A Theory of Socially Responsible Investment *

Martin Oehmke† Marcus Opp‡

April 26, 2022

Abstract

We characterize conditions under which socially responsible investors impact firm behavior, in a setting in which firm production generates social costs and is subject to financing constraints. In this setting, impact requires a broad mandate: Socially responsible investors need to internalize social costs irrespective of whether they are investors in a given firm. If firms face binding financial constraints, impact is optimally achieved by enabling a scale increase for clean production, and socially responsible and financial investors are complementary: Jointly they can achieve higher surplus than either investor type alone. Scarcely socially responsible capital should be allocated according to a social profitability index (SPI). This micro-founded ESG metric captures not only a firm’s social status quo but also the counterfactual social costs produced in the absence of socially responsible investors.

Keywords: Socially responsible investing, ESG, Social Profitability Index (SPI), 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, Anastasia Kartasheva, Lukasz Pomorski, Ailsa Roël, Joel Shapiro, Lin Shen, 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 investment. Assets under management in socially responsible 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). From an asset management perspective, this trend raises immediate questions about the financial performance of such investments (Hong and Kacperczyk, 2009, Chava, 2014, Barber, Morse and Yasuda, 2021). However, if socially responsible investing is to generate real impact, it must affect firms’ production decisions. This raises additional, fundamental questions: Under which conditions can socially responsible investors impact firm behavior? How should scarce socially responsible capital be allocated across firms?

Answering these questions requires taking a corporate finance view of socially responsible investment. To this end, we incorporate socially responsible investors and the choice between clean and dirty production into an otherwise standard model of corporate financing with agency frictions, 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 socially responsible investors can indeed push firms to adopt clean production. However, because this requires them to fund investments that profit-motivated

\(^1\) 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.
investors would not fund, impact is only possible if socially responsible investors are willing to sacrifice financial returns. In our setting, a necessary condition for impact is that socially responsible investors follow a broad mandate, in the sense that they internalize social costs generated by firms regardless of whether they are investors in these firms. In the presence of binding financial constraints, 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 resulting increase in clean production raises total surplus, even compared to the scenario in which all capital is held by socially responsible investors, reflecting a complementarity between financial and socially responsible capital. When faced with an investment decision across many heterogeneous firms, scarce socially responsible capital should be allocated according to a social profitability index (SPI). One key feature of this micro-founded ESG metric is that avoided social costs are relevant for ranking investments. Hence, investments in “sin” industries are not necessarily inconsistent with a socially responsible investment mandate.

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 limit 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. A socially responsible (SR) fund also cares about financial returns but, in addition, internalizes some of the social costs generated by firms. We distinguish between two mandates for the SR fund. Under a narrow mandate, the SR fund cares about social
costs of firms it has invested in. Under a broad mandate, the SR fund cares about social costs independent of its own investments.

We first develop two benchmark cases. In the first benchmark, 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 constraint, the planner can implement the first best via a Pigouvian tax on the externality. In contrast, if the firm is financially constrained, a Pigouvian tax alone does not achieve first best and must be complemented with a subsidy. The broader point is that regulation that targets one source of inefficiency (externalities) but does not address the other (financing constraints) has limited effectiveness in our setting.

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 socially responsible investors can address this inefficiency. Our model demonstrates that the SR fund can have impact (i.e., change the firm’s technology choice) if it follows a broad mandate. In this case, 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, so that it only cares about social costs generated by firms it has invested in, the equilibrium allocation would be unchanged relative to the benchmark case in which only financial
investors are present. In this case, dirty firms remain dirty and obtain financing from financial investors with abundant capital.

The optimal financing agreement in the presence of the SR fund can be implemented by issuing two bonds, a green bond purchased by socially responsible investors 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 et al. (2018) 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.

If the firm is financially constrained under the clean technology, the optimal way for the SR fund to achieve impact is by facilitating an increase in the scale of clean production. In this case, a complementarity arises between socially responsible and financial 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 socially responsible investors and, thereby, generates additional financing capacity. Since binding financing constraints imply that clean production is below the socially optimally scale, this additional financing capacity enables a surplus-enhancing increase in the scale of (socially valuable) clean production.

While socially responsible 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 socially responsible capital is invested most efficiently. Which firms should impact investors target? A multi-firm extension of our model yields a micro-founded investment criterion for scarce socially responsible capital, the Social Profitability Index (SPI).
Similar to the (standard) profitability index, the SPI measures “bang for buck”— in this case, value created for socially responsible investors per unit of socially responsible capital consumed. However, 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 socially responsible investors. Therefore, investment metrics for socially responsible investors should include estimates of carbon emissions that can be avoided if the firm adopts a cleaner production technology. Because avoided externalities matter, it can be efficient for socially responsible investors to invest in firms that, in an absolute sense, generate a high level of social costs even under clean production. Accordingly, investments in sin industries (see Hong and Kacperczyk, 2009) can be consistent with a socially responsible investment mandate. In contrast, it is efficient to not invest in firms that are already committed to clean production (e.g., because, intrinsically, the entrepreneur cares sufficiently about the environment), because clean production will occur regardless of investment by socially responsible investors.

We conclude with a brief discussion of two extensions. First, we consider how a SR fund with a broad mandate can emerge from a setting in which small individual investors delegate their investment. This analysis highlights the importance of overcoming a free-rider problem. One way to achieve this is for sovereign funds to invest on behalf of citizens, rather than paying out their resource income to individual citizens. Second, we briefly consider social goods. While our analysis is primarily motivated by the mitigation of negative production externalities (such as carbon emissions) by a SR fund, an extension to positive externalities reveals an interesting asymmetry. Specifically, in the case of positive externalities, a narrow mandate (i.e., only accounting for the positive externalities generated by one’s own investment) generates larger impact than a broad mandate.
Related Literature. The theoretical literature on socially responsible investing consists of two main strands. Following the pioneering paper by Heinkel, Kraus and Zechner (2001), the first strand studies the effects of exclusion, such as investor boycotts, divestment, or portfolio underweighting of 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. The vast majority of this literature\(^2\) generates this effect via a reduction in risk sharing that raises the firm’s cost of capital. However, Heinkel et al. (2001) and Broccardo, Hart and Zingales (2020) point out that the effect on risk premia is small if profit-seeking investors can substitute for divested capital.\(^3\) In Landier and Lovo (2020), divestment can be effective because of a matching friction between firms and investors, which implies that a boycott by socially responsible investors (probabilistically) leaves the firm without access to financing.

