{u \geq \eta\bar{U}^E + (1-\eta)\underline{U}^E} \max_{K,\Xi} \hat{v}_\tau K - u \tag{A.9}$$

subject to

$$K \geq 0 \tag{A.10}$$

$$K \leq \bar{K} \tag{A.11}$$

$$\Xi \geq -(A + u) + \left(pR - \gamma^E\phi_\tau\right) K \tag{A.12}$$

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

$$\Xi \geq 0 \tag{LL}$$

Constraint (A.12) ensures that upfront consumption is weakly greater than zero, $c \geq 0$, using the definition of u in (A.8) and the aggregate resource constraint (2). Constraint (LL) ensures that the security offers limited liability to investors by guaranteeing a weakly positive expected payoff (this constraint will be irrelevant for the determination of equilibrium scale and technology). As the problem formulation suggests, it is useful to sequentially solve the optimization in three steps to exploit that Ξ only enters the linear program via the constraints (A.12), (LL), and (IC) but not the objective (A.9).

It is clear from Problem 1^{**} that 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.[30] We now consider the inner problem: For a fixed technology τ with $\hat{v}_\tau > 0$ and a fixed utility $u \geq \eta \bar{U}^E + (1-\eta)\,\underline{U}^E$, we solve for the optimal vector (K, Ξ) as a function of τ and u.

Lemma A.3 *For any technology τ with $\hat{v}_\tau > 0$ and $u \geq \eta \bar{U}^E + (1-\eta)\,\underline{U}^E$, the solution to the inner problem, i.e., $\max_{K,\Xi} \hat{v}_\tau K - u$ subject to (A.10), (A.11), (A.12), (IC) and (LL) implies a maximum scale*

$$K^*_\tau(u) = \min\left\{\frac{A+u}{\xi - \gamma^E \phi_\tau}, \bar{K}\right\} > 0. \tag{A.13}$$

The minimum expected repayment to investors is

$$\Xi_\tau(u) = \max\left\{\left(pR - \gamma^E \phi_\tau\right) K^*_\tau(u) - (A+u), 0\right\}. \tag{A.14}$$

Proof: The feasible set for (K, Ξ) as implied by the five constraints (A.10), (A.11),

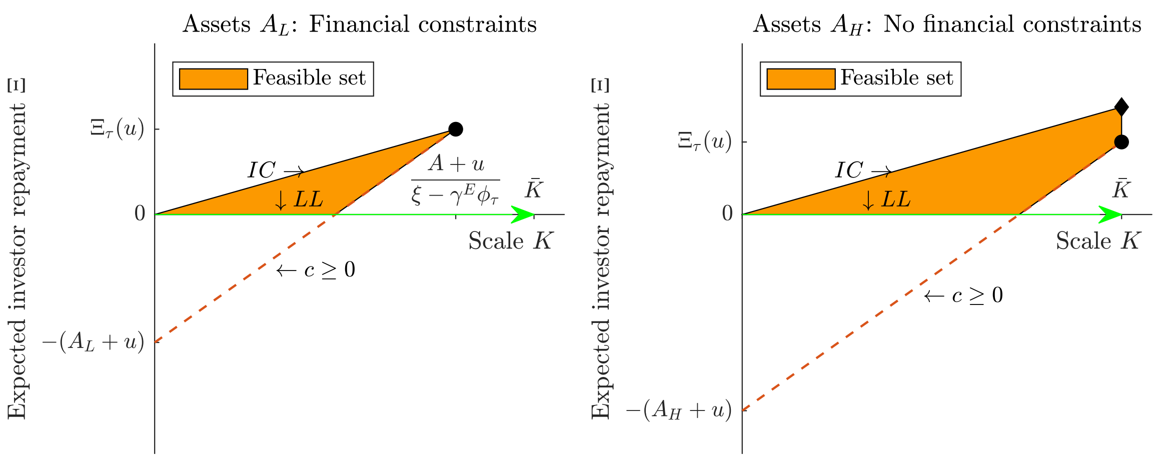

Figure 1. Feasible set of the inner problem: The set of feasible solutions is depicted in orange and forms a polygon. The objective function is increasing in the direction of the green arrow (up to $\bar{K}$). The left panel plots the case of low entrepreneur assets A_L, so that financial constraints bind. The right panel plots the case of high entrepreneur assets A_H, so that the efficient scale $\bar{K}$ is achievable.

(A.12), (IC), and (LL) forms a polygon (the orange region in Figure 1). Choosing the maximal scale $K^*_\tau(u)$ is optimal, since, for any given τ with $\hat{v}_\tau > 0$ and any fixed

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

$u \geq \eta \bar{U}^E + (1-\eta)\, \underline{U}^E$, the objective function $\hat{v}_\tau K - u$ is strictly increasing in K for $K \leq \bar{K}$ and independent of Ξ. The solution (indicated by the black dot) depends on whether financial constraints are binding (left panel) or not (right panel).

