Socially Optimal Eligibility Criteria for ESG Funds

Roman Inderst* Marcus Opp†

August 2022

Abstract

Agencies around the world are in the process of developing taxonomies for sustainable (or ESG) investment products. When a social planner takes into account only real effects, as households merely derive (decisional) "warm-glow" utility, the introduction of a sustainable investment category is only beneficial when (environmental) regulation is weak, e.g., due to policy failure, or when firms are sufficiently financially constrained. Otherwise, the resulting indirect subsidy for higher sustainability is inefficient. In case of financial constraints, an ESG classification relaxes a trade-off faced by a social planner who wants to increase sustainability for given output (internal margin) while achieving an efficient size of the economy (external margin). Then, the optimal classification maximizes the resulting subsidy in view of a downward sloping supply of sustainable capital. Its stringency depends on the scarcity of sustainable capital and whether social norms lead to a feedback effect in its supply, on the severity of financial constraints, and on the availability of more direct instruments such as environmental taxes.

Keywords: Sustainability; ESG; green financing.

*Goethe Universität Frankfurt. inderst@finance.uni-frankfurt.de

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

1 Introduction

Initiatives to harness financing to create a more sustainable economy have gained much momentum over the last years. This is particularly visible in the progress that the European Union and its various institutions made in 2021 to draft the technical standards that will allow to classify investment vehicles according to their sustainability contributions, comprising both environmental and social characteristics.[1] While such a taxonomy is likely to be used also for various regulatory purposes, the declared key purpose is to channel funding also from notably retail investors to more sustainable firms and projects. In this paper, we thus start from households' preferences for sustainable investment.[2] Though the evidence on this is still mixed (cf. below), we assume that these preferences make them accept a lower return. Households' preferences are assumed to be non-consequentialist, on which we comment in detail below: For each Euro invested, a household derives a (decisional) warm-glow utility, which does not enter the social planner's objective. Focussing thus only on real effects, there are no immediate benefits from the introduction of a sustainable investment classification. Depending also on the availability of more direct policy instruments, such as a minimum standard for all production or a tax on externalities, we analyze when the introduction of a sustainable investment classification is beneficial and if so, what factors then determine the optimal stringency of the respective threshold. Our analysis is set in a corporate finance model where agency problems generate frictions in outside financing, for which we employ the workhorse model of Holmström and Tirole (1997). For our welfare analysis, our model is closed on the product market side.

We derive the following key messages from our analysis. The introduction of a sustainable investment category is beneficial only in the following two cases, while otherwise it reduces welfare. First, it is beneficial when it is not feasible to set a sufficiently high minimum standard with which all firms must comply or a sufficiently high tax on externalities. Recently, the argument that notably environmental standards may be too low due to political failure has gained traction (cf. Tirole 2012), which in Europe has even lead some courts to impose obligations on governments.[3] While limited sustainable capital

[1]See the Final Report on Draft Regulatory Standards, issued in October 2021 jointly by various institutions, including ESMA (the European Securities and Market Authority): https://www.esma.europa.eu/sites/default/files/library/jc_2021_50_-_final_report_on_taxonomy-related_product_disclosure_rts.pdf.

[2]In line with the language chosen for the European taxonomy project, we use the term "sustainable investment", rather than the term "ESG investment".

[3]Notably the case of the Netherlands has been widely covered, where in 2019 its supreme court upheld a ruling ordering the government to step up actions to cut carbon emissions. Also the German surpreme

does not affect the external margin of the considered economic activity, it positively affects the internal margin by providing a subsidy that incentivizes a higher standard. Given a downward sloping supply of sustainable capital, a social planner faces a trade-off. The classification is optimally less stringent when such capital is (still) scarce or when there is a feedback-effect as households' preferences depend on endogenous social norms.

The second case in which the introduction of a sustainable investment category is beneficial, now despite the optimal choice of other, more direct policy instruments, is when the economy is sufficiently financially constrained. Then, the resulting subsidy for more sustainable investments mitigates a trade-off between imposing a higher minimum standard on the internal margin and, at the same time, inefficiently shrinking the economy on the external margin. When a sustainable investment classification is introduced, the respective minimum standard should optimally be increased (or a tax on externalities should likewise be increased).

The introduction of a sustainable investment classification thus provides an indirect way to subsidize investment that is more sustainable than required by the prevailing minimum standard or that exceeds the level that a prevailing tax on externalities would induce. Such a subsidy is strictly suboptimal if in the absence of financial constraints a more direct instrument, such as a minimum standard, can be optimally chosen. When however financial constraints generate a trade-off between the two goals of raising sustainability at the internal margin and achieving an efficient size of the respective economic activity (external margin), the direct instrument alone is insufficient to achieve the social optimum and the indirect subsidy through a sustainable investment classification becomes strictly beneficial. The optimal choice of the classification threshold maximizes the resulting subsidy in view of a downward sloping supply of sustainable capital.