Our model completely shuts off the exclusion channel by considering a risk-neutral setting with perfectly elastic supply of profit-motivated capital. Our setting therefore echoes Broccardo et al. (2020), who, building on Heinkel et al. (2001) and Hart and Zingales (2017), conclude that “voice” (engagement) is more effective than “exit” (divestment).

Consequently, our paper is more closely related to the second strand of the literature, which studies the effects of engagement by activist investors who intrinsically care about social costs, thereby providing a corporate finance perspective on the economics of motivated agents (see e.g., Besley and Ghatak, 2005; Bénabou and Tirole, 2006). Rather than imposing costs on dirty firms via the threat of divestment, socially responsible activists effectively subsidize firms to adopt clean technologies through the direct pricing of social preference. Chowdhry, Davies and Waters (2018) show that such 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)


\(^3\)As Davies and Van Wesep (2018) point out, divestment can have other unintended consequences, for example, by inducing firms to prioritize short-term profit at the expense of long-term value.
compares impact investing with grants, highlighting the ability of investors to withdraw capital as an advantage over grants.\footnote{A corollary of subsidizing clean production is that socially responsible investors sacrifice financial return. In contrast, in Gollier and Pouget (2014) a large activist investor can generate positive abnormal returns by reforming firms and selling them back to the market.}

Our contribution relative to this strand of literature is twofold. First, we highlight that socially responsible investors can achieve impact by relaxing financing constraints for clean production and that, in doing so, they generate a complementarity with financial investors. Second, our framework endogenizes the allocation of socially responsible capital across firms using a micro-founded investment criterion, the social profitability index.

2 Model Setup

We study the role of socially responsible investing in a setting in which \textit{production externalities} interact with \textit{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). The main innovation of our model is that the firm has access to two different production technologies, one of them “clean” (i.e., associated with low social costs) and the other “dirty” (i.e., associated with 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 $R \cdot \min(K, \bar{K})$ 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.\footnote{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.}

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 $k_C > k_D$.\footnote{The assumption that $0 \leq \phi_C < \phi_D$ reflects that our analysis focuses on the mitigation of negative production externalities by socially responsible investors. We discuss the case of positive production externalities in Section 5.2.} 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 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 $\tau$ 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.\footnote{Once we allow for $N$ technologies (see Online Appendix B.1), 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 socially responsible investment, since even purely profit-motivated capital would ensure clean production in this case.}

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 $\xi$ 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 $\tau$, the agency cost per unit of capital $\xi := p \frac{B}{\Delta p}$ satisfies

$$\pi_\tau < \xi < pR - \frac{p}{\Delta p} \pi_\tau.$$  \hspace{1cm} (1)

This assumption states that the moral hazard problem, as characterized by the agency cost per unit of capital $\xi$, 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, $\pi$ and $\nu$ 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 the need 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. Both investor types care about expected cash flows, but only the socially responsible fund’s objective is affected by the social costs of production, $\phi_r \bar{K}$, with intensity $\gamma_{SR}$. We normalize $\gamma_{SR} + \gamma^E \leq 1$, so that
jointly investors and the entrepreneur do not internalize more than 100% of social costs.\footnote{The assumption that at least some investors internalize social costs is consistent with evidence in Riedl and Smeets (2017), Bonnefon et al. (2019), and Hartzmark and Sussman (2019).}

We assume that financial investors literally do not care about social costs. However, an alternative setting in which financial investors do care about social costs but do not act on them because of a free-rider problem would yield equivalent results.

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

**Definition 1 (Broad versus Narrow Mandate)** A SR fund has a **broad mandate** if it accounts for externalities regardless of whether it has invested in the firm. A SR fund has a **narrow mandate** if it only accounts for externalities caused by firms it has invested in.

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 $K \leq \bar{K}$,\footnote{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 $K > \bar{K}$.} the total promised repayment $X$, the effort decision $a$, upfront consumption by the entrepreneur $c$, and the technology choice $\tau \in \{C, D\}$,

$$U^E(K, X, \tau, c, a) = p(RK - X) - (A - c) - \gamma^E \phi_\tau K + 1_{a=0} [BK - \Delta p(RK - X)]. \quad (U^E)$$

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, \( \gamma^E \phi_r K \), 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

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

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, \( K k_r \), must equal the total investments made by the entrepreneur and outside investors,

\[
K k_r = A - c + IF + ISR, \tag{2}
\]

where \( IF \) and \( ISR \) 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 \( ISR = XSR = 0 \). The entrepreneur’s objective
is then to choose a financing arrangement (consisting of scale $K \in [0, \bar{K}]$, 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

$$U^F := pX^F - I^F \geq 0$$

As a preliminary step, it is useful to analyze the financing arrangement that maximizes scale for a given technology $\tau$ 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 $\tau$ 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, $\xi := pR \Delta r$, and the per-unit financial value under technology $\tau$, $\pi_\tau$. 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 $\tau$ 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 \( \tau_F \) 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 \( (\gamma^E = 0) \), 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 \( \tau \) that maximizes her utility

\[
U^E_F = \max_\tau (\pi_F - \gamma^E \phi_F) K^F_F. \tag{3}
\]

where

\[
K^F_F := \min \left\{ \frac{A}{\xi - \pi_F}, \bar{K} \right\}. \tag{4}
\]

According to Lemma 1, if financing is raised from financial investors only, the entrepreneur chooses the technology \( \tau_F \) 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 \( K^F_F \). Maximum scale (up to \( \bar{K} \)) is optimal because, under the equilibrium technology \( \tau_F \), the project generates positive surplus for the entrepreneur and financial investors. It follows that the entrepreneur adopts the dirty technology whenever

\[
(\pi_D - \gamma^E \phi_D) K^F_D > (\pi_C - \gamma^E \phi_C) K^F_C. \tag{5}
\]

Given that the dirty technology is financially more profitable, \( \pi_D > \pi_C \), and the scale is larger under the dirty technology, \( K^F_D \geq K^F_C \), this condition is satisfied whenever the
entrepreneur’s concern for externalities $\gamma^E$ lies below a strictly positive cutoff $\tilde{\gamma}^E$.

**Corollary 1 (Benchmark: Conditions for Dirty Production)** If only financial investors are present, the entrepreneur adopts the dirty production technology if $\gamma^E < \tilde{\gamma}^E := \frac{\pi_D K_D^E - \pi_C K_C^E}{\phi_D K_D^E - \phi_C K_C^E}$. 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.\(^\text{10}\)

### 3.2 The Planner’s Problem

Before turning to socially responsible investment, we outline, as a second benchmark, what a welfare-maximizing planner would do. In our setting, welfare is defined as the total surplus created by production,

$$\Omega := \min\{K, \bar{K}\} \cdot v_\tau. \quad (6)$$

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, $A < \bar{K} (\xi - \pi_C)$, first-best welfare can be achieved by a Pigouvian tax of $\phi_\tau$ per unit of scale.