In both panels, the upper bound of Ξ defined by (IC) is an increasing affine function of K that runs through the origin, whereas the lower bound defined by Equation (A.12) is an increasing affine function of K with negative intercept $-(A+u)$. These bounds intersect at a positive value of K, since the slope coefficient in Equation (A.12), $pR - \gamma^E \phi_\tau$, is strictly greater than the slope of Equation (IC), $pR - \xi$:

$$\left(pR - \gamma^E \phi_\tau\right) - (pR - \xi) = \xi - \gamma^E \phi_\tau > \pi_\tau - \gamma^E \phi_\tau \geq \hat{v}_\tau > 0,$$

where the first inequality follows from Assumption 1 (i.e., $\xi > \pi_\tau$).

Financial constraints bind (left panel): In the left panel, entrepreneurial assets are sufficiently low, $A = A_L$, so that the upper bound (IC) and the lower bound (A.12) intersect at scale $\frac{A+u}{\xi - \gamma^E \phi_\tau} < \bar{K}$, which implies that $\bar{K}$ is outside of the feasible region. Financial constraints bind. Given the optimal scale $K^*_\tau(u) = \frac{A+u}{\xi - \gamma^E \phi_\tau}$, the expected repayment (A.14) is uniquely determined by the binding IC constraint (i.e., $\Xi_\tau(u) = (pR - \xi)\frac{A+u}{\xi - \gamma^E \phi_\tau}$), as indicated by the black circle in Figure 1.

Financial constraints do not bind (right panel): In the right panel, assets are sufficiently high, $A = A_H$, so that the intercept of constraint that defines the lower bound of Ξ (i.e., constraint (A.12), which ensures $c \geq 0$), shifts down by enough so that the efficient scale, $K^*_\tau(u) = \bar{K}$ can be achieved. In this case, there is a continuum of solutions for Ξ to support scale $\bar{K}$, indicated graphically by the line segment connecting the black diamond and the black circle. These solutions yield the same payoff to the SR fund, $\hat{v}_\tau \bar{K} - u$, and only differ in terms of the entrepreneur's upfront consumption c and the associated income pledged to investors. By convention, we focus on the solution with the lowest upfront payment to the entrepreneur and, accordingly, the minimum expected repayment to investors (A.14), indicated by the black circle. ■

Given a 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 $u \geq \eta \bar{U}^E + (1-\eta)\, \underline{U}^E$.

Lemma A.4 *In any solution to Problem 1**, the entrepreneur obtains her reservation utility from the bargaining game $u = \eta \bar{U}^E + (1-\eta)\, \underline{U}^E$.*

Proof: It suffices to show that the objective in (A.9) is strictly decreasing in u. (As long as $K^*_\tau(u) = \bar{K}$, the objective $\hat{v}_\tau \bar{K} - u$ trivially decreasing in u). Now consider the case where $K^*_\tau(u) = \frac{A+u}{\xi - \gamma^E \phi_\tau}$. Then, using $\hat{v}_\tau = \pi_\tau - \left(\gamma^E + \gamma^{SR}\right)\phi_\tau$, we obtain that:

$$\hat{v}_\tau K^*_\tau(u) - u = \frac{\hat{v}_\tau}{\xi - \gamma^E \phi_\tau} A - \frac{\xi + \gamma^{SR}\phi_\tau - \pi_\tau}{\xi - \gamma^E \phi_\tau} u \tag{A.15}$$

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.15) is strictly decreasing in u. ■

Given that the entrepreneur's utility is given by $u = \eta \bar{U}^E + (1-\eta)\, \underline{U}^E$, we can now define the (relevant) scale as a function of the bargaining power η, i.e.,

$$\hat{K}_\tau(\eta) := K_\tau^*\left[\eta\bar{U}^E + (1-\eta)\underline{U}^E\right] \tag{A.16}$$

The payoff to the SR fund *for a given* τ (at the optimal scale) is then given by:

$$U_B^{SR} = \hat{v}_\tau\hat{K}_\tau(\eta) - \left[\eta\bar{U}^E + (1-\eta)\underline{U}^E\right]. \tag{A.17}$$

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

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

$$\hat{\tau} = \arg\max_\tau \hat{v}_\tau\hat{K}_\tau(\eta). \tag{A.18}$$

Proof: In the relevant case $\hat{v}_D > 0$, we need to compare payoffs (A.17) under the two technologies. The clean technology is chosen if and only if $\hat{v}_C\hat{K}_C(\eta) > \hat{v}_D\hat{K}_D(\eta)$, which simplifies to (A.18). If $\hat{v}_D \leq 0$, then A.18 trivially holds as only $\hat{v}_C > 0$. ■

Lemmas A.3 to A.5 jointly characterize the solution to Problem 1**, which solves the original Problem 1 and allows us to determine the respective maximum feasible utilities:

$$\bar{U}^E = \underline{U}^E + \hat{v}_{\hat{\tau}}\hat{K}_\tau\left(\bar{U}^E\right) - \hat{v}_{\tau_F}K_{\tau_F}^F, \tag{A.19}$$

$$\bar{U}_I^{SR} = \underline{U}^{SR} + \hat{v}_{\hat{\tau}}\hat{K}_\tau\left(\underline{U}^E\right) - \hat{v}_{\tau_F}K_{\tau_F}^F \tag{A.20}$$