We now defend our assumption on households' warm-glow preferences. Survey evidence seems to confirm that at least some investors take also an ethical perspective (e.g., Riedl and Smeets 2017). Baker et al. (2018) and Zerbib (2019) find that, ceteris paribus, green bonds are issued at a premium. Hartzmark and Sussman (2019) relate capital flows into mutual funds to their sustainability ratings.[4] The theoretical literature has mainly focussed on consequentialist (institutional) investors, as in Chowdhry et al. (2014) and Oehmke and

court has obliged the government to take actions so as to prevent irreversible damages suffered by future generations.

[4] We acknowledge that there is also evidence of positive abnormal returns, e.g., Gibson and Krueger (2018) and Henke (2016) for the relation between a portfolio's sustainability and return and Andersson et al. (2016) for the outperformance of green stock indices (albeit these results have also been attributed to market underreaction).

Opp (2020).[5] In our model, households have non-consequentialist preferences, deriving a (decisional) warm-glow utility from every Euro invested in a, using the language of the European disclosure regulation, "taxonomy-conform" sustainability (ESG) fund. Such preferences provide us with additional justification for why the social planner in ou rmodel is interested only in real effects. The specification of such preferences receives some support by a large literature in environmental and resource economics that elicits preferences for so-called "non-use value" related, in our case, to sustainability.[6] For instance, as pointed out in this literature, elicited willingness-to-pay frequently fails a so-called "scope test" or "adding-up test": Subjects' willingness-to-pay is relatively insensitive to the actual impact of the respective scenario change, e.g., the number of animals saved. Individuals may exhibit a high willingness-to-pay even when the environmental impact is miniscule and their willingness-to-pay does not increase sufficiently when the impact, i.e., the "scope", is much larger.[7] Put differently, the warm-glow effect that is experienced from a purchase of "moral satisfaction" (Kahneman et al. 1992) remains thus relatively unaffected by the actually achieved outcome. We thus treat it as a mere decisional utility, which does not enter the social planner's objective. In an extension we also analyze the case where such preferences depend on an endogenous social norm.

Ours is not the first analysis of the implications of sustainable investment preferences, albeit the literature has largely focussed on consequentialist preferences. In Heinkel et al. (2001) a boycott of some investors leads to more concentrated shareholding and thereby to a higher risk premium for less sustainable firms, and in Landier and Lovo (2020) lower sustainability increases capital costs by making it harder to locate an investor when the capital market has frictions. According to the model in Oehmke and Opp (2020), firms can negotiate lower capital costs with an institutional investor in exchange for implementing a less polluting technology. None of these contributions has addressed the two key questions of our analysis, i.e., whether such subsidized financing is socially beneficial in view of alternative, more direct policy instruments and how to optimally design a sustainable investment classification. The latter issue taps into a large theoretical literature on optimal certification (see Bizzotto and Harstad 2021 for a recent contribution and a detailed survey). This literature, however, typically considers the perspective of a profit-maximizing

[5] Oehmke and Opp (2020) also apply the Holmström and Tirole (1997) model of financial constraints, which we share also with Heider and Inderst (2020).

[6] We note that the assumption that households care only about the respective classification could also be motivated by lack of information or excessive complexity associated with a more detailed assessment.

[7] In the area of contingent valuation analysis, prominent studies that document these effects are Boyle et al. (1994) and more recently Desvousges et al. (2012).

certifier and the design of labels for products or services when purchasers have limited information about quality.

The rest of this paper is organized as follows. In Section 2 we introduce the modelling framework. Section 3 analyzes the market equilibrium, while Section 4 solves for the optimal policy. Section 5 introduces social preferences, and Section 6 concludes. All proofs are collected in an appendix.

2 Model Framework

We consider a single-sector economy and analyze the technology choice and investment decisions of a continuum of firms that produce a homogeneous good. Firms (owner-managers) are endowed with internal funds that are essential to mitigate a moral-hazard problem vis-à-vis outside investors. Outside investors are households that have heterogeneous sustainability preferences. These relate to what we generally refer to as the sustainability dimension of firms' production, e.g., to what extent the firm positively contributes to mitigating global warming or social development. Investors' sustainability preferences are "warm-glow" in the sense that they depend on whether a firm has been labelled sustainable or not. One should think about these preferences in terms of the following intermediation arrangement. If a firm obtains the "sustainability" classification, it is eligible for investment by an ESG fund. Upon investing in an ESG fund, households derive a warm glow utility (relative to a "normal" investment fund). The design of the eligibility criterion for the ESG fund is our main policy variable.

We next describe the model ingredients and then characterize the equilibrium, where households optimally allocate their capital and firms optimally decide which technology to use and how much outside investment to raise. This also determines the equilibrium price of capital raised from investors.