\(^{10}\)Because the entrepreneur is constrained under self-financing, $A < k_D \bar{K}$, she prefers the clean technology if and only if $\frac{A}{k_C} (\pi_C - \gamma^E \phi_C) \geq \frac{A}{k_D} (\pi_D - \gamma^E \phi_D)$. Hence, the entrepreneur is “corrupted” by financial markets when $\gamma^E \in (\tilde{\gamma}^E, \bar{\gamma}^E)$ where $\bar{\gamma}^E := \frac{k_C \pi_D - k_D \pi_C}{k_C \phi_D - k_D \phi_C}$.\}
2. Otherwise, a Pigouvian tax alone cannot achieve first best, but needs to be complemented with a subsidy of $\bar{K} (\xi - \pi_C) - A$.

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 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.\(^{11}\)

If instead financial constraints are binding for the clean technology, a Pigouvian tax of $\phi_r$ (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 subsidy could either be administered 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 such 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 scales.\(^{12}\)

\(^{11}\) 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).

\(^{12}\) 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 Socially Responsible Investment

We now turn to our main question: whether and how a SR fund can impact 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.

4.1 Single-Firm Analysis

In contrast to financial investors, the SR fund cares not only about financial payoffs $X^{SR}$ but also about social costs $\phi_{r}K$. The extent to which social costs are internalized depends both on whether the fund has a broad or narrow mandate, $M \in \{B, N\}$ (see Definition 1), and the associated internalization parameter $\gamma^{SR} \leq 1 - \gamma^{E}$. If $\gamma^{E} + \gamma^{SR} = 1$, agents in the model jointly internalize all externalities. If $\gamma^{E} + \gamma^{SR} < 1$, some externalities (e.g., those imposed on future generations) are not accounted for. The partial internalization of externalities could result from imperfect coordination among individual socially responsible investors who delegate their investments to a SR fund.\(^1\)

Depending on the SR fund’s mandate (broad or narrow), the fund’s objective is given by

\[ U^{SR}_{B} = pX^{SR} - I^{SR} - \gamma^{SR}\phi_{r}K, \]

\[ U^{SR}_{N} = pX^{SR} - I^{SR} - \gamma^{SR}\phi_{r}K \cdot 1_{I^{SR} > 0}. \]

Accordingly, under a broad mandate the fund internalizes social costs independent of whether the fund has invested in the company, whereas a narrow mandate only internalizes social costs if the fund has invested in the firm. It is useful to note that even under

\(^{13}\)See Section 5.1 for a brief discussion on how the SR fund’s mandate could be endogenized.
a broad 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.$^{14}$

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, $U^E$ given in Equation (3), now becomes the entrepreneur’s outside option. If the socially responsible fund remains passive, $I^{SR} = 0$, its payoff under a broad mandate satisfies

$$U^E_{SR} = -\gamma^{SR} \Phi_{\tau_F} K^F_{\tau_F} < 0.$$  \hspace{1cm} (7)

This expression, which acts as the SR fund’s reservation utility under a broad mandate, accounts for the social costs generated when the entrepreneur raises financing exclusively from financial investors and chooses technology $\tau_F$ and scale $K^F_{\tau_F}$ (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 $U^E_{SR} = 0$. This 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 $U^E_{SR}$ and $U^E$, the SR fund can engage with the entrepreneur and agree on a financing contract that specifies the technology $\tau$, scale $K$, as well as the required financial investments and cash

---

$^{14}$ 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. However, whereas the recent wave of inflows into ESG funds and the empirical evidence cited in footnote 8 all suggest that some investors explicitly care about externalities, there is no evidence that these investors also value rents to insiders (managers).
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

$$\max_{I^F, I^{SR}, X^{SR}, X^F, K, c, \tau} U^M_{SR}$$

subject to the entrepreneur’s IR constraint:

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

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

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, $U^E$. 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. Under a narrow mandate, the equilibrium technology choice and scale are identical to the benchmark equilibrium described in Lemma 1.

2. Let $\hat{v}_\tau := \pi_\tau - (\gamma^E + \gamma^{SR}) \phi_\tau \geq v_\tau := \pi_\tau - \phi_\tau$ denote bilateral surplus (per unit of scale) for the SR fund and the entrepreneur. The equilibrium technology choice
under a broad mandate is given by

\[ \hat{\tau} = \arg \max_{\tau} \hat{v}_\tau \hat{K}_\tau, \]

where the scale given technology \( \tau \) satisfies

\[ \hat{K}_\tau = \min \left\{ \frac{A + U^E}{\xi - \gamma^E \phi}, \bar{K} \right\}. \]

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 responsible fund, in what follows we focus on the broad fund mandate. Under a broad 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, \( \frac{A + U^E}{\xi - \gamma^E \phi} \), ensures that the entrepreneur earns the same utility as her outside option \( U^E \). (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)** There is a continuum of financing arrangements that achieve equivalent impact. 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:

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

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

If the firm is financially constrained, so that $\hat{K}_+ = \frac{A + U^E}{\xi - \gamma^E \phi_f}$, the entrepreneur optimally co-invests all her wealth, $\hat{c} = 0$, and the financing arrangement exhausts the entrepreneur’s pledgeable income, $p\hat{X} = (pR - \xi) \hat{K}_+$. 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, $\hat{K}_+ = \bar{K}$, 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.\textsuperscript{15}

**Corollary 3 (Implementation)** The following securities implement the optimal financing agreement under a broad 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 $\lambda$ 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.

**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.\textsuperscript{16} 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 a broad mandate and $\gamma^{SR} \geq \bar{\gamma}^{SR}$, where the threshold $\bar{\gamma}^{SR}$ is decreasing in $\gamma^E$.

\textsuperscript{15}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.