■

Proof of Corollary 2: Since the SR fund has all the bargaining power, we set $u = \underline{U}^E$. Then (A.14) implies that the expected repayment to investors satisfies $p\hat{X} = \Xi_{\hat{\tau}}\left(\underline{U}^E\right)$. Because any financing 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 observation that financial investors break even (Lemma A.1), which implies that $pX^F - I^F = 0$ and $X^F \in [0, R]$. The entrepreneur's upfront consumption follows from setting U^E to $\underline{U}^E$ and solving for c. ■

Proof of Corollary 3: The result follows from the cash-flow rights described in Corollary 2 and the fact that equity and debt are identical in our setup (given that the cash flow of the firm's project is zero in the low state). ■

Proof of Corollary 4: The statements follow directly from the impact-mandate condition in Proposition 2 and the observation that the difference in joint surplus, $\hat{v}_D - \hat{v}_C$, is strictly decreasing in $\gamma^{SR} + \gamma^E$, with $\hat{v}_D - \hat{v}_C < 0$ for $\gamma^{SR} + \gamma^E = 1$. ■

Proof of Proposition 3: The proof of this proposition follows from Lemmas A.6 and A.7 below. ■

Lemma A.6 *The firm is financially constrained under the clean technology both in the benchmark equilibrium with financial investors only and in the equilibrium with the SR* *fund only,* $\max\left\{K_C^F, K_C^{SR}\right\} < \bar{K}$*, if and only if*

$$\frac{A}{\bar{K}} < \min\left\{\xi - \pi_C, k_D \frac{\xi - \gamma^E \phi_C}{pR - \gamma^E \phi_D}\right\}. \tag{A.21}$$

Proof: We first prove that $K_C^F < \bar{K}$ if and only if $\frac{A}{\bar{K}} < \xi - \pi_C$. This follows directly from the definition of $K_C^F = \min\left\{\frac{A}{\xi - \pi_C}, \bar{K}\right\}$ given in Equation (4). Second, to see that $K_C^{SR} < \bar{K}$ if $\frac{A}{\bar{K}} < \min\left\{\xi - \pi_C, k_D \frac{\xi - \gamma^E \phi_C}{pR - \gamma^E \phi_D}\right\}$, note that analogous to Equation (10), K_C^{SR} can be expressed as

$$K_C^{SR} = \min\left\{\frac{A + \underline{U}_{SF}^E}{\xi - \gamma^E \phi_C}, \bar{K}\right\}. \tag{A.22}$$

In contrast to (10), $\underline{U}_{SF}^E$ now refers to the entrepreneur's outside option *under self-financing*, which yields scales $\frac{A}{k_D}$ and $\frac{A}{k_C}$ for the dirty and clean technology, respectively:

$$\underline{U}_{SF}^E := \max\left\{\frac{A}{k_D}(\pi_D - \gamma^E \phi_D), \frac{A}{k_C}(\pi_C - \gamma^E \phi_C)\right\}. \tag{A.23}$$

Equations (A.22) and (A.23) imply that $K_C^{SR} < \bar{K}$ if and only if

$$\frac{A}{\bar{K}} < \min\left\{k_D \frac{\xi - \gamma^E \phi_C}{pR - \gamma^E \phi_D}, k_C \frac{\xi - \gamma^E \phi_C}{pR - \gamma^E \phi_C}\right\}. \tag{A.24}$$

Therefore, if $\frac{A}{\bar{K}} < \min\left\{\xi - \pi_C, k_C \frac{\xi - \gamma^E \phi_C}{pR - \gamma^E \phi_C}, k_D \frac{\xi - \gamma^E \phi_C}{pR - \gamma^E \phi_D}\right\}$, we obtain that both $K_C^{SR} < \bar{K}$ and $K_C^F < \bar{K}$. Since $k_C \frac{\xi - \gamma^E \phi_C}{pR - \gamma^E \phi_C} > \xi - \pi_C$, this expression simplifies to (A.21).[31] This proves that $\max\left\{K_C^F, K_C^{SR}\right\} < \bar{K}$ if (A.21) holds. If (A.21) is not satisfied, $\frac{A}{\bar{K}} > \min\left\{\xi - \pi_C, k_D \frac{\xi - \gamma^E \phi_C}{pR - \gamma^E \phi_D}\right\}$, the above arguments imply that we obtain $K_C^F = \bar{K}$ or $K_C^{SR} = \bar{K}$ (or both). ■

Lemma A.7 *There is a strict complementarity,* $\hat{K}_C > \max\left\{K_C^F, K_C^{SR}\right\}$ *if and only if (A.21) holds. Else, there is no complementarity,* $\hat{K}_C = \max\left\{K_C^F, K_C^{SR}\right\} = \bar{K}$*.*

Proof: The proof consists of two parts. We first prove that $\hat{K}_C = \min\left\{\frac{A + \underline{U}^E}{\xi - \gamma^E \phi_C}, \bar{K}\right\} > K_C^{SR} = \min\left\{\frac{A + \underline{U}_{SF}^E}{\xi - \gamma^E \phi_C}, \bar{K}\right\}$ if and only if $K_C^{SR} < \bar{K}$ (see the condition in Lemma A.6). This follows directly from the fact that the outside option in the presence of financing from competitive financial investors exceeds the outside option under self-financing, i.e., $\underline{U}^E > \underline{U}_{SF}^E$.