Firms. We consider the following one-good economy. There is a mass one of firms, each endowed with initial funds of size A (which is, thus, also the aggregate size of internal funds). We index firms by i. Their investment technology is described below. Individual investment size is denoted by I_i so that aggregate investment is $I = \int_0^1 I_i di$. Firms can use internal funds $A_i \leq A$ and raise external (loan) funding L_i for such investment, so that $I_i = A_i + L_i$. The latter requires the promise of some repayment D_i, which we will later express in terms of the required rate of return r_i. Whether this promise can be honored depends on the firm's success, which is subject to an agency problem adopted from the workhorse model of Holmström and Tirole (1997). Specifically, investment is only successful with probability one if the owner-manager exerts unobservable effort. If she shirks, she obtains a private benefit $B > 0$ per unit of investment, but with probability $\Delta p > 0$ the technology fails and returns no output. The produced output quantity is denoted by q_i. For simplicity, we assume constant-returns-to-scale production with $q_i = I_i$ in case of success and $q_i = 0$ in case of failure. The product market is perfectly competitive and, to focus on the role of the provision of sustainable capital, we suppose that more sustainable firms cannot secure a premium on the output market.[8] Then, given aggregate firm output $q = \int_0^1 q_i di$, firms reap a market price of $P(q)$ where $P' < 0$ and $P(q) = 0$ for q above a finite threshold. For ease of exposition, we omit a positive cost of production (though production will not be costless if firms produce more sustainably, as described next). As is standard in the literature, we restrict attention to parameters such that shirking is off-equilibrium.[9]

Social costs and environmental regulation. Production or consumption generates per unit social costs of ϕ leading to aggregate social costs of $q \cdot \phi$.[10] In equilibrium, owner-managers are incentivized not to shirk, so that social costs can be equally written as $I \cdot \phi$. One example is that of emissions, in which case the parameter ϕ can be interpreted as the emission intensity. Likewise, production can involve health hazards for workers, or it may inflict harm on animals. Alternatively, social costs can be related to consumption of the respective output. In the absence of any investment in sustainability, each firm generates per-unit social costs of $\phi = \bar{\phi}$, which can be reduced to $\phi = \bar{\phi} - \theta$ at cost $c(\theta)$, where c is a strictly increasing and convex function with $c(0) = c'(0) = 0$ and $\lim_{\theta \to \bar{\phi}} c'(\theta) = \infty$.

One objective of our analysis is to investigate the relevance of sustainable finance and ESG classification in the light of alternative policy instruments. We first consider a maximum emission standard $\phi_M \leq \bar{\phi}$. In light of our subsequent discussion of an ESG classification, it proves however convenient to perceive of this as a minimum investment or minimum sustainability standard θ_m (so that $\phi_M = \bar{\phi} - \theta_m$). For instance, this may require mandatory air filters of a certain quality. In an extension we also consider a possible tax of τ per unit of social cost and show how our results extend when this instrument is used

[8] See Hakenes and Schliephake (2021) for a model that considers both socially responsible consumption and socially responsible investment.

[9] See the proof of Proposition 1 for the exact condition.

[10] If externalities relate to production, with this specification they are realized only in case of success, rather than when the respective investment is sunk. Our results would be unaffected under the alternative speficiation where, instead, aggregate social costs are given by $I \cdot \phi$.

either alone or in combination with a minimum sustainability standard.[11]

Sustainable finance and ESG classification. Funds for investment are provided by atomistic households with aggregate capital K. Households are risk-neutral and identical in the valuation of cash flows. In addition, as motivated by the recent literature on ESG funds, we suppose that some households derive a payoff from owning an ESG fund, which can only consist of firms that exceed a more stringent sustainability standard $\theta_s > \theta_m$. The choice of θ_s will represent our main policy variable of interest. We will refer to firms that satisfy the standard θ_s as sustainable (and otherwise as non-sustainable).

In the Introduction, we already referred to the European Union's Sustainable Finance Disclosure Regulation (SFDR), for which the European Commission adopted on 6 April 2022 the necessary technical standards relating presently primarily the disclosure of environmental objectives. Investment products have to disclose, both on a pre-contractual basis and through periodic statements, how and to what extent the product's investments are aligned with the EU Taxonomy. An example for the likely high-level disclosure to consumers is contained in the consumer-testing documents, where it is indicated whether (or not) the product "has sustainable investments at its objective - in activities aligned with the EU taxonomy".[12] Even where the information provided to retail investors is less binary, e.g., indicating a percentage of such taxonomy-conform investments in a fund, there is still an underlying decision on the respective threshold that must be met so that a particular activity is deemed to be taxonomy-conform.

