Are Socially Responsible Funds Viable?*

Amil Dasgupta, Dirk Jenter, Richmond Mathews, Paul Voss

This version: January 2026

Abstract

We study the viability of socially responsible (SR) funds when investors are small, have social preferences, and can cheaply invest in other ways. In a static setting, SR "impact" funds, which engage with dirty firms at a cost, cannot exist due to free riding, regardless of the type of social preferences. SR exclusion funds with reduced ownership of dirty firms can exist, but only if investors' social preferences are "warm glow," i.e., tied to the extent of ownership. In a dynamic setting, we demonstrate that such exclusion funds endogenously make impact funds viable. Impact funds profit by making dirty firms investable for exclusion funds.

*Dasgupta is at the London School of Economics; email: a.dasgupta@lse.ac.uk. Jenter is at the London School of Economics; email: d.jenter@lse.ac.uk. Mathews is at the University of Maryland; email: rmathews@umd.edu. Voss is at HEC Paris; email: voss@hec.fr. We thank audiences at the Tsinghua Conference in Theory and Finance (2025), as well as Augustin Landier, Mark Loewenstein, Michelle Lowry, Stefano Lovo, Martin Oehmke, Marcus Opp, Vincent Maurin, and Josef Zechner for helpful conversations. Dasgupta and Jenter acknowledge financial support from the LSE's Global School of Sustainability Research Fund.

Are Socially Responsible Funds Viable?

This version: January 2026

Abstract

We study the viability of socially responsible (SR) funds when investors are small, have social preferences, and can cheaply invest in other ways. In a static setting, SR "impact" funds, which engage with dirty firms at a cost, cannot exist due to free riding, regardless of the type of social preferences. SR exclusion funds with reduced ownership of dirty firms can exist, but only if investors' social preferences are "warm glow," i.e., tied to the extent of ownership. In a dynamic setting, we demonstrate that such exclusion funds endogenously make impact funds viable. Impact funds profit by making dirty firms investable for exclusion funds.

1 Introduction

The net-zero transition demands sustainable investment on a scale that public budgets and policy alone cannot deliver. The IEA estimates clean-energy spending must rise to roughly $4.5 trillion per year by the early 2030s, up from about $1.8 trillion in 2023 (IEA, 2024). Against this backdrop, the equity asset management sector is critical because of its sheer financial weight and concentrated ownership of listed firms. Global assets under management reached $128 trillion in 2024 (BCG, 2025) and – in the largest six economies globally – such institutional investors hold progressively larger percentages of corporate equity, including numerous and significant concentrated positions (Dasgupta et al, 2021). The size and concentration of the equity holdings of the asset management sector creates the possibility of two potential channels for affecting corporate strategy: a passive, exclusion channel and an active, engagement channel. By forming *exclusion funds* to overweight transition leaders and underweight firms that lag behind, asset managers can, in principle, affect firms' cost of capital, thereby penalizing polluting strategies. By forming *impact funds* that intensively engage their portfolio firms on climate issues, asset managers can, in principle, directly shape corporate sustainability actions.

Unlike proprietary investors, asset managers represent pools of delegated capital. They must therefore keep their investors on board while undertaking any actions aimed at corporate sustainability. If such actions are costly—exclusion may result in lost diversification benefits while engagement imposes direct costs—investors may be tempted to "free ride" (Grossman and Hart 1980) on the fund's efforts. Even if individual investors are individually socially-minded, as long as they are small, they will be tempted to escape the fund and enjoy any benefits from its activities for free, thus drastically limiting the effectiveness of pooled capital to shape corporate sustainability. This is because small investors will not internalize the externality caused by their individual actions.

Such free-riding is particularly problematic for the active, *engagement* activities of asset managers, i.e., for impact funds. While a degree of exclusion may be tolerated (or even welcomed) by socially-minded investors who experience disutility from direct ownership of dirty firms, i.e., those with so called "warm glow" or "values aligned" preferences (to adopt the nomenclature laid out in Dangl et al, 2025)[1], no such mitigation exists for impact funds that engage directly: such funds

[1]For evidence of the existence of such preferences, please see Reidl and Smeets 2017, Bauer, et al 2021, and Cole, et al 2023.

must both hold dirty assets in their portfolio (generating disutility for investors with values aligned preferences) and expend costs to clean up such firms (generating direct costs for all fund investors *regardless* of preferences). Yet, there is now growing evidence that passive, exclusion strategies have limited effect on corporate behavior (Heath et al, 2023, Berk and van Binsbergen, 2025, and Pedersen, 2025). Thus, for asset managers to aid the net zero transition, it is key to understand the viability of green impact funds.

In this paper, we characterize such conditions under which green impact funds can exist. We first demonstrate that, in a static setting, across a full range of feasible types of social preferences for individual investors, *only* "imperfect" exclusionary funds can exist: such funds will only *partially* exclude dirty firms from their portfolios. However, we also show that the very existence of imperfect exclusionary funds engender conditions for impact funds to exist. In a dynamic setting, the existence of exclusion funds creates the opportunity for impact funds to buy low (because dirty firms trade at a discount), expend costs in cleaning up dirty firms, and then sell high (because clean firms trade at a premium), thus recouping the costs of engagement. In other words, our theoretical analysis provides a foundation for the full landscape of green funds: it is possible for exclusion funds and impact funds to coexist in a symbiotic relationship, reinforcing each other's effect on the net zero transition. At a positive level, our analysis is consistent with the finding in Abis et al (2024) and Yang and Yasuda (2025) that only a small fraction of US socially responsible funds seek impact, while the rest are exclusionary funds, and with Frankel et al.'s (2024) observation that US socially responsible funds report their objective as variants of "long-term capital appreciation" rather than "impact."

While providing potentially reassuring conditions for the existence of green impact funds, our model also highlights significant limits on the role that the asset management sector can play in promoting sustainability. Such limits follow from the symbiosis at the core of our theoretical mechanism. First, the size of the price effect of exclusion directly affects the efficiency of impact capital: when exclusion generates only a small discount (as seems the empirical evidence seems to suggest), it will take a large amount of impact capital to make a difference. This is because the per-investor willingness to pay for engagement within an impact fund depends on the per-investor trading profit gained from leveraging the presence of exclusion funds. When exclusion funds do not generate large profit opportunities, the per-impact-investor willingness to pay is small, and so

many such investors, i.e., a large allocation to impact funds, will be necessary to undertake even relatively small engagements. Second, even if they are unconstrained in size, and if engagement is inexpensive, impact funds can only clean up a strict *subset* of dirty firms in order to be viable. This is because, in order to profit from the presence of exclusion funds, impact funds must be able to buy low, which—in turn—means that the market cannot know precisely which subset of dirty firms they will purchase in order to clean up. Otherwise, the full price impact of their clean-up activities would be fully reflected in the purchase price, suppressing the symbiosis at the heart of our model. Indeed, it is clear that the larger the fraction of dirty firms the impact fund cleans up, the lower its informational advantage, and thus the lower its achievable profits to offset its engagement costs.

Our contribution to the literature is novel along several dimensions. First, we provide a novel characterization of what impact funds are. In Oehmke and Opp (2025)—who study a setting where profit seeking capital is in perfectly elastic supply, so that exclusion has no price impact—impact funds can only exist by trading off financial returns for social good. We show that when there is *any* non-zero price effect from exclusion, impact funds can exist *without* such a trade off. These funds will be no less profitable than other funds in equilibrium, because they profit from the existence of the exclusionary sector. This is consistent with US impact funds reporting value maximization, rather than impact, as their objective (Frankel et al., 2024). By the same token, in contrast to Broccardo et al (2022) who treat engagement and exclusion as alternative strategies, our model connects the price effects of exclusion to the feasibility of engagement.

Second, we provide a novel characterization of when impact funds can exist. Oehmke and Opp (2025) show that impact funds can attract capital but require either that investors experience warm glow utility specifically from investing in an impact *fund* (which excludes the possibility of free riding),[2] or that there are some consequentialist investors who internalize the effect of not joining an impact fund, because they are large or because they coordinate with other investors. We present a different story. In our model, all investors are small and unable to coordinate, and experience warm glow utility from holding clean *assets* in their portfolio, whether through a fund or directly: thus free riding is always possible. Nevertheless, impact funds exist because they benefit from the existence of endogenously formed exclusion funds. In other words, we provide a characterization of

[2]Similar preferences feature in Landier and Lovo (2025), who are primarily interested in the consequences of socially responsible investment as opposed to the existence of funds per se.

the full landscape of socially responsible funds: such funds exist via intricate interdependencies; it is not possible to think of different types of socially responsible funds in isolation.

Finally, while our buy-low-sell-high mechanism has a ring of the classic microstructure literature started by Kyle and Vila (1991), our mechanism is distinct. In those papers, a large (proprietary) investor buys into a firm and engages management to increase value. This generates free riding—the engaging entity pays the cost but everyone who owns the shares enjoys the benefits—that can be mitigated if the large investor has some private information. However, green engagement in our model does not *directly* generate increased cash flows that justify a higher price, so even in the presence of private information, such engagement would not occur, as there would be no gains to capture. Instead, the price impact of green investment is generated *indirectly* because of the (endogenous) existence of exclusionary funds.