Second, we show that $\hat{K}_C = \min\left\{\frac{A + \underline{U}^E}{\xi - \gamma^E \phi_C}, \bar{K}\right\} > K_C^F := \min\left\{\frac{A}{\xi - \pi_\tau}, \bar{K}\right\}$ if and only if $K_C^F < \bar{K}$. If $\hat{K}_C = \bar{K}$, the results follows immediately from $K_C^F < \bar{K}$. It remains to be

[31] Notice that $k_C \frac{\xi - \gamma^E \phi_C}{pR - \gamma^E \phi_C} - (\xi - \pi_C) = \left(\pi_C - \gamma^E \phi_C\right) \frac{pR - \xi}{pR - \gamma^E \phi_C} > 0$.

shown that $\frac{A+\underline{U}^E}{\xi-\gamma^E\phi_C} > K_C^F$. We obtain

$$\frac{A+\underline{U}^E}{\xi-\gamma^E\phi_C} - K_C^F = \frac{A+(\pi_D-\gamma^E\phi_D)K_D^F - \left(\xi-\gamma^E\phi_C\right)K_C^F}{\xi-\gamma^E\phi_C} \tag{A.25}$$

$$\geq \frac{(\pi_D-\gamma^E\phi_D)K_D^F - \left(\pi_C-\gamma^E\phi_C\right)K_C^F}{\xi-\gamma^E\phi_C} > 0, \tag{A.26}$$

where the first equality uses the definition $\underline{U}^E = (\pi_D-\gamma^E\phi_D)K_D^F$. The weak inequality follows from $A \geq K_C^F(\xi-\pi_C)$, see (4). The final, strict inequality follows from the fact that the dirty technology was optimally chosen by the entrepreneur in the benchmark equilibrium with financial investors only, $(\pi_D-\gamma^E\phi_D)K_D^F > \left(\pi_C-\gamma^E\phi_C\right)K_C^F$, see (3). Taken together, $\hat{K}_C > \max\left\{K_C^F, K_C^{SR}\right\}$ if and only if both $K_C^F < \bar{K}$ and $K_C^{SR} < \bar{K}$. This is satisfied if and only if Condition (A.21) holds (by Lemma A.6). ■

Proof of Corollary 5: Given that financial investors break even in expectation, see Lemma A.2, we can focus, without loss of generality, on the financing arrangement in which all external cash flow rights, $p\hat{X}$, are pledged to the SR fund.
Case 1: The proof first considers the case $K_C^F < \bar{K}$. In this case, Lemma A.7 implies that the equilibrium scale offered by the SR fund is strictly greater than that offered by competitive financial investors, i.e., $\hat{K}_C > K_C^F$. Since K_C^F is the largest possible clean scale that allows any investor to break even on financial terms, it must be the case that the SR fund makes a loss.
Case 2: We now consider the case $K_C^F = \bar{K}$. The financial resource constraint implies that

$$\hat{I} = \bar{K}k_C + \hat{c} - A = \underline{U}^E - \left(\pi_C-\gamma^E\phi_C\right)\bar{K} + p\hat{X}, \tag{A.27}$$

where the second equality uses the definition of $\hat{c}$ in (12). The net financial payoff is then given by

$$p\hat{X} - \hat{I} = \left(\pi_C-\gamma^E\phi_C\right)\bar{K} - (\pi_D-\gamma^E\phi_D)\bar{K} < 0, \tag{A.28}$$

where the inequality follows from the fact that the entrepreneur prefers the dirty technology under the respective benchmark agreements offered by financial investors (with respective scale $K_D^F = \bar{K}$ and $K_C^F = \bar{K}$). ■

Proof of Proposition 4: Ranking investments based on the social profitability index is optimal under the same conditions as for the standard profitability index ranking (see, e.g. Berk and DeMarzo, 2020). First, there must be a single resource constraint, which is satisfied given that the SR fund faces a single capital constraint κ in our setting. Second, the resource must be completely exhausted, which is satisfied because firms are of infinitesimal size in our setting. ■

Proof of Proposition 5: The social profitability index is defined as

$$\text{SPI:} = \frac{\Delta U^{SR}}{I^{SR}}. \tag{A.29}$$

The minimum investment that is sufficient to induce a change in production technology is given by pledging all cash flow rights to financial investors. Using the same steps as in the derivation of (A.28), we obtain that this minimum investment is given by

$$I_{\min}^{SR} = \left(\pi_D - \gamma^E \phi_D\right) K_D^F - \left(\pi_C - \gamma^E \phi_C\right) \hat{K}_C. \tag{A.30}$$

Given the definition of ΔU^{SR}, see Equation (14), the corresponding (maximum) SPI is given by

$$\begin{aligned}\text{SPI}_{\max} &= \frac{\hat{v}_C \hat{K}_C - \hat{v}_D K_D^F}{\left(\pi_D - \gamma^E \phi_D\right) K_D^F - \left(\pi_C - \gamma^E \phi_C\right) \hat{K}_C} \\ &= \gamma^{SR} \frac{\Delta\phi + \phi_C \left(1 - \frac{\hat{K}_C}{K_D^F}\right)}{\Delta\pi - \gamma^E \Delta\phi + \left(\pi_C - \gamma^E \phi_C\right)\left(1 - \frac{\hat{K}_C}{K_D^F}\right)} - 1.\end{aligned}$$