We stipulate that if household j makes an investment in such an ESG product, this generates a warm-glow effect w_j per unit of investment, which, thus, acts akin to an additional per-unit return. We refer to the introduction for our motivation of such preferences. The demand for ESG funds (and, hence, the extra-willingness to provide financing to sustainable firms) is assumed to be heterogeneous across households. The degree of heterogeneity is captured by the CDF $G(w)$ over $[0, \overline{w}]$. We suppose that a positive mass of households corresponds to a standard financial investor who only cares about financial payoffs, i.e.,

[11] We note that while (Pigouvian) taxes are a prominent tool for economists, they are not widely used in practice. In fact, outside environmental regulation they seem to be hardly used at all, while there are also only few cases where taxes on pollutants, such as CO_2 or SO_x, are used to govern emissions. One such case is the well-known European Union's Emissions Trading System (EU ETS), though even there the primary policy instrument are emission rights and thus quantity. While fuel taxes are more common in various jurisdictions, their prominence precedes environmental concerns. Also these observations justify that we focus first on a minimum sustainability standard as a primary policy instrument.

[12] See p. 18 on file:///C:/Users/inder/Downloads/consumer_testing_afm_report.pdf.

$G(0) > 0$.

While we suppose that total capital K is sufficiently large relative to the funding needs of all firms, we suppose that capital owned by households with warm-glow preferences $K(1 - G(0))$ is not sufficient to provide funding for all firms.[13] This assumption is aimed to capture the fact that ESG capital is still small relative to purely financial capital (despite its tremendous growth) and, hence, an economy cannot be run at full scale without financial capital. In addition, households have access to a storage technology that offers a fixed return r_0.

Social planner's objective function. We envisage an objective function of the social planner that encompasses only the consumption and production side, but not investors' perception of a warm glow. We thus conceive of w_j as merely a decisional utility (cf. the Introduction). As a result, the introduction of an ESG classification does not have an obvious, direct welfare effect, and we can focus on the real effects. When firm i chooses the sustainability level θ_i, invests I_i, and produces q_i, while total output is q, the respective welfare criterion is thus given by

$$\Omega = \int_0^q P(q)dq - \int I_i(1 + r_o)di - \int q_i[(\bar{\phi} - \theta_i) + c(\theta_i)]di. \tag{1}$$

This comprises, first, gross consumer welfare; second, investment costs; and third, costs of production, including the social externality. Since owner-managers do not shirk in equilibrium, this expression simplifies by using $q_i = I_i$ and $q = I$. Throughout the analysis we restrict consideration to cases where, first, it is profitable for firms to operate given some exogenously set minimum standard θ_m and, second, the economic activity is socially beneficial at least when the considered policy instruments are optimally chosen. As we will see, for the latter it is sufficient that

$$P(0) > (1 + r_0) + \min_\theta[c(\theta) + (\bar{\phi} - \theta)]. \tag{2}$$

Timing. We consider the following timing of events. At $t = 0$ the social planner chooses the minimum production standard θ_m (or, in our extension, a tax on externalities) as well as the threshold for sustainable investing θ_s. At $t = 1$, given this regulatory environment, firms then simultaneously choose their sustainability level θ_i as well as their desired outside financing L_i. Households optimally allocate funds, including to the storage technology. At

[13] See the proof of Proposition 2 for the exact condition.

$t = 2$, firm output q_i is realized, which is sold at product price $P(q)$. At this stage, firms also honor their promises D_i to households.

3 Preliminary analysis

We characterize the equilibrium investment by all firms for a given regulatory environment. In a first step, this requires us to determine the effect of financial constraints for each firm.

3.1 Financial constraints

When an owner-manager does not shirk, she realizes $I_i[P(I) - c(\theta_i)] - D_i$, where $c(\theta_i)$ depends on the chosen sustainability level θ_i. In case of shirking, the firm fails with probability Δp but she realizes private benefits BI_i. Incentive compatibility, thus, requires that $I_i[P(I) - c(\theta_i)] - D_i$ exceeds the sum of $(1 - \Delta p)[I_i[P(I) - c(\theta_i)] - D_i]$ and BI_i, which transforms to the requirement

$$D_i \leq I_i \left[P(I) - c(\theta_i) - \frac{B}{\Delta p} \right]. \qquad (IC)$$

This incentive compatibility constraint limits the external financing capacity, where $P(I) - c(\theta_i)$ measures the firm's gross return and $B/\Delta p$ can be interpreted as the per-unit agency rent.

We next incorporate the outside investors' participation constraint (IR). Given a required return r_i (and no shirking) for firm i the IR constraint is satisfied if investors are promised

$$D_i \geq (I_i - A_i)(1 + r_i). \qquad (IR)$$

The required return r_i will be endogenized below as a function of a firm's sustainability classification. Note that we need not distinguish between different classes of investors for a given firm and their required rate of return as firms in our model are of negligible size compared to the capital market.[14] Combining the manager's IC constraint and the investors' IR constraint we obtain an upper bound on investment for firm i

$$I_i \leq k_i A_i, \qquad (3)$$

where the investment-capacity multiplier k_i satisfies

$$k_i = \frac{1 + r_i}{B/\Delta p - [P(I) - c(\theta_i) - (1 + r_i)]}. \qquad (4)$$

[14] As we discuss in the Introduction, this is different in models where firms could negotiate with different, large-scale active investors with consequentialist preferences.