\textsuperscript{16}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, $\hat{K}_\tau = K_{\tau_F}^F$, 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.
Impact therefore requires a broad mandate and that the SR fund cares sufficiently about social costs \( (\gamma^S_R \geq \hat{\gamma}^S_R) \). If the entrepreneur and the SR fund jointly internalize all externalities, \( \gamma^E + \gamma^S_R = 1 \), 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 \( K^F_C \) and \( K^S_R \) are below \( \bar{K} \), financial capital and socially responsible capital act as complements: The equilibrium clean scale with both investor types, \( \hat{K}_C \), is larger than the clean scale that can be financed in an economy with only one of the two investor types,

\[
\hat{K}_C > \max \{ K^F_C, K^S_R \}. \tag{13}
\]

2. Otherwise, there is no complementarity and \( \hat{K}_C = \bar{K} = \max \{ K^F_C, K^S_R \} \).

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, \( \hat{K}_C > K^E_C \). If \( \gamma^E < \hat{\gamma}^E \), a clean scale of \( K^E_C \) 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 \( \hat{K}_C > K^E_C \).

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, \( U^\text{SR} = -\gamma^\text{SR} \phi_D K^F_D \). This unlocks additional financing capacity, so that \( \hat{K}_C > K^\text{SR}_C \).

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 care about social costs regardless of its own investments (the “broad mandate”). The broad mandate implies that 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.

It is useful to highlight how this complementarity result interacts with additional
government regulation. When the complementarity arises – binding financing constraints and a SR fund with a broad mandate – the introduction of a Pigouvian tax would strictly reduce welfare.\textsuperscript{17} 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 a subsidy, see Proposition 1, first-best could be achieved regardless of whether a SR fund is present or not.

**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,

\[
\Delta U^{SR} := \hat{v}_C \hat{K}_C - \hat{v}_D \hat{K}^F_D > 0,
\]

the SR fund does not break even in financial terms.

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

\[
p\hat{X}^{SR} - \hat{I}^{SR} = (\pi_C - \gamma^E \phi_C) \hat{K}_C - (\pi_D - \gamma^E \phi_D) \hat{K}^F_D < 0.
\]

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

\textsuperscript{17} The result that Pigouvian taxes generally do not achieve first best in the presence of financial constraints echoes the findings of Hoffmann et al. (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.
reflects the loss in bilateral surplus for financial investors and the entrepreneur relative to their preferred agreement, which yields a joint payoff of \((\pi_D - \gamma^E \phi_D) K^E_D\). If the entrepreneur is purely profit-motivated \((\gamma^E = 0)\) she needs to be compensated for the loss of total profits, \(\pi_C \hat{K}_C - \pi_D K^E_D\).

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. (2018) and Zerbib (2019). However, if socially responsible investors were to sell their 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 \(p \hat{X}^{SR}\)).\(^{18}\)

### 4.2 The Social Profitability Index

We now derive a micro-founded investment criterion to guide scarce socially responsible capital. To do so, we extend the single-firm analysis presented in Section 4 to a multi-firm setting with limited socially responsible capital. As motivated by the conditions for impact (see Corollary 4), we focus on a SR fund with a broad mandate.

Let \(\kappa\) be the aggregate amount of socially responsible capital (we initially continue to assume that financial capital is abundant) and consider an economy with a continuum of infinitesimal firms grouped into distinct firm types.\(^{19}\) 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

\(^{18}\) 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).

\(^{19}\) The assumption that firms are infinitesimally small rules out well-known difficulties that arise when ranking investment opportunities of discrete size.
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

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

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

Given this aggregate social cost, how should a SR fund 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 from competitive financial investors only. For the remaining firm types, the payoff to the SR fund from reforming a firm of type $j$ is given by:

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

(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 of the SR fund resulting from inducing 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 associated reduction in social cost internalized by the SR fund.

Given limited capital $\kappa$, the SR fund is generally not able to reform all firms. It should therefore prioritize investments in firm types that maximize impact per dollar invested. This is achieved by ranking firms according to a variation on the classic profitability index, the social profitability index (SPI). The SPI divides the change in payoffs to the SR fund, $\Delta U_{j}^{SR}$, by the amount the SR fund needs to invest to impact the firm’s investment choice,
\[ I_j^{SR}, \]

\[ \text{SPI}_j = \mathbb{1}_{\gamma_j^E < \bar{\gamma}_j^E} \frac{\Delta U_j^{SR}}{I_j^{SR}}. \]  

(18)

**Proposition 4 (The Social Profitability Index (SPI))** The SR fund should rank firms according to the social profitability index, \( \text{SPI}_j \). There exists a threshold \( \text{SPI}^* (\kappa) \geq 0 \) such that a SR fund with scarce capital \( \kappa \) should invest in all firms for which \( \text{SPI}_j \geq \text{SPI}^* (\kappa) \).

According to Proposition 4, it is optimal to invest in firms ranked by the SPI until no funds are left, which happens at the cutoff \( \text{SPI}^* (\kappa) \). SR capital is scarce if and only if the amount \( \kappa \) 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, \( \gamma_j^E \), 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 the SR fund 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 SPI also has implications for the assortative matching between the social-mindedness of entrepreneurs and socially responsible capital (see also Green and Roth, 2021).21 As 20 The change in the payoff to the SR fund \( \Delta U_j^{SR} \) 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.

21 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.
long as the SR fund is needed to generate impact, \( \gamma_j^E < \bar{\gamma}_j^E \), there is a sense of positive assortative matching: Firms with more socially minded entrepreneurs should be prioritized because they generate larger bilateral surplus and require a smaller investment from the SR fund in order 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. The SR fund should not invest in these firms.\(^{22}\)

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 \( \lambda_j \) of a firm’s cash flow rights. This minimum cash-flow stake then pins down \( I_j^{SR} \).\(^{23}\) Given these assumptions, the SPI is given by,

\[
SPI_j = \frac{\Delta \phi_j - \Delta \pi_j}{\Delta \pi_j + \lambda_j \min \left\{ p_j R_j - \xi_j, k_{D,j} - \frac{A_j}{K_j} \right\}}.
\]

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 \( SPI_j = \frac{\Delta \phi_j - \Delta \pi_j}{\Delta \pi_j + \lambda_j (p_j R_j - \xi_j)} \). 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, \( SPI_j = \frac{\Delta \phi_j - \Delta \pi_j}{\Delta \pi_j + \lambda_j (k_{D,j} - A_j / K_j)} \). 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 \( \phi_D K \) (and therefore independent of \( A \)).

\(^{22}\) Note that this policy is only socially optimal as long as these firms are financially unconstrained.