The ratio $\frac{\hat{K}_C}{K_D^F}$ depends on entrepreneurial assets A. It is easily verified that in all cases (constrained and unconstrained) $\text{SPI}_{\max}$ is increasing in γ^E and $\Delta\phi$ and decreasing in $\Delta\pi$ given that $\xi - \pi_\tau > 0$ (see Assumption 1).
Case 1: If assets A are sufficiently high, so that $\hat{K}_C = K_D^F = \bar{K}$, we obtain:

$$\text{SPI}_{\max} = \frac{\gamma^{SR}}{\frac{\Delta\pi}{\Delta\phi} - \gamma^E} - 1. \tag{A.31}$$

Case 2: If assets A are intermediate, so that $K_D^F = \bar{K}$ and $\hat{K}_C = \frac{A + \bar{K}(\pi_D - \gamma^E \phi_D)}{\xi - \gamma^E \phi_C}$, we obtain:

$$\text{SPI}_{\max} = \frac{\gamma^{SR}\left[\Delta\phi\xi + \phi_C\left(\xi - \pi_C - \Delta\pi - \frac{A}{\bar{K}}\right)\right]}{\Delta\pi\xi + \pi_C\left(\xi - \pi_C - \Delta\pi - \frac{A}{\bar{K}}\right) - \gamma^E\left[\phi_C\left(\xi - \pi_C - \frac{A}{\bar{K}}\right) + \Delta\phi\left(\xi - \pi_C\right)\right]} - 1. \tag{A.32}$$

To see that $\text{SPI}_{\max}$ is increasing in γ^E note that $\xi - \pi_C - \frac{A}{\bar{K}} > 0$ since $\bar{K} > K_C^F = \frac{A}{\xi - \pi_C}$. As a result, the denominator is strictly decreasing in γ^E.
Case 3: If assets A are sufficiently low, so that $\hat{K}_C \leq K_D^F < \bar{K}$, then

$$\text{SPI}_{\max} = \frac{\gamma^{SR}}{\frac{\Delta\pi}{\Delta\phi} - \frac{\gamma^E}{\xi}\left[\xi - \pi_C + \frac{\Delta\pi}{\Delta\phi}\phi_C\right]} - 1. \tag{A.33}$$

■

Proof of Lemma 2: Since financial capital is abundant relative to the financing needs of firms, an increase in κ only operates through the set of reformed firms, i.e.,

$$\Delta\Omega = \int_{j:\gamma_j^E < \bar{\gamma}_j^E \ \&\ SPI_j \geq SPI^*(\kappa)} \left(v_{C,j}\hat{K}_{C,j} - v_{D,j}K_{D,j}^F\right) d\mu(j). \tag{A.34}$$

An increase in κ only affects the threshold $SPI^*(\kappa)$. Since $v_{C,j} > 0 > v_{D,j}$, each term in the integral is positive, leading to a strictly positive effect as long as additional capital leads to reform. ■

Proof of Proposition 6: The proof consists of two parts. We first consider the case in which financial constraints are absent, $K_{C,j}^{SR} = K_{C,j}^{F} = \bar{K}_j$. In this case, the SR fund will ensure that all firms in the economy choose the clean technology (since $\gamma^{SR} + \gamma^{E} = 1$ implies that $\hat{v}_{C,j} > 0 > \hat{v}_{D,j}$) and operate at the socially optimal scale $\bar{K}_j$. Therefore, first-best welfare is achieved for $x^{SR} = 1$ (see Equation (6)). Moreover, as long as some firms would choose the dirty technology if only financial investors were present (i.e., $\gamma_j^E < \bar{\gamma}_j^E$), giving all capital to financial investors, $x^{SR} = 0$, would yield strictly lower welfare. This proves the first statement.

To prove that it may be strictly optimal to have $x^{SR} < 1$ consider the following case. Suppose that all firm types are financially constrained (i.e., $\max\left\{K_{C,j}^{SR}, K_{C,j}^{F}, K_{D,j}^{F}\right\} < \bar{K}_j$) and that total investor capital is large enough such that the following two conditions are jointly met for some $\tilde{x}^{SR} \in (0,1)$:

1. Financial investors (with a fraction $1 - \tilde{x}^{SR}$ of total capital) could finance dirty production by all firms at scale $K_{D,j}^{F}$.
2. The SR fund (with a fraction $\tilde{x}^{SR}$ of total capital) could finance all firms at a clean scale of $\frac{A_j + \bar{U}_j^E}{\xi_j - \gamma_j^E \phi_{C,j}}$.