This investment-capacity multiplier implies a total debt capacity of $(k_i - 1)\, A_i$ and leverage of $\frac{k_i - 1}{k_i}$. The debt capacity is higher if the agency problem (as measured by $B/\Delta p$) is smaller, the equilibrium funding rate r_i is lower, and profitability is higher. In contrast to the standard Holmström-Tirole setup, an individual firm's profitability and, hence, debt capacity depends on the aggregate investment by all firms via $P(I)$.

To ensure that the debt capacity of any individual firm is finite, we presume that the agency problem is sufficiently severe relative to profitability (even at the highest product price and most favorable funding rate), i.e.,[15]

$$B/\Delta p > P(0) - c\,(\theta_m) - (1 + r_0 - \bar{w}). \tag{5}$$

This does not yet imply that aggregate industry output is limited by financial constraints. We return to this issue below. Also, it will depend on parameters whether a firm strictly prefers to lever up to (3).

If the owner-manager invests all of her own funds $A_i = A$, her expected payoff is given by $U_i = I_i\,[P(I^*) - c\,(\theta_i)] - D_i$.[16] Using binding (IR), we obtain

$$U_i = I_i\,[P(I^*) - c\,(\theta_i)] - (I_i - A)(1 + r_i). \tag{6}$$

3.2 Equilibrium

We first characterize optimal firm behavior and the resulting equilibrium both in the product market and the financial market given exogenous regulation parameters θ_m and $\theta_s > \theta_m$. We characterize the optimal planner problem in the subsequent section. In equilibrium, firms and households make optimal decisions.

Supply. Since households only care about whether a firm is classified as sustainable or not, which determines eligibility for the ESG fund, we only need to distinguish between two rates of return. The required rate of return for non-sustainable investments is equal to that obtained from the alternative investment opportunity, r_0. We denote the required rate of return for sustainable investment by r_s. "*Investor optimality*" requires that a household with warm-glow preferences $w_j \geq 0$ chooses to invest in a sustainable firm if and only if $w_j \geq \Delta r = r_0 - r_s$, where Δr measures the return subsidy for sustainable investments. Given this simple cut-off strategy the fraction of households that invests sustainably is $1 - G(\Delta r)$ so that the total supply of sustainable investment funds is given

[15] Hence, it is only a sufficient condition, as it is derived at $I = 0$, so that $P(I)$ is highest.

[16] In equilibrium, full coinvestment is always weakly optimal as the return on internal assets always weakly exceeds the outside option r_0.

by $K[1-G(\Delta r)]$. Intuitively, this supply is strictly decreasing in the return subsidy Δr for $\Delta r \in (0, \overline{w})$.

Demand. We now turn to our second equilibrium condition, "*firm optimality*" regarding the choice of the sustainability level θ_i. An individual firm takes as given both the market conditions of outside financing, r_0 and r_s, as well as the product price $P(I)$. Given the binary nature of the return required by outside investors and the costs of higher sustainability, a firm will optimally choose either the minimum standard θ_m or θ_s. Given the sustainability subsidy Δr, the firm makes this choice, as well as that of the external (loan) funding L_i and thus the investment size I_i, so as to maximize U_i in (6) subject to the outside financing (incentive) constraint (3). As a result of these optimal financing and production decisions, the aggregate amount of loans demanded by sustainable firms is given by $L_s(\Delta r) = \int_{\theta_i=\theta_s} L_i(\Delta r)\, di$.

Our third equilibrium condition relates to the market for sustainable funding ("*sustainable market clearing*"). The equilibrium sustainability subsidy Δr is determined by the intersection of demand and supply

$$L_s(\Delta r) = K[1 - G(\Delta r)]. \tag{7}$$

Characterization. Since we assume that warm-global capital is sufficiently scarce as to finance investment of all firms in the economy, a fraction of firms will optimally operate at the minimum production standard θ_m. Production at this minimum standard is profitable if and only if aggregate investment is below the unconstrained competitive investment capacity of $\bar{I}$ solving the zero-profit condition

$$P(\bar{I}) = 1 + r_0 + c(\theta_m). \tag{8}$$

We will refer to the economy as unconstrained (constrained) by the moral hazard problem if $I = \bar{I}$ $\left(I < \bar{I}\right)$. Note that $\bar{I}$ is a function of the minimum sustainability standard as it affects production costs.