Static analysis and results We first study a static, CARA-Normal model with two risky assets and a risk-free asset with return normalized to unity. One of the risky assets is clean while the other is dirty: the dirty asset generates a non-tradable externality. All asset and externality cash flows are independent. There are two types of investors: nonfinancial investors who have social preferences—either consequentialist (wherein agents experience disutility from the *existence* of the externality) or warm glow (wherein agents experience disutility from *holding* units of the dirty asset)—and financial investors who have standard preferences. Financial investors have full market access. Nonfinancial investors can also trade on their own, but face an *infinitesimal* cost of market access, which is eliminated if they join a collective fund of any size. We do not model fund formation or contracting in this paper, but refer the reader to Dasgupta and Mathews (2025) for a full analysis. The fact that nonfinancial investors can access financial markets at an infinitesimal cost means that any fund that imposes strictly positive costs on investors, e.g., in the form of lost diversification and/or engagement costs, would be exposed to the loss of investors due to free riding.

We consider two possible types of socially responsible funds, and ask whether they can exist in isolation given the potential for investor free riding. The two fund types are:

1. An impact fund (IF) that holds both assets and pays the cost c to clean up the brown asset, eliminating the externality.

2. An exclusionary fund (EF) that does not pay the clean up cost and trades both assets, but purchases *less* of the dirty asset to take into account nonfinancial investors' disutility. The adjusted mean of the externality distribution in case an exclusionary fund exists is lower than the ex ante mean by an amount that all investors forecast correctly in equilibrium.

A *fully* exclusionary fund, which does not hold the dirty asset at all, is a special case of the EF fund type.

We first show that IFs never exist in the static model with *either* consequentialist or warm glow preferences. The reason is as follows. If any type of collective investment vehicle is to exist in equilibrium, it must offer each investor utility at least equivalent to what she can get by investing directly, *taking into account* the existence of the fund (and thus any effect the fund has on the externality). Suppose, counterfactually, an IF exists. Then, each IF investor will bear a proportional share of the clean-up cost, c, that is shared among all investors in the fund. The investor's alternative is to trade directly to the same portfolio (giving them identical risk-sharing benefits), without bearing any of the clean-up cost, at the incremental, infinitesimal cost of direct market access. Since the proportionate share of c is of greater magnitude, this deviation payoff is clearly higher. Further, because the externality is eliminated by virtue of the existence of the IF, the specific preference type (consequentialist or warm glow) does not matter. In both cases, therefore, free riding motives preclude the existence of the IF.

We then show that EFs also cannot exist in a static model with consequentialist investors. The reason is very similar to the case of impact funds. Consequentialist investors only care about the overall level of the externality. Any EF that actually holds less of the dirty asset than the level required by perfect risk sharing in a competitive equilibrium—and thus causes a reduction in the externality—would impose a non-inifinitesimal diversification loss on its investors. If investors are consequentialist, they will prefer to simply pay the infinitesimal cost to directly trade to optimal risk sharing, and enjoy the fund's exclusion effect on the externality without paying the diversification cost. However, when investors have warm glow preferences, the implications are different. We now show that a uniquely (privately) optimal EF can exist, which features a degree of exclusion that is limited by each warm glow nonfinancial investor's willingness to sacrifice risk sharing benefits to avoid the direct disutility from the externality. Thus, such a fund would trade to a position that

aggregates the positions that individual warm glow investors would trade to without coordination, and all warm glow investors would join such a fund in order to save on the infinitesimal cost of direct market access. It is noteworthy that such an exclusionary fund would not, other than in cases where warm glow preferences are extreme, entirely exclude the dirty asset; this is because the fund's holdings are determined by trading off diversification (that encourages ownership of the dirty asset) and warm glow disutility (that discourages it). However, it is possible to interpret this fund as made up of a "fully" exclusionary fund and an index fund, where investors split their investment between the two.

Dynamic analysis and results To model the dynamic symbiosis between EFs and IFs, we extend our static model into a dynamic context, while simplifying the problem as much as possible. In particular, we assume that there are now three risky assets: one clean and two (symmetric) dirty assets. The need for multiple dirty assets is a minimal change required to endow the IF with the ability to trade at prices that are not fully revealing. Apart from this, the dynamic model is deliberately kept as simple as possible: there are only two trading dates and full temporal independence in cash flows and stochastic externality payoffs across periods. Further, also for simplicity, we let the measure of financial agents (with non-social preferences) shrink to zero, and endow nonfinancial agents with warm glow preferences (as those are the conditions under which the EF exists in the static model).

The timing of the dynamic model is as follows. At the outset, non-financial investors sort into the exclusionary fund (EF) and impact fund (IF) to maximize expected utility. At $t = 1$, the EF and IF simultaneously trade and first period payoffs are realized. The impact fund then decides whether to clean up *one* dirty asset at cost c. At $t = 2$: EF and IF observe previous $t = 1$ holdings and which dirty asset has been cleaned up, if any. Then, the EF and IF again trade and second period payoffs are realized. We conjecture an equilibrium in which at $t = 1$, the IF and EF invest identically—on a per investor basis—in the clean asset, while the EF demands the two dirty assets symmetrically. The IF randomly selects a dirty asset to clean up and then demands more of that asset and less of the other dirty asset. At $t = 2$, the EF and IF select identical per-investor portfolio, i.e., there is no engagement in the second period.

In our main result, we show that there exist two equilibria in which both an IF and an EF form,

the IF randomizes over which dirty asset to clean up, and nonfinancial investors are indifferent between joining the IF vs the EF. The two equilibria differ in the weight of capital that accumulates in the IF vs the EF: in one equilibrium the IF attracts a small number of investors while the EF attracts many; in the other, the IF features many investors, while the EF is small. Intuitively, when the IF is small and the EF is large, there will be a large difference in the first and second period price of the dirty asset that is cleaned up (little money is invested in the first period by those with access to the private information that this asset will be cleaned up). Simultaneously, with a small IF, there are fewer investors to share in the cost of cleanup, so they must be compensated for this cost with high profits via the large price change upon sale of the cleaned up asset. On the other hand, when the IF is large, the price difference across the two periods is smaller as more first period investors are "in the know." In this case, IF investors have to spread the cost of cleanup more widely since they get lower profits from selling the cleaned up asset in the second period. Thus, as discussed above, the size of the price effect of exclusion directly affects the efficiency of impact capital: when exclusion generates only a small discount it will take a large amount of impact capital to make a difference. This is because the per-investor willingness to pay for engagement within an IF depends on the per-investor trading profit gained from leveraging the presence of the EF. When the EF is small and doesn't generate large profit opportunities for the IF, the per-IF-investor willingness to pay is small, and so many such investors, i.e., large IF, will be necessary to undertake even relatively small engagements.

Related literature

Our paper contributes to the burgeoning set of papers on sustainable investing—too numerous to survey exhaustively here—wherein corporate equity owners can influence firms either by engagement (see Broccardo et al, 2022, as well as a recent survey by Sautner, 2025) or via trading strategies that affect their cost of capital (see Heath et al, 2023, Berk and van Binsbergen, 2025, Green and Roth, 2025, as well as surveys by Pastor, Stambaugh, and Taylor, 2024, and Pedersen, 2025). Within this literature, our analysis relates most closely to papers that focus on the role of socially responsible *funds*, in particular, Oehmke and Opp (2025) and Landier and Lovo (2025). The focus of these papers is different from ours: these authors are primarily interested in the *consequences* of the existence of impact funds.

Aspects of our paper are also connected directly to Dangl et al's (2025) results. While they do not model investment funds, they find that—though consequentialist preferences have no impact in equilibrium without explicit coordination into a fund (which we show cannot occur)—warm glow preferences can have impact via exclusion even without coordination. Our static model results on exclusionary funds are closely aligned with this finding: indeed, absent the infinitesimal cost of direct market access, the optimal exclusionary fund could be replicated by individual trading (and thus warm glow investors can have aggregate impact without coordination).

Several papers in the sustainable investment literature focus directly or indirectly on free riding. For example, the free riding arising from consequentialist preferences is discussed in Hakenes and Schliephake (2024). Further, Gryglewicz et al (2025) focus on the interaction of internal and external free rider problems in sustainability. However, none of these papers focus on the endogenous existence of socially responsible funds.

Gollier and Pouget (2022) consider whether the price impact of exclusion can induce a large activist proprietary blockholder to engage. While in that sense their analysis bears a similarity to aspects of our dynamic analysis, they do not consider endogenous fund formation in the presence of free riding. Their blockholder is assumed to have long horizons and to be (exogenously) sufficiently socially responsible; whereas in our model the blockholder is endogenously formed from small investors, and the size of the fund scales endogenously to adjust to the cost of engagement and the price impact of exclusion. Gupta, Kopytov, and Starmans (2026) show that the opportunity to realize gains from trade by selling dirty firms to socially responsible investors who desire to clean them up can cause current firm owners to delay clean-up efforts. They assume the existence of a large socially responsible investor that can acquire the firm and clean it up with no need to attract or retain capital. The contemporaneous work of Guo (2025) has a focus that is similar to ours. Guo (2025) analyzes a model that provides a distinct justification for the existence of impact funds; however, in her model impact funds must be supported by subsidized lending provided by banks that themselves benefit from subsidized depositor financing due to a preference for liquid assets.