The first condition ensures that all firms have the outside option of dirty production at scale $K_{D,j}^{F}$ by raising financing from financial investors. The second condition ensures that, given this threat, the SR fund has sufficient capital to induce all firms to adopt the clean production technology by offering a (larger) clean scale of $\frac{A_j + \bar{U}_j^E}{\xi_j - \gamma_j^E \phi_{C,j}} > K_{C,j}^{SR}$ (see Proposition 3). This scale increase is socially valuable, implying that welfare is strictly higher for $\tilde{x}^{SR} < 1$ than for $x^{SR} = 1$. ■

Proof of Result 1: The investor's objective (24) is affine in κ_i^{SR} with coefficient $\gamma^i \frac{\phi_D K_D^F - \phi_C \hat{K}_C}{\pi_D K_D^F - \pi_C \hat{K}_C} - 1$. If $\gamma^i = 1/n$, then the coefficient turns negative for n sufficiently high. Therefore, the optimal contribution is $\kappa_i^{SR} = 0$.

■

Proof of Corollary 6: The result follows immediately from the Proof of Result 1. In particular, it is required that $\gamma^i \frac{\phi_D K_D^F - \phi_C \hat{K}_C}{\pi_D K_D^F - \pi_C \hat{K}_C} - 1 > 0$ for at least one agent. ■

Proof of Corollary 7: See Proof of Corollary 6 and replace γ^i with $\sum_{i=1}^{n_1} \gamma^i$. ■

Proof of Corollary 8: When there is an additional warm-glow benefit, the investor's objective is affine in κ_i^{SR} with associated coefficient $w^i + \gamma^i \frac{\phi_D K_D^F - \phi_C \hat{K}_C}{\pi_D K_D^F - \pi_C \hat{K}_C} - 1$, which is positive if w^i is sufficiently high. ■

References

Aghion, Philippe, Roland Bénabou, Ralf Martin, and Alexandra Roulet, "Environmental Preferences and Technological Choices: Is Market Competition Clean or Dirty?," 2019. Working Paper.

Albuquerque, Rui, Yrjö Koskinen, and Chendi Zhang, "Corporate Social Responsibility and Firm Risk: Theory and Empirical Evidence," *Management Science*, 2019, *65* (10), 4451–4469.

Andreoni, James, "Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving," *The Economic Journal*, 06 1990, *100* (401), 464–477.

Angelis, Tiziano De, Peter Tankov, and Olivier David Zerbib, "Climate Impact Investing," *Management Science*, forthcoming.

Baker, Malcolm, Daniel Bergstresser, George Serafeim, and Jeffrey Wurgler, "The Pricing and Ownership of US Green Bonds," *Annual Review of Financial Economics*, 2022, *14*, 415–437.

Barber, Brad M, Adair Morse, and Ayako Yasuda, "Impact investing," *Journal of Financial Economics*, 2021, *139* (1), 162–185.

Bergstrom, Theodore, Lawrence Blume, and Hal Varian, "On the private provision of public goods," *Journal of Public Economics*, 1986, *29* (1), 25–49.

Berk, Jonathan and Jules H van Binsbergen, "The Impact of Impact Investing," 2021. Working Paper, Stanford University and Wharton.

_ **and Peter DeMarzo**, *Corporate Finance*, 5th ed., Pearson, 2020.

Biais, Bruno and Augustin Landier, "Emission caps and investment in green technologies," 2022. Working Paper, HEC Paris.

Bonnefon, Jean-Francois, Augustin Landier, Parinitha Sastry, and David Thesmar, "The Moral Preferences of Investors: Experimental Evidence," 2019. Working Paper, TSE, MIT and HEC.

Broccardo, Eleonora, Oliver Hart, and Luigi Zingales, "Exit versus Voice," *Journal of Political Economy*, 2022, *130* (12), 3101–3145.

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.

Coase, Ronald H., "The Problem of Social Cost," *The Journal of Law and Economics*, 1960, *3*, 1–44.

Dangl, Thomas, Michael Halling, Jin Yu, and Josef Zechner, "Social Preferences and Corporate Investment," 2023. Working Paper, WU Vienna.

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

Dewatripont, Mathias and Jean Tirole, "The Morality of Markets and the Nature of Competition," 2020. Working Paper, University of Brussels and Toulouse School of Economics.

Edmans, Alex, Doron Levit, and Jan Schneemeier, "Socially Responsible Divestment," 2022. European Corporate Governance Institute–Finance Working Paper.

Financial Times, "Bill Gates-Backed Fund Aims to Invest $15bn in Clean Tech," https://www.ft.com/content/f25fd95d-e2d8-43f8-b786-1552a1f0059e 2022. Accessed March 28, 2022.

Friedman, Milton, "A Friedman Doctrine: The Social Responsibility of Business is to Increase its Profits," 1970. The New York Times Magazine, September 13, 1970.

Gollier, Christian and Sébastien Pouget, "The "Washing Machine": Investment Strategies and Corporate Behavior with Socially Responsible Investors," 2014. Working Paper, Toulouse School of Economics.

Green, Daniel and Benjamin Roth, "The Allocation of Socially Responsible Capital," 2021. Working Paper, Harvard Business School.

Gupta, Deeksha, Alexandr Kopytov, and Jan Starmans, "The Pace of Change: Socially Responsible Investing in Private Markets," *Available at SSRN 3896511*, 2022.

Harstad, Bård, "Buy Coal! A Case for Supply-Side Environmental Policy," *Journal of Political Economy*, 2012, *120* (1), 77–115.