Since both sustainable and non-sustainable firms are active in equilibrium, by the equilibrium conditions firms must be indifferent between 1) obtaining a financing subsidy of Δr as a compensation for higher production cost $c(\theta_s)$ and 2) incurring lower production cost $c(\theta_m)$ but unsubsidized financing conditions. Regardless of the sustainability choice, as firms are always financially constrained, the payoff of any firm can be expressed as the product of the per-unit agency rent, $\frac{B}{\Delta p}$, and maximal scale, Ak_i:[17]

$$U_i = \frac{B}{\Delta p} A k_i. \tag{9}$$

Maximal scale is always strictly optimal as long as the economy is constrained in the aggregate, i.e., $I < \bar{I}$, since entrepreneurs earn a scarcity rent. Interestingly, this expression even holds if $I = \bar{I}$ and non-sustainable firms do not have a strict incentive to lever up to the maximum scale. In this case, (4) implies that $k_i = \frac{1+r_0}{B/\Delta p}$, so that $U_i = A(1+r_0)$ simply reflects the return on the storage outside option.

Indifference between being sustainable or not requires that the investment capacity of sustainable and non-sustainable firms be equalized in equilibrium (see (9)). This observation allows us to express the resulting equilibrium investment capacity multiplier for all firms, k, solely as a function of aggregate investment I (given primitives (r_0, θ_m)):

$$k(I) = \frac{1+r_0}{B/\Delta p - [P(I) - c(\theta_m) - (1+r_0)]}.^{18} \tag{10}$$

Intuitively, the dependency of the debt capacity multiplier on I arises because larger aggregate investment increases the quantity of supplied goods, which pushes down product prices (and, hence profitability).

We now turn to the equilibrium value of aggregate investment $I^* \leq \bar{I}$ and the resulting equilibrium investment capacity multiplier $k^* := k(I^*)$.

Proposition 1 (Aggregate investment I^*) *If total internal funds A are sufficiently constrained, $A < \bar{I}\frac{B/\Delta p}{1+r_0}$, the economy remains financially constrained. The equilibrium value of aggregate investment, I^*, is then the unique solution to*

$$I = Ak(I) \tag{11}$$

If $A \geq \bar{I}\frac{B/\Delta p}{1+r_0}$, the economy is not constrained by the moral hazard problem and $I^* = \bar{I}$ as defined in (8).

Intuitively, Proposition 1 highlights that the moral hazard problem constrains aggregate investment only if firm assets A are sufficiently low. Given I^*, we immediately obtain the equilibrium product prices $P(I^*)$ as well as the investment capacity multiplier for all firms $k^* := k(I^*)$.

[17] It is useful to recall that k_i is strictly decreasing in $B/\Delta p$ and that, taken together, also U_i is strictly decreasing.

[18] This expression results from substituting the primitives (r_0, θ_m) into (4).

We now turn to the equilibrium effect of sustainability preferences. While the distribution of investor preferences G does not affect aggregate investment I^*, it has important compositional implications for production, i.e., the fraction of firms that produce sustainably versus the ones that produce at the minimum standard. To derive these compositional effects, we initially determine the equilibrium subsidy for producing sustainably, as given by the equilibrium capital cost differential Δr^*. Firm indifference requires that this interest rate differential is simply the levered cost differential using the equilibrium multiplier k^*

$$\Delta r^* = \Delta c \frac{k^*}{k^* - 1}. \tag{12}$$

Given Δr^*, a fraction $1 - G(\Delta r^*)$ of total capital K is deployed to the sustainable fund, which in conjunction with co-financing by entrepreneurs results in total sustainable investment of

$$I_s^* = K[1 - G(\Delta r^*)] \frac{k^*}{k^* - 1}. \tag{13}$$

This allows us to write the (equilibrium) fraction of sustainable investment in the economy as follows:

Proposition 2 (The sustainability subsidy and the composition) *In equilibrium, given the interest rate differential Δr^* in (12), the share of sustainable investment equals*

$$\sigma^* = \frac{I_s^*}{I^*} = \frac{K\left[1 - G\left(\Delta c \frac{k^*}{k^*-1}\right)\right] \frac{k^*}{k^*-1}}{\min\left\{\bar{I}, Ak^*\right\}}. \tag{14}$$

This characterization in Proposition 2 yields unambiguous comparative statics. We first analyze the effect of a demand trend for holding ESG funds.

Corollary 1 *As preferences for sustainability increase (by way of a First-Order Stochastic Dominance shift in $G(w)$), the share of sustainable investment σ^* increases, while the respective financing conditions, r_0 and Δr^*, remain unchanged.*

Intuitively, when ceteris paribus there is a greater supply of sustainable capital, this results in a greater share of sustainable investment. Still the financing subsidy Δr^* remains unchanged, as in equilibrium this is pinned down by firms' endogenous decision to become more sustainable and the resulting indifference condition (12).[19] Further, we can analyze

[19] In this sense, Corollary 1 compares the outcome after the respective equilibrium adjustments. If over a shorter time horizon firms' sustainablity levels remained unchanged, Δr^* would decrease. As we have shown in a working paper version, availability of sustainable funding would still have real effects by increasing the share of sustainable investments.

the effects of policy changes in the form of a stricter threshold for sustainability θ_s. The following comparative results follow directly from inspection of (12) and (13).