More generally, our paper relates to the literature on the role of equity holders in monitoring the firm via exit and voice surveyed by Edmans and Holderness (2016) and the role of asset managers in such monitoring, surveyed by Dasgupta et al (2021). Finally, in providing a justification for the landscape of socially responsible funds—impact and exclusionary funds coexist in our model—and

relating them to investor preferences, our paper bears a thematic connection to Malenko, et al (2025). They present an imperfect competition model based on Salop (1979) to study how growing heterogeneity in investor preferences affect the fund industry. In their model, however, investors cannot invest by themselves, funds do not affect externalities, and funds are symmetric, so they do not speak to either the existence of impact funds or the coexistence of impact and exclusionary funds.

A recent empirical literature analyzes the prospectuses of active equity funds in the US to document the increasing use of information about firms' environmental, social, and governance (ESG) performance in investment strategies.[3] Abis et al. (2024) and Yang and Yasuda (2025) show that most socially responsible funds use ESG information for negative or positive screening, i.e., to exclude 'bad' firms or restrict their investments to 'good' firms. Our model's explanation is that they are serving investors with warm-glow preferences. Fewer SR funds engage in what Abis et al. call "impact activism," i.e., are actively trying to change firms for the better. Our model suggests that such impact funds can exist only because of the significant exclusionary funds sector documented by these papers. Finally, Frankel et al. (2024) note that SR funds report variants of "long-term capital appreciation" as their investment objective, rather than "environmental / social impact," which is consistent with free-riding forcing SR funds to deliver competitive financial performance.

2 A Static Model of Socially Responsible Funds

2.1 Assets and Investors

We consider a static model with a single trading period. The gross risk-free rate is normalized to 1. There are two risky assets, green (G) and brown (B), with terminal cash flows:

$$\tilde{z}_G \sim \mathcal{N}(\mu_G, \sigma_G^2), \quad \tilde{z}_B \sim \mathcal{N}(\mu_B, \sigma_B^2).$$

The assets are in exogenously fixed supply (s_G, s_B). The brown asset creates a non-tradeable externality $\tilde{y} \sim \mathcal{N}(\mu_y, \sigma_y^2)$. The externality asset and the terminal cash flows of the two assets are all independently distributed.

[3] See, for example, Abis et al. (2024), Frankel et al (2024), Andrikogiannopoulou et al. (2025), and Yang and Yasuda (2025).

Each investor has exponential utility defined over their terminal wealth, W:

$$u = -e^{-\frac{W}{\rho}}.$$

All investors start with zero endowment in both risky assets. There are two types of investors: financial investors with mass $1-\lambda$, and nonfinancial investors with mass λ. Financial investors have standard preferences, are not directly affected by the externality, and trade on their own at no cost. Their terminal wealth (ignoring cash holdings WLOG) is given by

$$W = \sum_{i \in \{G,B\}} \alpha_i(\tilde{z}_i - P_i),$$

where P_i is the price of asset i and α_i is the investor's final ownership stake in asset i.

Nonfinancial investors can trade on their own at an infinitesimal cost of market access, m, and hold a non-tradable short position of size b in the externality asset, which generates disutility (background risk) that affects their terminal wealth. We consider two types of preferences for these investors, which, following Dangl, et al (2025), we label as "consequentialist" or "warm glow"/"cold prickle" preferences, respectively. Note that we treat warm glow and cold prickle as being equivalent in the paper throughout, as in Dangl, et al (2025) and use the term cold prickle henceforth. The difference in these preferences lies in whether the investor's ownership of the brown asset affects the intensity of their disutility. Specifically, we assume the disutility measured in units of terminal wealth is given by $-\alpha_B(1_{cp}+\frac{1_c}{\alpha_B})b\tilde{y}$, where 1_c is an indicator function for consequentialist preferences, and 1_{cp} is an indicator function for cold prickle preferences. The terminal wealth of nonfinancial investors is therefore given by

$$W = \alpha_B(\tilde{z}_B - P_B - (1_{cp} + \frac{1_c}{\alpha_B})b\tilde{y}) + \alpha_G(\tilde{z}_G - P_G).$$

2.2 Funds and Externality Reduction

Since nonfinancial investors face a small cost m of trading on their own, they will prefer to invest through investment funds if they offer a lower cost of market access without otherwise sacrificing utility. For simplicity we assume all investment funds offer market access at a cost of zero—our

qualitative results will remain unchanged as long as these costs remain weakly lower than the cost of direct access for nonfinancial investors. We abstract from the formation and internal contracting of such funds, and assume that funds' mandates (i.e., their investment objectives) are exogenously given. See Dasgupta and Mathews (2025) for details of fund formation and contracting. As we let the cost m approach zero (which we do for all of our existence results below), any nonfinancial investor will invest through a larger fund only if it gives them at least equivalent utility to investing on their own. This is a very strict condition for existence of a given fund type, which we argue is justified by the existence of plentiful traditional investment options like index funds in the real world that offer easy market access and ample diversification. We further assume that all investors are price-takers, even when aggregated into funds. This can be justified in the case of large funds by arguing that if funds trade strategically, investors can always do better by trading on their own with no price impact, which will always provide them with superior risk-sharing benefits.

Our interest is in the existence and potential externality-reducing effects of different types of socially responsible (SR) funds given nonfinancial investors' disutility from the externality. Such funds take different forms, with their mandates generally involving (full or partial) exclusion of some assets, or some level of direct governance activity ("impact") designed to specifically address externalities generated by their portfolio assets. To model the potential effects of these policies as simply as possible, we assume the brown asset's externality can be reduced if the price of the brown asset is low (due to some level of exclusion), or if an impact fund exerts effort to clean up the asset. We assume the externality is completely eliminated if a fund pays the fixed cost c. In the absence of such a clean-up effort, we assume the adjusted mean of the externality distribution in case an exclusionary fund exists is given by $\hat{\mu_y} \leq \mu_y$, which all investors will forecast correctly in equilibrium.

We consider two possible types of SR funds, and ask whether they can exist in equilibrium given the investors' outside option of investing on their own at a vanishingly small cost. The two fund types are:

1. An impact fund (IF) that trades both assets and pays the cost c to clean up the brown asset, eliminating the externality.

2. An exclusionary fund (EF) that does not pay the clean up cost and trades both assets, but

purchases less of the brown asset to take into account nonfinancial investors' disutility.

Note that a fully exclusionary fund type, which does not trade the brown asset at all, is just a special case of the EF fund type, so these will be analyzed together.

2.3 Existence of Socially Responsible Funds

If any type of SR fund is to exist in equilibrium, it must offer investor utility at least equivalent to what they can get by investing on their own *taking into account* the existence of the collective fund (and thus any impact the fund has on the externality). First consider the potential existence of an impact fund (IF) that pays the cost c to clean up the brown asset. In such a fund, both the green asset and brown asset will be held, and nonfinancial investors will suffer no disutility since the externality is eliminated. Thus, the optimal holdings of the green and brown asset by an impact fund (if it maximizes investor utility, which gives it the best chance of survival) are exactly what they would be if the nonfinancial investors traded on their own. However, each IF investor will bear a proportional share of the clean-up cost, c, that is shared among all investors in the fund. Their alternative is to trade on their own (giving them identical risk-sharing benefits), and not bear any of the clean-up cost. Since the externality is eliminated, the preference type (consequentialist or cold prickle) does not matter. In both cases, therefore, free riding incentives give us a straightforward result.

Proposition 1. *With either consequentialist or cold prickle preferences, no impact fund can exist in equilibrium at any size because of free riding.*

This is similar to the finding in Oehmke and Opp (2025) for consequentalist investors and extends their result to cold prickle preferences that are not tied to investment in a particular fund. In Corollary 8 of their paper, Oehmke and Opp (2025) assume that investors experience a warm glow specifically from joining the impact fund, and that this can justify the existence of an impact fund. However, we show here that if—instead—investors experience warm glow preferences at the level of their holdings of the actual *asset*, impact funds cannot exist.

The situation with exclusionary funds is different, as there are no direct costs shared by fund investors that can be avoided by defecting to trade on their own. To proceed, we first derive the holdings of an "optimal" EF that trades both assets in a way that maximizes the utility of

investors *taking the level of the externality as given.* Note that since the green and brown assets are independent, the optimal holdings are also independent. It is straightforward to show that the optimal holding of the green asset irrespective of the type of preference is therefore $\alpha_G = s_G$ for both financial and nonfinancial investors, with accompanying price $P_G = P_G^c \equiv \mu_G - \frac{s_G \sigma_G^2}{\rho}$, just as they would be in a competitive equilibrium with financial investors only (i.e., if we set $\lambda = 0$). Thus, the crux of the optimization boils down to the fund's optimal holding of the brown asset.

First consider the case with consequentialist preferences for nonfinancial investors. Taking the mean of the externality as given, an optimal EF assumes that it cannot affect the size of the externality via its trading decisions. As such, it trades to the competitive allocation for the brown asset, i.e, a per-investor holding of $\alpha_B = \alpha_B^c \equiv s_B$, leading to a price of $P_B = P_B^c \equiv \mu_B - \frac{s_B \sigma_B^2}{\rho}$ (as in a competitive equilibrium with only financial investors). Thus, there is no exclusion and the mean of the externality is μ_y. Using a similar free-riding argument as above, this will be the *only* feasible exclusionary fund with consequentialist preferences; if there is any actual exclusion of the brown asset, an individual nonfinancial investor will find it optimal to defect and trade on their own like a financial investor, enjoying any effect of exclusion on the asset's price without suffering from the concomitant suboptimal risk sharing.