\(^{23}\) 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 \( \lambda_j \) 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.
To conclude this section, we analyze how the composition of investor capital (and not simply its aggregate amount) matters for total surplus, motivated by the recent growth in ESG investing. Increases in the amount of capital deployed by the SR fund do not immediately translate into welfare increases because the ranking implied by the SPI does not necessarily coincide with the planner’s ranking even if $\gamma_j^E + \gamma_j^{SR} = 1$. Even though, in this case, the SR fund’s payoff from reforming a firm $\Delta U_j^{SR}$ is identical to the welfare change resulting from reforming the firm, $v_C K_C - v_D K_D^F$, the planner would always increase scale up to the efficient scale $\bar{K}$, which is strictly larger than the scale funded by the SR fund as long as the firm is financially constrained post reform, $\bar{K} < \bar{K}$. This wedge arises because the SR fund does not internalize rents that accrue to the entrepreneur. The allocation implemented by the SR fund coincides with the planner’s solution only if the firm is financially unconstrained post reform, $\bar{K} = \bar{K}$. Therefore, binding financial constraints are a key determinant of the wedge between the planner’s solution and that implemented by the SR fund.\(^{24}\)

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

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

We then immediately obtain

**Lemma 2** Assume that financial capital is fixed and abundant. Aggregate welfare is increasing in the amount of socially responsible capital $\kappa$.

Intuitively, increasing the level of socially responsible capital has strictly positive welfare effects if it reduces externalities (that would have been financed by financial

\(^{24}\) A corollary of this statement is that, if all firms are financially unconstrained, the planner’s ranking coincides with the ranking implied by the SPI.
investors) and increases the scale of clean production for the set of reformed, financially constrained firms. Because 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 socially responsible 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, $K^{SR}_{C,j} = K^E_{C,j} = \bar{K}_j$ for all firm types $j$, welfare is maximized for $x^{SR} = 1$. Otherwise, it may be strictly 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 Discussion

5.1 Delegation to a SR Fund

Our analysis so far focused on the decisions of a SR fund with a given capital endowment, highlighting the importance of the fund’s mandate for generating impact, see Corollary 4. However, given that a broad mandate entails a financial loss (see Corollary 5) the question arises how the existence of a SR fund with a broad mandate can be endogenized. We first state a “negative” benchmark:

Result 1 If individual investors are affected by social costs but are uncoordinated, infinitesimally small, and consequentialist, no individual investor invests in a SR fund with a broad mandate.

Because the benefits of socially responsible investment by the SR fund (in the form of reduced externalities) are non-rival and non-excludable, each individual investor takes social costs generated by the firm as given and, therefore does not allocate resources to a SR fund with a broad mandate (a standard result in the literature on public goods following Samuelson, 1954). In case individual investors feel responsible for the social cost generated by firms in their portfolio, they may allocate funds to SR funds with a narrow mandate. However, either way no impact is achieved. While this result on the impact of impact investment is negative, consistent with the analysis of Berk and van Binsbergen (2021), it also suggests avenues for potential solutions.

First, coordination among disperse individual investors can help ensure that at least part of the social cost is internalized. Within our model, one can interpret the case of \( \gamma^{SR} < 1 - \gamma^E \) as the reduced-from effect of imperfect coordination. While coordination is typically difficult to achieve, it is facilitated by size (see, e.g., Dimson, Karakaş and Li, 2021). Landier and Lovo (2020) offer a potential solution by positing that individual investors derive warm-glow utility from the impact of the fund. Moisson (2020) studies divestment and shareholder activism considering various moral criteria, including consequentialism, rule consequentialism, and shared responsibility.
2019). One example for measures to facilitate coordination is the establishment of the Poseidon Principles, an initiative by eleven major banks to promote green shipping, (see Financial Times, 2019).

Second, large funds, such as sovereign wealth funds or large foundations should at least partially internalize their effect on the aggregate (even absent coordination). For example, the Breakthrough Energy Catalyst (BEC) fund by the Gates Foundation provides three types of capital—philanthropic donations, sub-market equity investments, and product offtake agreements—to fund large projects that would not otherwise be financially viable, consistent with the broad mandate posited in our paper (see Financial Times, 2022). Moreover, the presence of free-rider problems could rationalize the existence of state-owned funds that, like the Norwegian Sovereign Wealth fund, invest on behalf of all their citizens rather than paying out their resource income to individuals.\footnote{This idea is related to Morgan and Tumlinson (2019) who provide a model in which shareholders of a company value public good production but are subject to free-rider problems.}

Finally, public policy can help offset the financial disadvantage of SR funds with a broad mandate by taxing the returns of SR funds with a broad mandate at a lower rate.

5.2 Social Goods and General Production Technologies

We now analyze how Proposition 2 generalizes to more than two technologies and social goods. The formal results underpinning the discussion in this section can be found in Online Appendix B.1.

Suppose that the entrepreneur has access to $N \geq 2$ production technologies characterized by technology-specific cash flow, cost, and moral hazard parameters $R_r$, $k_r$, $p_r$, $\Delta p_r$, and $B_r$. 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 et al., 2019, Albuquerque et al., 2019). Moreover, we allow for the technology-specific social cost parameter $\phi_r$ to be negative, in which case the technology generates a positive externality (a social good).
Whereas the formal expressions are unaffected by whether the externality is negative or positive, there is one important difference. If externalities are negative, a broad mandate is necessary to ensure that socially responsible investors have impact. A broad mandate reduces the outside option for socially responsible investors (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 socially responsible investors is higher under a broad mandate than under a narrow mandate (the outside option is positive under a broad 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 by socially responsible 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 $\tau$ yields negative social surplus (i.e., $v_\tau < 0$ for all $\tau$). In this case, the socially optimal scale is zero and the entrepreneur is optimally rewarded with a transfer to shut down production.

6 Conclusion

A key question in today’s investment environment is to understand conditions under which socially responsible investors can achieve impact. For example, can investors with
social concerns influence firms to tilt their production technologies towards lower carbon emissions? 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 a broad mandate for socially responsible investors. Given an abundant supply of profit-motivated capital, it is not enough for socially responsible investors to simply internalize the social costs generated by the firms they are invested in. Rather, their concern for social costs must be unconditional—independent of their own investment. This condition generates both normative and positive implications. From a positive perspective, our model implies that as most current ESG funds lack such a broad mandate, they do not have impact, consistent with Berk and van Binsbergen (2021). From a normative perspective, it states that, if society wants responsible investors to have impact, then their mandate needs to be broad. Moreover, because a broad mandate entails the sacrifice of financial returns, 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), which summarizes the interaction of environmental, social and governance (ESG) aspects. Importantly, in line with the broad mandate, the SPI accounts for social costs that would have occurred in the absence of engagement by socially responsible investors. Accordingly, it can be optimal to invest in firms that generate relatively low social returns (e.g., a firm with significant carbon emissions), provided that the potential increase in social costs, if only financially-driven investors were to invest, is sufficiently large. This contrasts with many common ESG metrics that focus on firms’ social status quo. 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. Estimating the SPI using increasingly detailed data available on emissions 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 where socially responsible investors are homogeneous. These results can be extended in a straightforward way if socially responsible investors have the same directional preferences (e.g., to lower carbon emissions), albeit with different intensity. More challenging is the case in which socially responsible investors’ objectives conflict or are multi-dimensional (e.g., there is an 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 $\hat{\nu}_r = \pi_r - \gamma^E \phi_r$. 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 $\hat{\nu}_r K^*_r(u) - u = 0$, using Equation (A.15).