Corollary 2 *An increase in the sustainability standard θ_s decreases the share of sustainable investment σ^* and increases the equilibrium differential Δr^*.*

While a higher sustainability standard does not have an effect on aggregate investment (see Proposition 1), it increases the cost differential Δc for producing sustainably relative to the minimum standard. This higher cost differential, in turn, requires the capital cost subsidy for sustainable firms to go up, so as to keep sustainable production equally attractive, see (12). The required increase in the subsidy needs to be paid by households and, hence, reduces the attractiveness of the ESG fund for all households. As a result, previously marginal households no longer invest sustainably. This comparative statics highlights a key trade-off that a regulator is facing in our upcoming normative analysis. While increasing the sustainability cutoff reduces the negative externalities of more sustainable firms, it reduces the fraction of firms that choose to produce sustainably.[20]

For subsequent derivations we conclude with the following comparative result:

Corollary 3 *If the minimum standard θ_m increases, holding θ_s constant, total investment I^* decreases while sustainable investment I_s^* increases (implying also an increase in the respective share σ^*).*

The comparative result for I^* is immediate as with a higher standard θ_m costs of production increase for firms that do not satisfy the higher standard θ_s (and which account for the marginal investment). As now the less sustainable production becomes less attractive, the capital cost differential that is required to make firms indifferent in case of a more sustainable choice decreases. In equilibrium this increases the supply of the respective sustainable capital and with this sustainable investment.

4 Optimal policy

4.1 Optimal sustainability classification

For now we still take the minimum sustainability standard θ_m as given, so that the single policy instrument is the possible classification of sustainable investments through a higher

[20] Note that the lower equilibrium differential Δr^* following a reduction of the threshold θ_s is thus not an immediate effect of investors' lower appreciation for sustainability when the threshold is lower. Recall that in our model households have non-consequentialist preferences, so that the warm-glow effect that they experience is independent of the threshold.

standard θ_s. Recall from the Introduction that in many instances, notably outside the area of environmental sustainability, such a minimum standard may in fact be very low and not subject to frequent changes. For instance, for firms operating outside the jurisdiction or procuring from other countries, supply chain regulation may prescribe compliance only with a limited set of fundamental human rights, such as prohibiting the use of slave labour. In this case, the standard θ_s could prescribe, in addition, certain rights for employees or trade union activity. Also, in the case of environmental pollution, for a given pollutant θ_m may be chosen across all economic activities with a view only on of avoiding immediate health risks. Again, θ_s then represents more stringent requirements.

Recall now the social planner's objective function (1). Based on our preceding equilibrium characterization, this can be rewritten, using thereby, first, that in equilibrium all owner-managers are incentivized not to shirk; second, that firms choose only between θ_m and θ_s; and, third, our definition of I^* and I_s^* for the respective total investment sizes:

$$\Omega = \int_0^{I^*} P(I)dI - I^*[c(\theta_m) + (\bar{\phi} - \theta_m) + (1 + r_0)] + I_s^* [(\theta_s - \theta_m) - (c(\theta_s) - c(\theta_m))] . \quad (15)$$

The first term in (15) captures welfare when all investment was non-sustainable, with $\theta_i = \theta_m$. The second term captures the effect of sustainable investment of size I_s^*. As total investment does not depend on θ_s, this instrument only affects the social planner's objective through the second term, so that the social planner's program reduces to that of maximizing the product of I_s^* and $(\theta_s - \theta_m) - (c(\theta_s) - c(\theta_m))$. For any given unit of investment and production, the social planner would want to implement the level of sustainability θ_s that maximizes the second factor of this product, which is the case when $c'(\theta_s) = 1$. For future reference it is useful to denote this level of sustainability by θ_{int}: the level that would be optimal if the social planner had to consider only the *internal margin*. The introduction of a sustainability classification $\theta_s > \theta_m$ is therefore only beneficial when the prevailing minimum standard is below θ_{int}. The social planner can, however, not dictate firms to choose θ_s, but the respective incentives must come from the capital market. Consequently, a higher level θ_s necessarily implies a higher subsidy of capital costs Δr^*. With the constrained and downward sloping supply of sustainable financing, I_s^* must then decrease. This trade-off becomes attenuated when the supply of sustainable financing increases, giving rise to the following result:

Proposition 3 (Optimal sustainable investment classification) *Taking for now the minimum sustainability standard θ_m as given, when $\theta_m \geq \theta_{int}$, it is strictly suboptimal* *to introduce a higher standard $\theta_s > \theta_m$ for the classification of sustainable investment. When $\theta_m < \theta_{int}$, it is instead strictly optimal to introduce such a classification with $\theta_m < \theta_s^* < \theta_{int}$. Then, as investor preferences shift upwards in the sense of a Monotone Likelihood Ratio Property shift in $G(w)$, the optimal threshold $\theta_s^* > \theta_m$ strictly increases.*