Next consider the case with cold prickle preferences. Taking the mean of the externality as given, an optimal EF will maximize the following with respect to its holding of the brown asset on a per-investor basis:

$$\underset{\alpha_B}{max} \mathbb{E}\left[-e^{-\frac{\alpha_B(\tilde{z}_B - P_B - b\tilde{y})}{\rho}}\right],$$

which is equivalent to maximizing the certainty equivalent

$$\alpha_B\left(\mu_B - P_B\right) - \frac{\alpha_B^2 \sigma_B^2 + \alpha_B^2 b^2 \sigma_y^2}{2\rho} - b\alpha_B \hat{\mu}_y,$$

where $\hat{\mu_y}$ is their belief about the final mean of the externality (which will be correct in equilibrium). The FOC is then

$$\mu_B - P_B - \frac{\alpha_B(\sigma_B^2 + b^2\sigma_y^2)}{\rho} - b\hat{\mu}_y = 0,$$

and yields

$$\alpha_B^*(EF) = \frac{\rho\left(\mu_B - P_B - b\hat{\mu}_y\right)}{\sigma_B^2 + b^2\sigma_y^2}.$$

The financial investors, also acting as price takers, will maximize

$$\alpha_B \left(\mu_B - P_B\right) - \frac{\alpha_B^2 \sigma_B^2}{2\rho}$$

which has FOC

$$\mu_B - P_B - \frac{\alpha_B \sigma_B^2}{\rho} = 0$$

and solution

$$\alpha_B^*(FI) = \frac{\rho\left(\mu_B - P_B\right)}{\sigma_B^2}.$$

Market clearing gives

$$(1-\lambda)\frac{\rho\left(\mu_B - P_B\right)}{\sigma_B^2} + \lambda\frac{\rho\left(\mu_B - P_B - b\hat{\mu}_y\right)}{\sigma_B^2 + b^2\sigma_y^2} = s_B,$$

which yields a market price of

$$P_B = \mu_B - \frac{\sigma_B^2\left(s_B\left(b^2\sigma_y^2 + \sigma_B^2\right) + b\lambda\rho\hat{\mu}_y\right)}{\rho\left(\sigma_B^2 + b^2(1-\lambda)\sigma_y^2\right)}.$$

The price clearly corresponds to P_B^c only if $b = 0$. Plugging this price into the optimal EF's per-investor holding yields a stake of

$$\alpha_B^*(EF) = \frac{s_B\sigma_B^2 - b(1-\lambda)\rho\hat{\mu}_y}{\sigma_B^2 + b^2(1-\lambda)\sigma_y^2},$$

which is clearly lower than the per-investor holdings in an equilibrium with only financial investors (i.e., $\alpha_B^* = s_B$). Thus, the optimal EF does have some level of exclusion as long as b is positive. To avoid dealing with short-sale constraints, we assume from here forward that b and $\hat{\mu}_y$ are such that this expression is non-negative. Then, we have the following result.

Proposition 2. *Exclusionary funds cannot exist in equilibrium with consequentialist preferences. With cold prickle preferences, the only exclusionary fund that can exist is an optimal exclusion fund that gives its investors a per-investor stake in the brown asset equal to* $\alpha_B^*(EF) = \frac{s_B\sigma_B^2 - b(1-\lambda)\rho\hat{\mu}_y}{\sigma_B^2 + b^2(1-\lambda)\sigma_y^2} \leq \alpha_B^c$.

Thus, even with cold prickle preferences, the result above shows that the optimal EF will only

partially exclude the brown asset, unless b is very large. This is because a fully exclusionary fund would impose too high a diversification cost on even socially-minded investors. However, it is noteworthy that we can interpret our optimal EF above as made up of a fully exclusionary fund and an index fund, where investors split their investment between the two.

Overall, in a static model where investors have the ability to invest on their own at a vanishingly small cost, impact funds designed to take costly actions to reduce externalities will never attract investors, while exclusionary funds can only exist if investors have cold prickle preferences. Furthermore, in the latter case the degree of exclusion (and presumably therefore the degree of externality reduction) is limited by investors' willingness to give up risk sharing benefits to avoid the direct disutility from the externality, and is therefore likely to be suboptimal from a social welfare point of view. In the next section we show that in a dynamic model, the existence of such funds could open the door to impact funds, who can exert effort to clean up assets and then sell them at higher prices to exclusionary funds.

3 A Dynamic Model of Socially Responsible Funds

3.1 The Model

We now extend the static model in several ways to allow for dynamic pricing effects, and to create the possibility of a profitable strategy for impact funds. There are two trading periods, $t \in \{1, 2\}$, and three risky assets: one green asset, G, and two symmetric brown assets $B1$ and $B2$. Their per-period cash flows at time t are given by:

$$\tilde{z}_{G,t} \sim \mathcal{N}(\mu_G, \sigma_G^2), \quad \tilde{z}_{B1,t} \sim \mathcal{N}(\mu_B, \sigma_B^2), \quad \tilde{z}_{B2,t} \sim \mathcal{N}(\mu_B, \sigma_B^2).$$

Supply of the three assets is fixed at (s_G, s_B, s_B), respectively. The brown assets create per-period externalities $\tilde{y}_{B1,t} \sim \mathcal{N}(\mu_y, \sigma_y^2)$ and $\tilde{y}_{B2,t} \sim \mathcal{N}(\mu_y, \sigma_y^2)$. All cash flows and externalities are independent across assets and across time.

We let $\lambda \to 1$ for tractability (i.e., there are no financial investors), and assume all investors have cold prickle preferences and face disutility (background risk) from any per-period externality

$\tilde{y_{i,t}}$ at time t equal to $-\alpha_{i,t} b \tilde{y_{i,t}}$.[4] Investors have no endowments in any assets. Terminal wealth is therefore given by

$$\begin{aligned} W &= \sum_{i \in \{G,B1,B2\}} \alpha_{i,1}(\tilde{z}_{i,1} - P_{i,1} - 1_{i \neq G} b\tilde{y}_{i,1}) + \alpha_{i,2}(\tilde{z}_{i,2} - 1_{i \neq G} b\tilde{y}_{i,2}) + (\alpha_{i,1} - \alpha_{i,2}) P_{i,2} \\ &= \sum_{i \in \{G,B1,B2\}} \alpha_{i,1}(\tilde{z}_{i,1} + P_{i,2} - P_{i,1} - 1_{i \neq G} b\tilde{y}_{i,1}) + \alpha_{i,2}(\tilde{z}_{i,2} - P_{i,2} - 1_{i \neq G} b\tilde{y}_{i,2}), \end{aligned} \tag{1}$$

where $1_{i \neq G}$ is an indicator function that equals one for the two brown assets and zero for the green asset.

Investors invest through one of two types of funds: (1) an optimal exclusion fund which, like that derived in Proposition 2 above, maximizes investor utility subject to given expected values of the externalities, or (2) an impact fund that can clean up one brown asset at cost c, leaving that asset's first period externality unchanged but eliminating its second period externality. Investors optimally distribute themselves across the two funds, with a proportion η investing via the impact fund and the remaining $1 - \eta$ investing via the exclusionary fund. For simplicity, exclusion by itself has no impact on the externalities.[5] The timing of the model is as follows:

1. Non-financial investors sort into the exclusionary fund (EF) and impact fund (IF) to maximize expected utility and stay with the same fund for both periods.[6]

2. $t = 1$: The EF and IF simultaneously trade as price takers, and first period payoffs are realized. The IF then decides whether to clean up one brown asset at cost c.

3. $t = 2$: The EF and IF observe previous period holdings and whether any brown asset (and which one) has been cleaned up. Then, the EF and IF again trade as price takers, and second period payoffs are realized.

[4] Adding financial investors to this setting would result in their joining the impact fund in equilibrium because of its trading profits.

[5] This assumption makes no qualitative difference to our analysis, but has some normative implications that are discussed after Proposition 3.

[6] The assumption that investors stay with the same fund in both periods is without loss of generality because, as will be clear later, in the second period impact funds and exclusions funds hold identical portfolios.

3.2 Analysis

We first conjecture a set of equilibrium strategies, and then verify the existence of such an equilibrium. In our conjectured equilibrium:

- At $t = 1$, the IF and EF invest the same amount in asset G on a per-investor basis, while the EF treats the two brown assets symmetrically (submitting the same demand schedule for each). The IF randomly selects a brown asset to clean up, then demands more of the brown asset it will clean up and less of the other brown asset. Without loss of generality, for notational convenience we assume asset $B1$ gets cleaned up but only the IF knows this.
- At $t = 2$, the EF and IF select the same per-investor portfolio.

Because of the nature of the preferences and the independence of all random variables, we have the following result.

Lemma 1. *Each fund optimizes its second period holdings irrespective of its first period holdings. Furthermore, second period actions affect funds' first period choices only through the anticipated second-period prices of the assets.*

The first statement in the lemma follows from the assumption that cash flows and externalities are independent across assets and across time. The second statement is a consequence of price-taking behavior: funds assume their actions cannot do not affect prices, so they just use their forecast of second period prices to inform their first period trades.