Proof of Corollary 1: The result follows directly from a comparison of the net payoff to the entrepreneur, $U^{JE}$, 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 $\tau$ per unit of capital, $\phi_r$. 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 $A \geq \hat{K} (\xi - \pi_C + \phi_C)$, the firm can finance the efficient scale $\hat{K}$ for the clean technology by raising financing from financial investors, taking in to account the associated tax $\phi_C \hat{K}$. This follows from Equation (4) adjusted for “after-tax” assets $\hat{A} = A - \hat{K} \phi_C$. This proves the first statement of Proposition 1.

Case 2: If $A < \hat{K} (\xi - \pi_C + \phi_C)$, Equation (4) implies that the efficient scale cannot be achieved when raising financing from financial investors. A subsidy of (at least) $s = \hat{K} (\xi - \pi_C + \phi_C) - A > 0$ is required for the entrepreneur to finance a scale of $\hat{K}$. This proves the second statement of Proposition 1.

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

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

$$U^N_{SR} = pX^N - I^{SR} - \gamma^{SR} \phi_r \hat{K} \cdot 1_{I^{SR} > 0} \leq 0.$$  \hspace{1cm} (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 $\tau$). Second, the externality term satisfies $-\gamma^{SR} \phi_r \hat{K} \cdot 1_{I^{SR} > 0} \leq 0$ with strict equality if $I^{SR} = 0$ or $\phi_r = 0$ (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.

Broad Mandate: The proof will make 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
\( \eta \), the entrepreneur gets to make a take-it-or-leave-it offer, giving her the maximum payoff, denoted by \( U^E \), while the SR fund remains at its reservation utility \( U^{SR} \). With probability \( 1 - \eta \), the SR fund gets to make a take-it-or-leave-it offer, leading to the analogous respective payoffs of \( \overline{U}^{SR} \) and \( U^E \) (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 \( \eta \), we obtain

**Problem 1** Under a broad 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 \( \eta \),

\[
U^E (K, X^{SR} + X^F, \tau, c, 1) \geq (1 - \eta) U^E + \eta \overline{U}^E, \tag{A.3}
\]

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 a broad 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. \( \blacksquare \)

**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 \((I^F, I^{SR}, X^{SR}, X^F, K, c, \tau)\) with \( I^F \neq 0 \). Now consider the following perturbation of the contract (leaving \( K, c \), and \( \tau \) 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

\[
\begin{align*}
\ p\tilde{X}^{SR} - \tilde{I}^{SR} - \gamma^{SR} \phi_{\tau} K &= pX^{SR} - I^{SR} + pX^F - I^F - \gamma^{SR} \phi_{\tau} K \\
&= pX^{SR} - I^{SR} - \gamma^{SR} \phi_{\tau} K, \tag{A.5}
\end{align*}
\]

37
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 $\tau$, 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 $\Xi$ and the expected utility provided to the entrepreneur $u$, which satisfy

$$\Xi := pX,$$  \hspace{1cm} (A.7)

$$u := (\pi_\tau - \gamma E \phi_\tau) K + I - pX.$$  \hspace{1cm} (A.8)

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

**Problem 1**

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

subject to

$$K \geq 0$$  \hspace{1cm} (A.10)

$$K \leq \bar{K}$$  \hspace{1cm} (A.11)

$$\Xi \geq -(A + u) + (pR - \gamma E \phi_\tau) K$$  \hspace{1cm} (A.12)

$$\Xi \leq (pR - \xi) K$$  \hspace{1cm} (IC)

$$\Xi \geq 0$$  \hspace{1cm} (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 $\Xi$ 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.\footnote{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.}

We now consider the inner problem: For a fixed technology $\tau$ with $\hat{v}_\tau > 0$ and a fixed utility $u \geq \eta U^E + (1-\eta) U^E$, we solve for the optimal vector $(K, \Xi)$ as a function of $\tau$ and $u$.

**Lemma A.3** For any technology $\tau$ with $\hat{v}_\tau > 0$ and $u \geq \eta U^E + (1-\eta) U^E$, 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

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

The minimum expected repayment to investors is

\[ \Xi_r (u) = \max \left\{ (pR - \gamma^E \phi_{\tau} ) K^*_r (u) - (A + u), 0 \right\}. \]  \hfill (A.14)

**Proof:** The feasible set for \((K, \Xi)\) as implied by the five constraints \((A.10), (A.11), (A.12), (IC),\) and \((LL)\) forms a polygon (the orange region in Figure 1). Choosing the maximal scale \(K^*_r (u)\) is optimal, since, for any given \(\tau\) with \(\hat{v}_{\tau} > 0\) and any fixed \(u \geq \eta \bar{U}^E + (1 - \eta) \underbar{U}^E\), the objective function \(\hat{v}_{\tau} K - u\) is strictly increasing in \(K\) for \(K \leq \bar{K}\) and independent of \(\Xi\). 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 \(\Xi\) 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 \(\bar{K} - (A + u)\). 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 \((IC)\), \(pR - \xi\):

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

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
Lemma A.5 We now turn to the final step, the optimal technology choice.

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 \( \Xi \) (i.e., constraint (A.12), which ensures \( c \geq 0 \), shifts down by enough so that the efficient scale, \( K^*_\tau (u) = \tilde{K} \) can be achieved. In this case, there is a continuum of solutions for \( \Xi \) to support scale \( \tilde{K} \), indicated graphically by the line segment connecting the black diamond and the black circle. These solutions yield the same payoff to the SR fund, \( \hat{\nu}_\tau \tilde{K} - u \), 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 \( \hat{\nu}_\tau K^*_\tau (u) - u \) subject to \( u \geq \eta \hat{U}_E + (1 - \eta) \underline{U}_E \).