Hart, Oliver and John Moore, "Default and Renegotiation: A Dynamic Model of Debt," *Quarterly Journal of Economics*, 1998, *113* (1), 1–41.

Hartzmark, Samuel M and Abigail B Sussman, "Do Investors Value Sustainability? A Natural Experiment Examining Ranking and Fund Flows," *Journal of Finance*, 2019, *74* (6), 2789–2837.

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.

Hoffmann, Florian, Roman Inderst, and Ulf Moslener, "Taxing Externalities Under Financing Constraints," *The Economic Journal*, 2017, *127* (606), 2478–2503.

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.

Inderst, Roman and Florian Heider, "A Corporate Finance Perspective on Environmental Policy," 2022. Working Paper, Goethe University Frankfurt.

_ **and Marcus M Opp**, "Socially Optimal Eligibility Criteria for ESG Funds," 2022. Working Paper, Stockholm School of Economics.

Landier, Augustin and Stefano Lovo, "ESG Investing: How to Optimize Impact," 2020. Working Paper, HEC.

Moisson, Paul-Henri, "Ethics and Impact Investment," 2020. Working Paper, TSE.

Morgan, John and Justin Tumlinson, "Corporate Provision of Public Goods," *Management Science*, 2019, *65* (10), 4489–4504.

Nguyen, Nam, Alejandro Rivera, and Harold H Zhang, "Incentivizing Investors for a Greener Economy," 2021. Working Paper, UT Dallas.

Pastor, Lubos, Robert F. Stambaugh, and Lucian Taylor, "Sustainable Investing in Equilibrium," *Journal of Financial Economics*, 2021, *142* (2), 550–571.

Pedersen, Lasse Heje, Shaun Fitzgibbons, and Lukasz Pomorski, "Responsible Investing: The ESG-Efficient Frontier," *Journal of Financial Economics*, 2021, *142* (2), 572–597.

Riedl, Arno and Paul Smeets, "Why Do Investors Hold Socially Responsible Mutual Funds?," *Journal of Finance*, 2017, *72* (6), 2505–2550.

Roth, Benjamin N., "Impact Investing: A Theory of Financing Social Enterprises," 2019. Working Paper, Harvard Business School.

Samuelson, Paul A., "The Pure Theory of Public Expenditure," *Review of Economics and Statistics*, 1954, *36* (4), 387–389.

Tirole, Jean, *The Theory of Corporate Finance*, Princeton University Press, 2006.

_ , "Some Political Economy of Global Warming," *Economics of Energy & Environmental Policy*, 2012, *1* (1), 121–132.

Welch, Ivo, "Why Divestment Fails," *New York Times*, May 10 2014, p. A23.

Zerbib, Olivier David, "The Effect of Pro-Environmental Preferences on Bond Prices: Evidence from Green Bonds," *Journal of Banking and Finance*, 2019, *98*, 39 – 60.

_ , "A Sustainable Capital Asset Pricing Model (S-CAPM): Evidence from green investing and sin stock exclusion," *Review of Finance*, 2022, *26* (6), 1345 – 1388.

Online Appendix

B Production technology specification

B.1 Many Production Technologies and Social Goods

In this section, we describe how Proposition 2 generalizes to more than two technologies and social goods. Suppose that the entrepreneur has access to $N \geq 2$ production technologies characterized by technology-specific cash flow, cost, and moral hazard parameters R_τ, $\bar{K}_\tau$, k_τ, p_τ, Δp_τ, and B_τ. The differences in parameters could reflect features such as increased willingness to pay for goods produced by firms with clean production technologies, implying $R_C > R_D$ (for models with this feature, see Aghion, Bénabou, Martin and Roulet, 2019; Albuquerque, Koskinen and Zhang, 2019). Moreover, we allow for the technology-specific social cost parameter ϕ_τ to be negative, in which case the technology generates a positive externality (a social good).

In analogy to the baseline model, we can then define, for each technology $\tau \in \{1, ..., N\}$, the financial value π_τ, the agency rent ξ_τ, and the maximum scale available from financial investors K_τ^F, maintaining the assumption that $\xi_\tau > \pi_\tau$ for all τ. A straightforward extension of Lemma 1 then implies that, in the absence of investment by the SR fund, the entrepreneur chooses technology

$$\tau_F = \arg\max_\tau \left(\pi_\tau - \gamma^E \phi_\tau\right) \min\left\{\frac{A}{\xi_\tau - \pi_\tau}, \bar{K}_\tau\right\}. \tag{B.1}$$

Equation (B.1) clarifies the entrepreneur's relevant outside option with N technologies: Any production technology dirtier than τ_F is not a credible threat. Given the credible threat τ_F, the induced technology choice in the presence of the SR fund $\hat{\tau}$ and the associated capital stock $\hat{K}$ are given by

$$\hat{\tau} = \arg\max_\tau \hat{v}_\tau \min\left\{\frac{A + \underline{U}^E}{\xi_\tau - \gamma^E \phi_\tau}, \bar{K}_\tau\right\}, \tag{B.2}$$

$$\hat{K} = \begin{cases} \min\left\{\frac{A+\underline{U}^E}{\xi_\tau - \gamma^E \phi_\tau}, \bar{K}_\tau\right\} & \text{if } \hat{v}_{\hat{\tau}} > 0 \\ 0 & \text{if } \hat{v}_{\hat{\tau}} \leq 0 \end{cases}, \tag{B.3}$$

which mirrors Proposition 2.