We solve the model by backward induction, starting at $t = 2$. In the conjectured equilibrium, assets $B1$ and G are both clean in the second period, so each will be held in the same proportion by both funds, with prices and per-investor allocations as in a competitive one-period equilibrium with only financial investors. That is, final holdings for all investors are $\alpha^*_{i,2} = s_i$ and prices are $P_{i,2} = \mu_i - s_i \frac{\sigma_i^2}{\rho}$ (as discussed above) for $i \in \{G, B1\}$. For asset $B2$, which is still dirty, the representative investor's problem is

$$\max_{\alpha_{B2,2}} -\mathbb{E}_2\Big[\exp\big(-\frac{\alpha_{B2,2}(\tilde{z}_{B2,2} - P_{B2,2} - b\tilde{y}_{B2,2})}{\rho}\big)\Big], \tag{2}$$

which is equivalent to maximizing the certainty equivalent

$$\underset{\alpha_{B2,2}}{max} \left\{ \alpha_{B2,2}(\mu_B - P_{B2,2} - b\mu_y) - \frac{1}{2\rho}\alpha_{B2,2}^2\left(\sigma_B^2 + b^2\sigma_y^2\right) \right\}. \tag{3}$$

The FOC is

$$\begin{aligned} & (\mu_B - P_{B2,2} - b\mu_y) = \frac{1}{\rho}\alpha_{B2,2}\left(\sigma_B^2 + b^2\sigma_y^2\right) \\ \iff & \alpha_{B2,2}^* = \frac{\rho(\mu_B - P_{B2,2} - b\mu_y)}{\left(\sigma_B^2 + b^2\sigma_y^2\right)} \end{aligned} \tag{4}$$

Market clearing yields

$$\begin{aligned} & 1\alpha_{B2,2}^* = s_B \\ \iff & \frac{\rho(\mu_B - P_{B2,2} - b\mu_y)}{\left(\sigma_B^2 + b^2\sigma_y^2\right)} = s_B \\ \iff & P_{B2,2} = \mu_B - b\mu_y - s_B\frac{\left(\sigma_B^2 + b^2\sigma_y^2\right)}{\rho}. \end{aligned} \tag{5}$$

Final allocations therefore are

$$\alpha_{B2,2}^* = \frac{\rho(\mu_B - \left(\mu_B - b\mu_y - s_B\frac{\left(\sigma_B^2 + b^2\sigma_y^2\right)}{\rho}\right) - b\mu_y)}{\left(\sigma_B^2 + b^2\sigma_y^2\right)} = s_B, \tag{6}$$

as required. Thus, all investors, independent of which fund they are in, hold identical stakes in asset $B2$, but the asset's price is reduced relative to asset $B1$ due to the anticipated externality.[7]

In the first period, for asset G both funds just maximize investor welfare identically. The representative investor's problem is

$$\underset{\alpha_G}{max} \left\{ \alpha_{G,1}(\mu_G + P_{G,2} - P_{G,1}) - \frac{1}{2\rho}\alpha_{G,1}^2\sigma_G^2 \right\}. \tag{7}$$

[7]This implies that the exclusion fund at $t = 2$ does not actually hold less of the dirty asset, but this is just an artifact of the simplifying assumption that there are no financial investors in the model to absorb excess supply.

The FOC is

$$\begin{aligned}(\mu_G + P_{G,2} - P_{G,1}) &= \frac{1}{\rho}\alpha_{G,1}\sigma_G^2 \\ \iff \alpha^*_{G,1} &= \frac{\rho(\mu_G + P_{G,2} - P_{G,1})}{\sigma_G^2}\end{aligned} \tag{8}$$

Market clearing yields

$$\begin{aligned}&1\alpha^*_{G,1} = s_G \\ \iff &\frac{\rho(\mu_G + P_{G,2} - P_{G,1})}{\sigma_G^2} = s_G \\ \iff &P_{G,1} = \mu_G + P_{G,2} - s_G\frac{\sigma_G^2}{\rho} = 2(\mu_G - s_G\frac{\sigma_G^2}{\rho}).\end{aligned} \tag{9}$$

Intuitively, the funds forecast that they will hold the same quantity of the green asset over both periods (i.e., there is no trade at $t = 2$), so they pay up front the full value of two periods worth of cash flows. Final allocations therefore are

$$\alpha^*_{G,1} = s_G. \tag{10}$$

For assets $B1$ and $B2$ the EF and IF will behave differently based on their different information. The IF, anticipating the clean-up and higher second-period price, will demand $\alpha_{B1,1}$ shares per investor of asset $B1$ according to

$$\underset{\alpha_{B1,1}}{max}\left\{\alpha_{B1,1}(\mu_B + P_{B1,2} - P_{B1,1} - b\mu_y) - \frac{1}{2\rho}\alpha^2_{B1,1}\left(\sigma_B^2 + b^2\sigma_y^2\right)\right\}. \tag{11}$$

so that the FOC becomes

$$\begin{aligned}(\mu_B + P_{B1,2} - P_{B1,1} - b\mu_y) &= \frac{1}{\rho}\alpha_{B1,1}\left(\sigma_B^2 + b^2\sigma_y^2\right) \\ \iff \alpha^*_{B1,1}(IF) &= \frac{\rho(\mu_B + P_{B1,2} - P_{B1,1} - b\mu_y)}{\left(\sigma_B^2 + b^2\sigma_y^2\right)}\end{aligned} \tag{12}$$

Let $P_{C,2}$ denote the second period price of a cleaned-up brown asset and $P_{D,2}$ the second period price of a still-dirty brown asset. The EF does not know whether $B1$ or $B2$ will be cleaned up and

therefore with respect to asset $B1$ maximizes:[8]

$$\underset{\alpha_{B1,1}}{max}\ \Big\{\alpha_{B1,1}(\mu_B + (\frac{1}{2}P_{C,2} + \frac{1}{2}P_{D,2}) - P_{B1,1} - b\mu_y) - \frac{1}{2\rho}\alpha_{B1,1}^2\big(\sigma_B^2 + b^2\sigma_y^2\big)\Big\}. \tag{13}$$

The FOC is

$$\begin{aligned}
&(\mu_B + (\frac{1}{2}P_{C,2} + \frac{1}{2}P_{D,2}) - P_{B1,1} - b\mu_y) = \frac{1}{\rho}\alpha_{B1,1}\big(\sigma_B^2 + b^2\sigma_y^2\big) \\
\iff &\alpha^*_{B1,1}(EF) = \frac{\rho(\mu_B + (\frac{1}{2}P_{C,2} + \frac{1}{2}P_{D,2}) - P_{B1,1} - b\mu_y)}{\big(\sigma_B^2 + b^2\sigma_y^2\big)}
\end{aligned} \tag{14}$$

Using the generic price notation, the IF investors demand, from above,

$$\alpha^*_{B1,1}(IF) = \frac{\rho(\mu_B + P_{C,2} - P_{B1,1} - b\mu_y)}{\big(\sigma_B^2 + b^2\sigma_y^2\big)}, \tag{15}$$

which is larger as a higher resale price is expected and uncertainty is lower. Market clearing yields

$$\begin{aligned}
&\eta\frac{\rho(\mu_B + P_{C,2} - P_{B1,1} - b\mu_y)}{\big(\sigma_B^2 + b^2\sigma_y^2\big)} + (1-\eta)\frac{\rho(\mu_B + (\frac{1}{2}P_{C,2} + \frac{1}{2}P_{D,2}) - P_{B1,1} - b\mu_y)}{\big(\sigma_B^2 + b^2\sigma_y^2\big)} = s_B \\
\iff &\frac{\rho(\mu_B - P_{B1,1} - b\mu_y)}{\big(\sigma_B^2 + b^2\sigma_y^2\big)} + \frac{\rho(\eta P_{C,2} + (1-\eta)(\frac{1}{2}P_{C,2} + \frac{1}{2}P_{D,2}))}{\big(\sigma_B^2 + b^2\sigma_y^2\big)} = s_B \\
\iff &P_{B1,1} = \mu_B + (P_{C,2} - (1-\eta)\frac{1}{2}(P_{C,2} - P_{D,2})) - b\mu_y - s_B\frac{\big(\sigma_B^2 + b^2\sigma_y^2\big)}{\rho}.
\end{aligned} \tag{16}$$

Replacing the prices $P_{C,2}$ and $P_{D,2}$ with the expressions derived above for $P_{B1,2}$ and $P_{B2,2}$, respectively, and solving for the demand from each fund, we end up with per-investor allocations of $\alpha^*_{B1,1}(EF) = s_B - \frac{\rho\frac{1}{2}\eta(b\mu_y + \frac{b^2\sigma_y^2}{\rho})}{\big(\sigma_B^2 + b^2\sigma_y^2\big)}$ for the EF and $\alpha^*_{B1,1}(IF) = s_B + \frac{\rho(1-\eta)\frac{1}{2}(b\mu_y + \frac{b^2\sigma_y^2}{\rho})}{\big(\sigma_B^2 + b^2\sigma_y^2\big)}$ for the IF.