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

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

\[
\hat{\nu}_\tau K^*_\tau (u) - u = \frac{\hat{\nu}_\tau}{\xi - \gamma^E \phi_\tau} A - \frac{\xi + \gamma^{SR} \phi_\tau - \pi_\tau}{\xi - \gamma^E \phi_\tau} u \quad \text{(A.15)}
\]

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 \( u = \eta \hat{U}_E + (1 - \eta) \underline{U}_E \), we can now define the (relevant) scale as a function of the bargaining power \( \eta \), i.e.,

\[
\hat{K}_\tau (\eta) := K^*_\tau [\eta \hat{U}_E + (1 - \eta) \underline{U}_E] \quad \text{(A.16)}
\]

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

\[
U_B^{SR} = \hat{\nu}_\tau \hat{K}_\tau (\eta) - [\eta \hat{U}_E + (1 - \eta) \underline{U}_E] \quad \text{(A.17)}
\]

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

Lemma A.5 The optimal technology choice is given by

\[
\hat{\tau} = \arg \max_\tau \hat{\nu}_\tau \hat{K}_\tau (\eta) \quad \text{(A.18)}
\]
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 $\hat{v}_C K_C (\eta) > \hat{v}_D K_D (\eta)$, 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:

\begin{align*}
\bar{U}^E &= U^E + \hat{v}_F K_F - \hat{v}_p K^F, \quad (A.19) \\
\bar{U}_{SR}^E &= U_{SR} + \hat{v}_F K_F - \hat{v}_p K^F. \quad (A.20)
\end{align*}

Proof of Corollary 2: Since the SR fund has all the bargaining power, we set $u = U^E$. Then (A.14) implies that the expected repayment to investors satisfies $p \hat{X} = \Xi (U^E)$. 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 $U^E$ 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 broad-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 $\{ K_C^F, K_C^{SR} \} < \tilde{K}$, if and only if

\begin{align*}
\frac{A}{K} < \min \left\{ \frac{\xi - \pi_C, k_D (\xi - \gamma^E \phi_C)}{pR - \gamma^E \phi_D} \right\}. \quad (A.21)
\end{align*}

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

\begin{align*}
K_C^{SR} &= \min \left\{ \frac{A + U^E_{SF}}{\xi - \gamma^E \phi_C}, \tilde{K} \right\}. \quad (A.22)
\end{align*}
In contrast to (10), \( U_{SF}^E \) 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:

\[
U_{SF}^E := \max\left\{ \frac{A}{k_D} (\pi_D - \gamma^E \phi_D), \frac{A}{k_C} (\pi_C - \gamma^E \phi_C) \right\}.
\] (A.23)

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

\[
\frac{A}{\bar{K}} < \min \left\{ k_d \frac{\xi - \gamma^E \phi_C}{pR - \gamma^E \phi_D}, k_c \frac{\xi - \gamma^E \phi_C}{pR - \gamma^E \phi_D} \right\}.
\] (A.24)

Therefore, if \( \frac{A}{\bar{K}} < \min \left\{ k_d \frac{\xi - \gamma^E \phi_C}{pR - \gamma^E \phi_D}, k_c \frac{\xi - \gamma^E \phi_C}{pR - \gamma^E \phi_D} \right\} \), we obtain that both \( K_c^{SR} < \bar{K} \) and \( K_c^F < \bar{K} \). Since \( k_c \frac{\xi - \gamma^E \phi_C}{pR - \gamma^E \phi_D} > \xi - \pi_C \), this expression simplifies to (A.21).\(^{28}\)

This proves that \( \max \{ K_c^F, K_c^{SR} \} < \bar{K} \) 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, \( \hat{K}_C > \max \{ K_c^F, K_c^{SR} \} \) if and only if (A.21) holds. Else, there is no complementarity, \( \hat{K}_C = \max \{ K_c^F, K_c^{SR} \} = \bar{K} \).

**Proof:** The proof consists of two parts. We first prove that \( \hat{K}_C = \min \left\{ \frac{A + U_c^E}{\xi - \gamma^E \phi_C}, \bar{K} \right\} > K_c^{SR} = \min \left\{ \frac{A + U_{SF}^E}{\xi - \gamma^E \phi_C}, \bar{K} \right\} \) 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., \( U_c^E > U_{SF}^E \).

Second, we show that \( \hat{K}_C = \min \left\{ \frac{A + U_c^E}{\xi - \gamma^E \phi_C}, \bar{K} \right\} > K_c^F := \min \left\{ \frac{A}{\xi - \pi_C}, \bar{K} \right\} \) 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 shown that \( \frac{A + U_c^E}{\xi - \gamma^E \phi_C} > K_c^F \). We obtain

\[
\frac{A + U_c^E}{\xi - \gamma^E \phi_C} - K_c^F = \frac{A + (\pi_D - \gamma^E \phi_D)K_D^F - (\xi - \gamma^E \phi_C)K_c^F}{\xi - \gamma^E \phi_C} \geq \frac{(\pi_D - \gamma^E \phi_D)K_D^F - (\pi_C - \gamma^E \phi_C)K_c^F}{\xi - \gamma^E \phi_C} > 0,
\] (A.25)

where the first equality uses the definition \( U_c^E = (\pi_D - \gamma^E \phi_D)K_D^F \). 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, \((\pi_D - \gamma^E \phi_D)K_D^F > (\pi_C - \gamma^E \phi_C)K_c^F \), see (3).

Taken together, \( \hat{K}_C > \max \{ K_c^F, K_c^{SR} \} \) 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).  

\(^{28}\) Notice that \( k_c \frac{\xi - \gamma^E \phi_C}{pR - \gamma^E \phi_D} - (\xi - \pi_C) = (\pi_C - \gamma^E \phi_C) \frac{pR - \xi}{pR - \gamma^E \phi_C} > 0 \).

42
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 \( KE_C < \bar{K} \). 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., \( \hat{K}_C > KE_C \). Since \( KE_C \) 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 \( KE_C = \bar{K} \). The financial resource constraint implies that

\[
\hat{I} = \hat{K}k_C + \hat{c} - A = U^E - (\pi_C - \gamma^E\phi_C) \hat{K} + p\hat{X},
\]

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

\[
p\hat{X} - \hat{I} = (\pi_C - \gamma^E\phi_C) \bar{K} - (\pi_D - \gamma^E\phi_D) \hat{K} < 0,
\]

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 \( KE_D = \bar{K} \) and \( KE_C = \bar{K} \)).