Whereas the formal expressions are unaffected by whether the externality is negative or positive, there is one important difference. If externalities are negative, an explicit impact mandate is necessary to ensure that the SR fund can affect the firm's choice of production technology. An impact mandate reduces the outside option for the SR fund (see Equation (7)), thereby unlocking the required additional financing capacity. In contrast, if the externalities under technology D are positive, $\phi_D < 0$, the outside option for the SR fund is higher under an impact mandate than under a narrow mandate (the outside option is positive under an impact mandate, whereas it is zero under a narrow mandate). Therefore, in the presence of positive externalities, impact is possible and, in fact, more likely to occur under a narrow mandate, revealing an interesting asymmetry between preventing social costs and encouraging social goods.

The more general technology specification additionally provides some insights about cases that we previously excluded. First, the entrepreneur's relevant outside option with N technologies is the technology that maximizes bilateral surplus for financial investors and the entrepreneur. Any technology that does not maximize this bilateral surplus is not a credible threat. Note that for some industries the cleanest technology may also be profit-maximizing (e.g., because of demand from SR consumers). In this case, there is no trade-off between doing good and doing well and, hence, socially responsible investors play no role. Second, it is also possible that, for some industries, any feasible technology τ yields negative social surplus (i.e., $v_\tau < 0$ for all τ). In this case, the socially optimal scale is zero and the entrepreneur is optimally rewarded with a transfer to shut down production.

B.2 Decreasing returns to scale

We now consider the case in which the two production technologies $\tau \in \{C, D\}$ exhibit standard decreasing returns to scale. In particular, suppose that the *marginal* financial value $\pi_\tau(K)$ is strictly decreasing in K. Then the first-best scale K_C^{FB} under the (socially efficient) clean technology is characterized by the first-order condition

$$\pi_C\left(K_C^{FB}\right) = \phi_C. \tag{B.4}$$

Note that the first-best scale K_C^{FB} corresponds to $\bar{K}$ in our baseline model.

Now consider the scenario in which technology D is chosen in the absence of the SR fund, with an associated scale of K_D^F. Moreover, for ease of exposition, focus on the case $\gamma^E + \gamma^{SR} = 1$, so that the SR fund has incentives to implement the first-best scale. The optimal financing agreement that the SR fund offers to induce the entrepreneur to switch to the clean technology then comprises three cases.

1. If the financing constraints generated by the agency problem are severe, i.e., assets are below some cutoff $A < \tilde{A}$, the optimal agreement offered by the SR fund rewards the entrepreneur exclusively through an increase in scale (rather than upfront consumption). The resulting clean scale, $\hat{K}_C$, is smaller than first-best scale (i.e., $\hat{K}_C < K_C^{FB}$). In our baseline model, this case corresponds to $\hat{K}_C = \frac{A+U^E}{\xi-\gamma^E\phi_\tau} < \bar{K}$

2. If the financing constraints generated by the agency problem are intermediate, i.e., $\tilde{A} < A < A^{FB}$, the optimal agreement specifies the first-best scale, $\hat{K}_C = K_C^{FB}$. In this case, it is efficient to increase clean scale up to the first-best level but no further, since scale above and beyond K_C^{FB} would reduce joint surplus. Inducing the entrepreneur to switch technologies solely through an increase in scale would require a production scale exceeding the first-best level K_C^{FB}. It is therefore optimal to partially compensate the entrepreneur through a reduction in repayment (or an upfront consumption transfer, as in Corollary 2). In our baseline model, this refers to the case where $K_C^F < \bar{K}$ but $\hat{K}_C = \bar{K}$.

3. If financing constraints do not bind, $A > A^{FB}$, we essentially obtain a Coasian solution (e.g., a downstream fishery might pay an upstream factory to reduce pollution, as in Coase 1960).[32] In this case, we distinguish between two sub-cases.

   (a) If $\phi_C = 0$, financial investors would provide the first-best scale of the clean technology, i.e., $K_C^F = K_C^{FB}$. In our baseline model, this case corresponds $K_C^F = \hat{K}_C = \bar{K}$. The SR fund simply needs to provide a subsidy to induce a switch in the production technology, as in Corollary 2.

   (b) If $\phi_C > 0$, financial investors would provide funding above and beyond the first-best scale of the clean production technology, i.e., $K_C^F > K_C^{FB}$. In our baseline model, this case cannot occur. The optimal financing agreement with the SR fund then ensures that the clean production technology is run at the first-best scale, $\hat{K}_C = K_C^{FB} < K_C^F$ via a lower repayment and/or upfront consumption, as in Corollary 2.

This case-by-case analysis shows that the insights from the reduced-form CRS specification of the baseline model extend to a standard specification with decreasing returns to scale.

[32] Note that, in cases 2 and 3, the agreement needs to explicitly limit the amount of firm investment (and not simply specify the technology). Otherwise, the entrepreneur would find it privately optimal to convert upfront consumption into additional firm investment.