Finally consider the equilibrium first period holdings in asset $B2$. Analogously to asset $B1$, the

[8] A discerning reader may wonder if additional volatility is introduced into the certainty equivalent expression due to the EF's uncertainty over the second period price of each brown asset. However, this is not the case because first period holdings are effectively non-stochastic *conditional* on an asset's second period status as cleaned-up or still-dirty (i.e., the fact that in the first period the EF will hold less of the cleaned-up asset and more of the still-dirty asset is entirely predictable based on the IF's randomization), and the prices $P_{C,2}$ and $P_{D,2}$ are also perfectly forecastable.

EF demands

$$\alpha^*_{B2,1}(EF) = \frac{\rho(\mu_B + (\frac{1}{2}P_{C,2} + \frac{1}{2}P_{D,2}) - P_{B2,1} - b\mu_y)}{\left(\sigma_B^2 + b^2\sigma_y^2\right)} \tag{17}$$

and the IF demands

$$\alpha^*_{B2,1}(IF) = \frac{\rho(\mu_B + P_{D,2} - P_{B2,1} - b\mu_y)}{\left(\sigma_B^2 + b^2\sigma_y^2\right)}. \tag{18}$$

Market clearing yields a price of

$$P_{B2,1} = \mu_B - b\mu_y + P_{D,2} + (1-\eta)\frac{1}{2}(P_{C,2} - P_{D,2}) - s_B\frac{\left(\sigma_B^2 + b^2\sigma_y^2\right)}{\rho}$$

and final equilibrium holdings are $\alpha^*_{B2,1}(EF) = s_B + \frac{\rho\frac{1}{2}\eta(b\mu_y + \frac{b^2\sigma_y^2}{\rho})}{\left(\sigma_B^2 + b^2\sigma_y^2\right)}$ for investors in the EF and $\alpha^*_{B2,1}(IF) = s_B - \frac{\rho(1-\eta)\frac{1}{2}(b\mu_y + \frac{b^2\sigma_y^2}{\rho})}{\left(\sigma_B^2 + b^2\sigma_y^2\right)}$ for investors in the IF.

Thus far, we have assumed that cleaning up asset $B1$ will be optimal for the IF after trading in period 1. We now confirm this via the following result.

Lemma 2. *The IF will optimally clean up asset $B1$ given the above trading equilibrium if $c \leq \eta s_B\,(b\mu_y) + \frac{\eta\rho(1-\eta)}{2\Sigma}B^2 + \eta^2\frac{1}{2\rho}s_B^2\left(b^2\sigma_y^2\right)$, where $B \equiv b\mu_y + \frac{b^2\sigma_y^2}{\rho}$ and $\Sigma \equiv \sigma_B^2 + b^2\sigma_y^2$.*

This result confirms that the conjectured strategy for the IF is subgame perfect as long as the cost of clean-up is not too high. From here on we assume that $c \leq \eta s_B\,(b\mu_y) + \frac{\eta\rho(1-\eta)}{2\Sigma}B^2 + \eta^2\frac{1}{2\rho}s_B^2\left(b^2\sigma_y^2\right)$.

Finally, we need to confirm that there exists a relative size of the IF, η, that makes investors indifferent between the two funds given the conjectured strategies. Note that both funds hold the same per-investor quantities of asset G in $t = 1$ and of all assets in $t = 2$. Thus, differences arise only from the holdings of the two brown assets in $t = 1$ as well as the clean-up costs. Indifference

thus requires

$$\begin{aligned}
&\alpha^*_{B1,1}(IF)(\mu_B + P_{C,2} - P_{B1,1} - b\mu_y) - \frac{1}{2\rho}\alpha^*_{B1,1}(IF)^2\big(\sigma_B^2 + b^2\sigma_y^2\big) - \frac{c}{\eta} \\
&+\alpha^*_{B2,1}(IF)(\mu_B + P_{D,2} - P_{B2,1} - b\mu_y) - \frac{1}{2\rho}\alpha^*_{B2,1}(IF)^2\big(\sigma_B^2 + b^2\sigma_y^2\big) = \\
&\alpha^*_{B1,1}(EF)(\mu_B + P_{C,2} - P_{B1,1} - b\mu_y) - \frac{1}{2\rho}\alpha^*_{B1,1}(EF)^2\big(\sigma_B^2 + b^2\sigma_y^2\big) \\
&+\alpha^*_{B2,1}(EF)(\mu_B + P_{D,2} - P_{B2,1} - b\mu_y) - \frac{1}{2\rho}\alpha^*_{B2,1}(EF)^2\big(\sigma_B^2 + b^2\sigma_y^2\big) \\
\iff &[\alpha^*_{B1,1}(IF) - \alpha^*_{B1,1}(EF)](\mu_B + P_{C,2} - P_{B1,1} - b\mu_y) - \frac{1}{2\rho}[\alpha^*_{B1,1}(IF)^2 - \alpha^*_{B1,1}(EF)^2]\big(\sigma_B^2 + b^2\sigma_y^2\big) - \frac{c}{\eta} = \\
&[\alpha^*_{B2,1}(EF) - \alpha^*_{B2,1}(IF)](\mu_B + P_{D,2} - P_{B2,1} - b\mu_y) - \frac{1}{2\rho}[\alpha^*_{B2,1}(EF)^2 - \alpha^*_{B2,1}(IF)^2]\big(\sigma_B^2 + b^2\sigma_y^2\big)
\end{aligned} \tag{19}$$

Analyzing this expression using the equilibrium holdings derived above, and B and Σ as defined in the proof of Lemma 2, yields the following result.

Proposition 3. *There exist two equilibria in which both an impact fund and an exclusion fund form, the impact fund randomizes over which brown asset to clean up, and the relative size of the impact fund, η, solves*

$$\eta(1-\eta)\frac{\rho}{2\Sigma}B^2 = c.$$

The two valid solutions for η correspond to a small IF with a large EF, and a large IF with a small EF. To understand intuitively why there are two possible solutions, note that when the IF is small, there will be a large difference in the first and second period price of asset $B1$ (little money is invested in the first period by those with access to the private information that $B1$ will be cleaned up). With a small IF, there are fewer investors to share in the cost of cleanup, so they must be compensated for this cost with high profits via the large price change and sale of the cleaned up asset. On the other hand, when the IF is large, the price difference across the two periods is smaller as more first period investors are "in the know." In this case, IF investors have to spread the cost of cleanup more widely since they get lower profits from selling the cleaned up asset at time 2. Both solutions are valid, so the model as it stands does not provide guidance on the relative sizes of the IF and EF sectors.

The model does, however, have implications, both positive and normative. Let us turn first to the positive implications. The size of the price effect of exclusion directly affects the *efficiency* of impact capital: when exclusion generates only a small discount, it will take a large amount of impact capital to make a difference. This is because the per-investor willingness to pay for engagement within an impact fund depends on the per-investor trading profit gained from leveraging the presence of exclusion funds. When exclusion funds do not generate large profit opportunities, the per-impact-investor willingness to pay is small, and so many such investors, i.e., large impact funds, will be necessary to undertake even relatively small engagements.

Second, turning to normative implications, the model suggests that, despite the fact that in impact capital is inefficient in the large-η equilibrium, it may be socially optimal to have a small exclusionary fund sector and a large impact fund sector. This is because, when both $t = 1$ dirty assets are taken into account, the largest amount of clean-up is achieved in the largeη equilibrium. In both equilibria, the impact fund (since it exists) achieves the same degree of clean-up: it cleans up asset $B1$. However, our model implies that the price of the other dirty asset, $B2$, is lower in the large-η equilibrium: Because the EF doesn't know which firm will be cleaned up, it demands the same amount of both brown firms. However, because the IF knows $B1$ will be cleaned up, it demands relatively more $B1$ and less $B2$. As a result, as η (IF's size) increases, the price of $B1$ increases and the price of $B2$ decreases. Thus, though we do not take a stand on the exact nature of the $\hat{\mu}$ exclusion function, as long as exclusion is monotonically increasing in the size of the price effect of exclusion, greater overall externality reduction would be achieved in the large-η equilibrium.