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 \( \kappa \) 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

\[
SPI = \frac{\Delta U^{SR}}{I^{SR}}.
\]

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

\[
I^{SR}_{\text{min}} = (\pi_D - \gamma^E\phi_D) K^F_D - (\pi_C - \gamma^E\phi_C) \bar{K}_C.
\]

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

\[
SPI_{\text{max}} = \frac{\hat{v}_C \hat{K}_C - \hat{v}_D K^F_D}{(\pi_D - \gamma^E\phi_D) K^F_D - (\pi_C - \gamma^E\phi_C) \bar{K}_C}
= \gamma^{SR} \frac{\Delta \phi + \phi_C \left(1 - \frac{\bar{K}_C}{K^F_D}\right)}{\Delta \pi - \gamma^E \Delta \phi + (\pi_C - \gamma^E\phi_C) \left(1 - \frac{\bar{K}_C}{K^F_D}\right)} - 1.
\]
The ratio $\frac{K_C}{K_D}$ depends on entrepreneurial assets $A$. It is easily verified that in all cases (constrained and unconstrained) $\text{SPI}_{\text{max}}$ is increasing in $\gamma^E$ and $\Delta \phi$ and decreasing in $\Delta \pi$ given that $\xi - \pi_r > 0$ (see Assumption 1).

Case 1: If assets $A$ are sufficiently high, so that $\hat{K}_C = K^F_D = \hat{K}$, we obtain:

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

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

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

To see that $\text{SPI}_{\text{max}}$ is increasing in $\gamma^E$ note that $\xi - \pi_C - \frac{A}{K} > 0$ since $\hat{K} > K^F_C = \frac{\hat{A}}{\xi - \pi_C}$.

As a result, the denominator is strictly decreasing in $\gamma^E$.

Case 3: If assets $A$ are sufficiently low, so that $\hat{K}_C \leq K^F_D < \hat{K}$, then

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

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

$$\Delta \Omega = \int_{j: \gamma^E_j < \gamma^E_P \& \text{SPI}_j \geq \text{SPI}^*(\kappa)} \left( v_{C,j} \hat{K}_{C,j} - v_{D,j} K^F_{D,j} \right) d\mu(j). \quad (A.34)$$

An increase in $\kappa$ only affects the threshold $\text{SPI}^*(\kappa)$. 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, $K^E_{C,j} = K^E_{C,j} = \hat{K}_j$. 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 $v_{C,j} > 0 > v_{D,j}$) and operate at the socially optimal scale $\hat{K}_j$. 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., $\gamma^E_j < \gamma^E_P$), 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., $\max \{ K^E_{C,j}, K^E_{C,j}, K^F_{D,j} \} < \hat{K}_j$) and that total investor capital is large enough such that the following two conditions are jointly met for some $\tilde{x}^{SR} \in (0,1)$:

44
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 $\frac{A_j + U^E}{\xi_j - \gamma_F \phi_{C,j}}$.

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 $\frac{A_j + U^E}{\xi_j - \gamma_F \phi_{C,j}} > K_{C,j}^{SR}$ (see Proposition 3). This scale increase is socially valuable, implying that welfare is strictly higher for $\tilde{x}^{SR} < 1$ than for $x^{SR} = 1$. ■
References


Online Appendix

B Production technology specification

B.1 Many production technologies

Suppose that the entrepreneur has access to $N$ production technologies characterized by technology-specific cash flow, cost, and moral hazard parameters $R_{\tau}$, $K_{\tau}$, $k_{\tau}$, $p_{\tau}$, $\Delta p_{\tau}$, and $B_{\tau}$. In analogy to the baseline model, we can then define, for each technology $\tau \in \{1, \ldots, N\}$, the financial value $\pi_{\tau}$, the agency rent $\xi_{\tau}$, and the maximum scale available from financial investors $K^F_{\tau}$, maintaining the assumption that $\xi_{\tau} > \pi_{\tau}$ for all $\tau$. A straightforward extension of Lemma 1 then implies that, in the absence of investment by socially responsible investors, the entrepreneur chooses technology

$$\tau_F = \arg \max_{\tau} \left( \pi_{\tau} - \gamma^E \phi_{\tau} \right) \min \left\{ \frac{A}{\xi_{\tau} - \pi_{\tau}}, K_{\tau} \right\}.$$  \hspace{1cm} (B.35)

Equation (B.35) clarifies the entrepreneur’s relevant outside option with $N$ technologies: Any production technology dirtier than $\tau_F$ is not a credible threat. Given the credible threat $\tau_F$, the induced technology choice in the presence of socially responsible investors $\hat{\tau}$ and the associated capital stock $\hat{K}$ are given by

$$\hat{\tau} = \arg \max_{\tau} \hat{v}_{\tau} \min \left\{ \frac{A + U_E}{\xi_{\tau} - \gamma^E \phi_{\tau}}, \hat{K}_{\tau} \right\},$$  \hspace{1cm} (B.36)

$$\hat{K} = \begin{cases} \min \left\{ \frac{A + U_E}{\xi_{\tau} - \gamma^E \phi_{\tau}}, \hat{K}_{\tau} \right\} & \text{if } \hat{v}_{\tau} > 0 \\ 0 & \text{if } \hat{v}_{\tau} \leq 0 \end{cases},$$  \hspace{1cm} (B.37)

which mirrors Proposition 2.

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

$$\pi_C \left( K_{C}^{FB} \right) = \phi_C.$$  \hspace{1cm} (B.38)

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

Now consider the scenario in which technology $D$ is chosen in the absence of socially responsible investors, 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 \( A < \bar{A} \), the optimal agreement offered by socially responsible investors 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., \( \hat{K}_C < K_{FB}^C \)). In our baseline model, this case corresponds to
\[
\hat{K}_C = \frac{A + \bar{E}}{1 - \gamma \bar{E}} < \bar{K}
\]

2. If the financing constraints generated by the agency problem are intermediate, i.e., \( \bar{A} < A < A^{FB} \), the optimal agreement specifies the first-best scale, \( \hat{K}_C = K_{FB}^C \). In this case, it is efficient to increase clean scale up to the first-best level but no further, since scale above and beyond \( K_{FB}^C \) 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 \( K_{FB}^C \). 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 \( K_{FB}^C < \bar{K} \), 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).\(^{29}\) 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., \( K_{FB}^C = K_{FB}^C \). In our baseline model, this case corresponds to \( K_{FB}^C = \bar{K} \). 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., \( K_{FB}^C > K_{FB}^C \). 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, \( \hat{K}_C = K_{FB}^C < K_{FB}^C \) 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.

\(^{29}\) 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.