Appendix

Proof of Proposition 1: Follows from the discussion in the text. ■

Proof of Proposition 2: The result for consequentalist preferences follows from the discussion in the text. For cold prickle preferences, first propose an exclusionary fund of any size that has an investment mandate different from the optimal EF. Then any proposed investor in the fund will prefer to deviate to trade optimally on their own given the existence of that fund (since the proposed fund does not optimize their risk sharing given its own existence). Now propose the optimal EF at its full size. Then an investor cannot do better by trading on their own given the existence of the optimal EF. Thus, an optimal EF that attracts all financial investors exists. ■

Proof of Lemma 1: Using backward induction, note that at $t = 2$, the maximization problem for either fund type is given by

$$\begin{aligned}
&\max_{\alpha_2} \mathbb{E}_2\big[-\exp\big(-\frac{\sum_{i\in\{G,B1,B2\}}\alpha_{i,1}(\tilde{z}_{i,2}+P_{i,2}-P_{i,1}-1_{i\neq G}b\tilde{y}_{i,1})+\alpha_{i,2}(\tilde{z}_{i,2}-P_{i,2}-1_{i\neq G}b\tilde{y}_{i,2})}{\rho}\big)\big]\\
=&\max_{\alpha_2} \mathbb{E}_2\big[-\exp\Big(-\frac{\sum_{i\in\{G,B1,B2\}}\alpha_{i,1}(\tilde{z}_{i,1}+P_{i,2}-P_{i,1}-1_{i\neq G}b\tilde{y}_{i1})}{\rho}\Big)\\
&\times\exp\Big(-\frac{\sum_{i\in\{G,B1,B2\}}\alpha_{i,2}(\tilde{z}_{i,2}-P_{i,2}-1_{i\neq G}b\tilde{y}_{i,2})}{\rho}\Big)\big] \qquad (20)
\end{aligned}$$

Standing at $t = 2$, $\sum_{i\in\{G,B1,B2\}}\alpha_{i,1}(\tilde{z}_{i,1}+P_{i,2}-P_{i,1}-1_{i\neq G}b\tilde{y}_{i1})$ is effectively a constant, so it does not change the maximization problem. Thus, the $t = 2$ maximization problem simplifies to

$$\max_{\alpha_2} \mathbb{E}_2\big[-\exp\Big(-\frac{\sum_{i\in\{G,B1,B2\}}\alpha_{i,2}(\tilde{z}_{i,2}-P_{i,2}-1_{i\neq G}b\tilde{y}_{i,2})}{\rho}\Big)\big] \qquad (21)$$

Backing up, at $t = 1$, the initial the maximization problem is

$$\max_{\alpha_1} \mathbb{E}_1\big[\max_{\alpha_2} \mathbb{E}_2\big[-\exp\big(-\frac{W}{\rho}\big)\big]\big]. \qquad (22)$$

The first period wealth of $\sum_{i\in\{G,B1,B2\}}\alpha_{i,1}(\tilde{z}_{i,1}+P_{i,2}-P_{i,1}-1_{i\neq G}b\tilde{y}_{i1})$ does not affect the second period choice as argued above. This implies that the choice at $t = 1$ is unaffected by $t = 2$ except with respect to the expectation of the second-period price, so that the investor's maximization

problem boils down to

$$\underset{\alpha_1}{max}\mathbb{E}_1\big[-\exp\bigg(-\frac{\sum_{i\in\{G,B1,B2\}}\alpha_{i,1}(\tilde{z}_{i,1}+P_{i,2}-P_{i,1}-1_{i\neq G}b\tilde{y}_{i1})}{\rho}\bigg)\big].\blacksquare \tag{23}$$

Proof of Lemma 2: Conditional on having a per-investor stake $\alpha^*_{B1,1}$ in asset $B1$, paying to clean up the asset is optimal if:

$$\begin{aligned}
&\eta\alpha^*_{B1,1}(\mu_B+P_{C,2}-P_1-b\mu_y)-\frac{1}{2\rho}\eta^2\alpha^{*2}_{B1,1}\big(\sigma_B^2+b^2\sigma_y^2\big)-c+\eta\left(s_B(\mu_B-P_{C,2})-\frac{1}{2\rho}\eta s_B^2\big(\sigma_B^2\big)\right)\geq\\
&\eta\alpha^*_{B1,1}(\mu_B+P_{D,2}-P_1-b\mu_y)-\frac{1}{2\rho}\eta^2\alpha^{*2}_{B1,1}\big(\sigma_B^2+b^2\sigma_y^2\big)+\eta\left(s_B(\mu_B-P_{D,2}-b\mu_y)-\frac{1}{2\rho}\eta s_B^2\big(\sigma_B^2+b^2\sigma_y^2\big)\right)\\
\iff&\eta\alpha^*_{B1,1}(P_{C,2}-P_{D,2})+\\
&\eta\left(\left(s_B(\mu_B-P_{C,2})-\frac{1}{2\rho}\eta s_B^2\big(\sigma_B^2\big)\right)-\left(s_B(\mu_B-P_{D,2}-b\mu_y)-\frac{1}{2\rho}\eta s_B^2\big(\sigma_B^2+b^2\sigma_y^2\big)\right)\right)\geq c\\
\iff&\eta\Big(s_B+\frac{\rho(1-\eta)\frac{1}{2}(b\mu_y+\frac{b^2\sigma_y^2}{\rho})}{\big(\sigma_B^2+b^2\sigma_y^2\big)}\Big)\big(b\mu_y+\frac{b^2\sigma_y^2}{\rho}\big)+\eta\left(-s_B\frac{b^2\sigma_y^2}{\rho}+\frac{1}{2\rho}\eta s_B^2\big(b^2\sigma_y^2\big)\right)\ \geq\ c\\
\iff&\eta s_B\left(b\mu_y\right)+\eta\frac{\rho(1-\eta)\frac{1}{2}(b\mu_y+\frac{b^2\sigma_y^2}{\rho})^{\ 2}}{\big(\sigma_B^2+b^2\sigma_y^2\big)}+\eta^2\frac{1}{2\rho}s_B^2\big(b^2\sigma_y^2\big)\ \ \geq\ c.
\end{aligned}\tag{24}$$

In the calculations above, the expressions

$$\eta\left(s_B(\mu_B-P_{C,2})-\frac{1}{2\rho}\eta s_B^2\big(\sigma_B^2\big)\right)$$

and

$$\eta\left(s_B(\mu_B-P_{D,2}-b\mu_y)-\frac{1}{2\rho}\eta s_B^2\big(\sigma_B^2+b^2\sigma_y^2\big)\right)$$

represent the aggregate second period CE for sticking to the clean-up strategy and deviating from it respectively. Clearly, the LHS is strictly positive. Defining B and Σ as in the proposition, which are both independent of η, gives a final condition of

$$\eta s_B\left(b\mu_y\right)+\frac{\eta\rho(1-\eta)}{2\Sigma}B^2+\eta^2\frac{1}{2\rho}s_B^2\big(b^2\sigma_y^2\big)\ \geq\ c.\blacksquare \tag{25}$$

Proof of Proposition 3: Note that

$$\begin{aligned}[\alpha^*_{B1,1}(IF) - \alpha^*_{B1,1}(EF)] =& s_B + \frac{\rho(1-\eta)\frac{1}{2}(b\mu_y + \frac{b^2\sigma_y^2}{\rho})}{\left(\sigma_B^2 + b^2\sigma_y^2\right)} - s_B + \frac{\rho\frac{1}{2}\eta(b\mu_y + \frac{b^2\sigma_y^2}{\rho})}{\left(\sigma_B^2 + b^2\sigma_y^2\right)} \\ =& \frac{\rho\frac{1}{2}(b\mu_y + \frac{b^2\sigma_y^2}{\rho})}{\left(\sigma_B^2 + b^2\sigma_y^2\right)} \\ =& \frac{\rho\frac{1}{2}B}{\Sigma}, \end{aligned} \tag{26}$$

$$\begin{aligned}[\alpha^*_{B1,1}(IF)^2 - \alpha^*_{B1,1}(EF)^2] =& \Big(s_B + \frac{\rho(1-\eta)\frac{1}{2}(b\mu_y + \frac{b^2\sigma_y^2}{\rho})}{\left(\sigma_B^2 + b^2\sigma_y^2\right)}\Big)^2 - \Big(s_B - \frac{\rho\frac{1}{2}\eta(b\mu_y + \frac{b^2\sigma_y^2}{\rho})}{\left(\sigma_B^2 + b^2\sigma_y^2\right)}\Big)^2 \\ =& \frac{\rho\, s_B}{\sigma_B^2 + b^2\sigma_y^2}\left(b\mu_y + \frac{b^2\sigma_y^2}{\rho}\right) + \frac{\rho^2}{4\left(\sigma_B^2 + b^2\sigma_y^2\right)^2}(1-2\eta)\left(b\mu_y + \frac{b^2\sigma_y^2}{\rho}\right)^2 \\ =& \frac{\rho\, s_B}{\Sigma}B + \frac{\rho^2}{4\Sigma^2}(1-2\eta)B^2, \end{aligned} \tag{27}$$

$$\begin{aligned}[\alpha^*_{B2,1}(EF) - \alpha^*_{B2,1}(IF)] =& s_B + \frac{\rho\frac{1}{2}\eta(b\mu_y + \frac{b^2\sigma_y^2}{\rho})}{\left(\sigma_B^2 + b^2\sigma_y^2\right)} - s_{B2} + \frac{\rho(1-\eta)\frac{1}{2}(b\mu_y + \frac{b^2\sigma_y^2}{\rho})}{\left(\sigma_B^2 + b^2\sigma_y^2\right)} \\ =& \frac{\rho\frac{1}{2}(b\mu_y + \frac{b^2\sigma_y^2}{\rho})}{\left(\sigma_B^2 + b^2\sigma_y^2\right)} \\ =& \frac{\rho\frac{1}{2}B}{\Sigma}, \end{aligned} \tag{28}$$

$$\begin{aligned}[\alpha^*_{B2,1}(EF)^2 - \alpha^*_{B2,1}(IF)^2] =& \Big(s_B + \frac{\rho\frac{1}{2}\eta(b\mu_y + \frac{b^2\sigma_y^2}{\rho})}{\left(\sigma_B^2 + b^2\sigma_y^2\right)}\Big)^2 - \Big(s_B - \frac{\rho(1-\eta)\frac{1}{2}(b\mu_y + \frac{b^2\sigma_y^2}{\rho})}{\left(\sigma_B^2 + b^2\sigma_y^2\right)}\Big)^2 \\ =& \frac{\rho\, s_B}{\sigma_B^2 + b^2\sigma_y^2}\left(b\mu_y + \frac{b^2\sigma_y^2}{\rho}\right) + \frac{\rho^2}{4\left(\sigma_B^2 + b^2\sigma_y^2\right)^2}(2\eta-1)\left(b\mu_y + \frac{b^2\sigma_y^2}{\rho}\right)^2 \\ =& \frac{\rho\, s_B}{\Sigma}B + \frac{\rho^2}{4\Sigma^2}(2\eta-1)B^2, \end{aligned} \tag{29}$$

$$
\begin{aligned}
&(\mu_B + P_{C,2} - P_{B1,1} - b\mu_y) \\
=&(\mu_B + (P_{C,2} - (1-\eta)\frac{1}{2}(P_{C,2} - P_{D,2})) - b\mu_y - s_B\frac{\left(\sigma_B^2 + b^2\sigma_y^2\right)}{\rho}) - b\mu_y) \\
=&((1-\eta)\frac{1}{2}(P_{C,2} - P_{D,2}) + s_B\frac{\left(\sigma_B^2 + b^2\sigma_y^2\right)}{\rho}) \\
=&((1-\eta)\frac{1}{2}(b\mu_y + \frac{b^2\sigma_y^2}{\rho}) + s_B\frac{\left(\sigma_B^2 + b^2\sigma_y^2\right)}{\rho}) \\
=&((1-\eta)\frac{1}{2}B + s_B\frac{\Sigma}{\rho}),
\end{aligned}
\tag{30}
$$

and

$$
\begin{aligned}
&(\mu_B + P_{D,2} - P_{B2,1} - b\mu_y) \\
=&(\mu_B + P_{D,2} - (\mu_B - b\mu_y + P_{D,2} + (1-\eta)\frac{1}{2}(P_{C,2} - P_{D,2}) - s_B\frac{\left(\sigma_B^2 + b^2\sigma_y^2\right)}{\rho}) - b\mu_y) \\
=&(P_{D,2} - (P_{D,2} + (1-\eta)\frac{1}{2}(P_{C,2} - P_{D,2}) - s_B\frac{\left(\sigma_B^2 + b^2\sigma_y^2\right)}{\rho})) \\
=&s_B\frac{\left(\sigma_B^2 + b^2\sigma_y^2\right)}{\rho} - (1-\eta)\frac{1}{2}(P_{C,2} - P_{D,2}) \\
=&s_B\frac{\left(\sigma_B^2 + b^2\sigma_y^2\right)}{\rho} - (1-\eta)\frac{1}{2}(b\mu_y + \frac{b^2\sigma_y^2}{\rho}) \\
=&s_B\frac{\Sigma}{\rho} - (1-\eta)\frac{1}{2}B.
\end{aligned}
\tag{31}
$$

Plugging everything into the indifference equation from the text yields:

$$
\begin{aligned}
&[\alpha^*_{B1,1}(IF) - \alpha^*_{B1,1}(EF)](\mu_B + P_{C,2} - P_{B1,1} - b\mu_y) - \frac{1}{2\rho}[\alpha^*_{B1,1}(IF)^2 - \alpha^*_{B1,1}(EF)^2]\big(\sigma_B^2 + b^2\sigma_y^2\big) - \frac{c}{\eta} = \\
&[\alpha^*_{B2,1}(EF) - \alpha^*_{B2,1}(IF)](\mu_B + P_{D,2} - P_{B2,1} - b\mu_y) - \frac{1}{2\rho}[\alpha^*_{B2,1}(EF)^2 - \alpha^*_{B2,1}(IF)^2]\big(\sigma_B^2 + b^2\sigma_y^2\big) \\
\iff &\frac{\rho\frac{1}{2}B}{\Sigma}((1-\eta)\frac{1}{2}B + s_B\frac{\Sigma}{\rho}) - \frac{1}{2\rho}[\frac{\rho\, s_B}{\Sigma}B + \frac{\rho^2}{4\Sigma^2}(1-2\eta)B^2]\Sigma - \frac{c}{\eta} = \\
&[\frac{\rho\frac{1}{2}B}{\Sigma}](s_B\frac{\Sigma}{\rho} - (1-\eta)\frac{1}{2}B) - \frac{1}{2\rho}[\frac{\rho\, s_B}{\Sigma}B + \frac{\rho^2}{4\Sigma^2}(2\eta-1)B^2]\Sigma \\
\iff &\frac{\rho\frac{1}{2}B}{\Sigma}((1-\eta)\frac{1}{2}B + s_B\frac{\Sigma}{\rho}) - \frac{1}{2\rho}[\frac{\rho\, s_B}{1}B + \frac{\rho^2}{4\Sigma}(1-2\eta)B^2] - \frac{c}{\eta} = \\
&[\frac{\rho\frac{1}{2}B}{\Sigma}](s_B\frac{\Sigma}{\rho} - (1-\eta)\frac{1}{2}B) - \frac{1}{2\rho}[\frac{\rho\, s_B}{1}B + \frac{\rho^2}{4\Sigma}(2\eta-1)B^2] \\
\iff &\eta(1-\eta)\frac{\rho\frac{1}{2}B^2}{\Sigma} = c \qquad (32)
\end{aligned}
$$

as given in the result. This clearly has two possible solutions in η. ■

Bibliography

Abis, S., A. Buffa, and M. Sadasivam (2024). Different shades of ESG funds. Working paper, SSRN.

Andrikogiannopoulou, A., P. Krueger, S.F. Mitali, and F. Papakonstantinou (2024). Empty ESG Promises: A Text Analysis of Mutual Fund Prospectuses. Working Paper. King's College London.

BCG (2025), Global Asset Management Report: From Recovery to Reinvention, https://on.bcg.com/3GujISl, April 2025.

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

Berk, J. B. and J. H. Van Binsbergen (2025). The impact of impact investing. Journal of Financial Economics 164, 103972.

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

Cole, S., L. Jeng, J. Lerner, N. Rigol, and B.N. Roth (2023). What do impact investors do differently? Working Paper.

Dangl, T., M. Halling, J. Yu, and J. Zechner (2025). Social Preferences and Corporate Investment. Journal of Financial Economics 172, 104-139.

Dasgupta, A., S. Fos, and Z. Sautner (2021). Institutional Investors and Corporate Governance. Foundations and Trends in Finance 12: 276-394.

Dasgupta, A. and R. Mathews (2025). Delegated Blocks. Working Paper, LSE and University of Maryland.

Frankel, R.M., X. Martin, H.I. Wang, and M.A. Yu (2024). Mutual Fund ESG Disclosures. Working Paper. Washington University.

Gollier, C., and S. Pouget (2022). Investment Strategies and Corporate Behaviour with SociallyResponsible Investors: A Theory of Active Ownership. Economica 89: 997-1023.

Green, D., and B. Roth (2025). The Allocation of Socially Responsible Capital. Journal of Finance 80: 755-782.

Grossman, S., and O. D. Hart, 1980, Takeover bids, the free-rider problem, and the theory of the corporation, Bell Journal of Economics 11, 42–64.

Gryglewicz, S., S. Mayer, and E. Morellec (2025). A Unifying Theory of Shareholder Activism, Ownership Dynamics, and Firm Policies. Working Paper, SSRN.

Guo, Y. (2025). From Preference to Influence: The Economics of Activist Investing. Working Paper, LSE.

Gupta, D., Kopytov, A., and J. Starmans (2025). The Pace of Change: Socially Responsible Investing in Private Markets. Review of Financial Studies 39, 30-78.

Hakenes, H., and E. Schliephake (2024). Responsible Investment and Responsible Consumption. Working Paper, SSRN.

Heath, D., D. Macciocchi, R. Michaely, and M. C. Ringgenberg (2023). Does socially responsible investing change firm behavior? The Review of Finance 27, 2057–2083.

Heeb, F., J. Koelbel, F. Paetzold, and S. Zeisberger (2023). Do investors care about impact? The Review of Financial Studies 36, 1737–1787.

IEA (2024), The Role of Carbon Credits in Scaling Up Innovative Clean Energy Technologies, IEA, Paris https://www.iea.org/reports/the-role-of-carbon-credits-in-scaling-up-innovative-clean-energy-technologies, Licence: CC BY 4.0.

Kyle, A. and J. Vila (1991). Noise Trading and Takeovers. RAND Journal of Economics 22: 54-71. Malenko, A., N. Malenko, and A. Tsoy (2025). Fund Proliferation, Decentralization, and Shareholder Power. Working Paper, Boston College.

Pastor, L., R. Stambaugh, and L. Taylor (2024). Sustainable Investing. Annual Review of Financial Economics 17: 321-341.

Pedersen, L. H. (2025). Can Sustainable Finance Save the Planet? Journal of Finance: Insights and Perspectives, forthcoming.

Riedl, A., and P. Smeets (2017). Why do investors hold socially responsible mutual funds? Journal of Finance 72, 2505–2549.

Sautner, Z. (2025). Institutional Investor Engagement: From Climate to Nature Risks. Journal of Finance: Insights and Perspectives, forthcoming.

Yang, K., and A. Yasuda (2025). Decoding Sustainable Investment Strategies: Bridging Intentions and Outcomes. Working Paper. UC Davis.