Proposition 3 is our first normative result. If the minimum standard, which we presently hold fixed, is already sufficiently large with $\theta_m \geq \theta_{int}$, it is strictly suboptimal to introduce a classification for sustainable investment that requires an even stricter standard $\theta_s > \theta_m$. To digest this result, recall that investors' warm-glow preferences do not enter into the social planner's objective, which thus comprises only real effects. While also for $\theta_s > \theta_m \geq \theta_{int}$ such a class of sustainable investment opportunities would attract investors and would thereby lead to subsidized capital costs for firms that choose a sufficiently high standard, in terms of real effects this would be strictly suboptimal.[21] While the choice of θ_s would then still be privately optimal for the respective firms, the resulting subsidized financing would now compensate for the incremental costs of achieving a sustainability standard that is too high. To the extent that we realistically capture (warm-glow) preferences of households investing in the respective ESG-funds, the first part of Proposition 3 thus qualifies the notion of a general social desirability of an ESG-classification of investment funds. Even when such investment opportunities meet with positive demand, this could represent "too much of a good thing".

As the second part of Proposition 3 shows, however, such a qualification clearly no longer applies when the minimum standard is sufficiently low with $\theta_m < \theta_{int}$. Then, it is efficient to introduce a sustainability classification for investments, which allows to harness the respective preferences of households so as to (indirectly) subsidize the choice of a higher sustainability level of some firms. As we discussed in the Introduction, there seems to be indeed a strong conviction that the prevailing (minimum) standard of sustainability is too low in many instances, such as with respect to CO_2 emissions. Recall next that the higher is θ_s^*, the larger must be the required subsidy Δr^* to make firms indifferent, which reduces the amount of sustainable funds that can be accessed. The second part of Proposition 3 formalizes how the social planner must bear in mind this trade-off arising from the downward sloping supply of sustainable capital: First, the social planner should always set θ_s^* strictly below the level that would be optimal when considering only the internal margin, thereby holding I_s^* fixed; second, the trade-off is mitigated when investors' sustainability

[21] Trivially, when θ_s was chosen so high that $c(\theta_s) - c(\theta_m) \geq \overline{w}$, where $\overline{w}$ also represents the maximum value for Δr^*, I_s^* would be zero.

preferences increase, so that θ_s^* should then be increased. When the respective shift in $G(w)$ occurs over time, rather than "challenging" the market with a high initial classification threshold, the social planner should optimally apply a lower threshold when sustainable preferences for investment are still less prevalent. In this sense, Proposition 3 provides also a rationale for increasing the standard over time. Below we will further investigate such dynamics by introducing a feedback loop when investors' preferences are shaped by an endogenous social norm.

4.2 Sustainability classification under optimal (environmental) policy

An important step in our overall analysis is the joint consideration of the classification of sustainable investment *and* other (environmental) policy instruments of a social planner, as presently represented by the minimum standard θ_m. We thus consider now applications where the setting of such a minimum standard is realistic. Our key question is whether in this case there is still scope for introducing additionally a higher standard $\theta_s > \theta_m$ and with this such a classification of investments. The optimality of this is a priori not obvious as the social planner could now directly target the sustainability level of all firms, irrespective of the supply of sustainable capital. Recall again that as the social planner's objective is restricted to real effects, the introduction of sustainable investment opportunities does not immediately lead to social benefits by satisfying the respective investor preferences.

We consider first, as a benchmark, the setting of the optimal minimum standard when the social planner does not avail herself of the additional instrument of a classification of sustainable investment. In a second step we consider the joint optimization over both instruments.

Benchmark: Optimal minimum standard. Ignoring for now θ_s and thus also setting $I_s^* = 0$, differentiating Ω in (15) with respect to θ_m we obtain

$$\frac{d\Omega}{d\theta_m} = \frac{dI^*}{d\theta_m}\left[P(I^*) - \left(c\left(\theta_m\right) + (\bar{\phi} - \theta_m) + (1 + r_0)\right)\right] - I^*[1 - c'(\theta_m)]. \tag{16}$$

Here, the first term captures the external margin, i.e., how a change in the minimum standard affects total investment and production, with social surplus comprising consumer surplus, cost of production, and externalities. The second term captures the internal margin for a given level of investment and production, i.e., the trade-off between higher costs of production and lower externalities.

It is now instructive to first consider the case where the economy is not constrained in the aggregate (as we slightly adjust θ_m), so that $I^* = \bar{I}$, as defined in (8). Substituting this back into (16), we have for this case the first-order condition for θ_m^*

$$\frac{d\Omega}{d\theta_m} = -(\bar{\phi} - \theta_m^*)\frac{d\bar{I}}{d\theta_m} - \bar{I}[1 - c'(\theta_m^*)] = 0. \tag{17}$$

Intuitively, when the economy is not financially constrained, the external margin reduces to the (negative) externality $-(\bar{\phi}-\theta_m^*)$. Observe now that total investment is strictly lower when θ_m increases, simply as the respective costs $c(\theta_m)$ increase, so that from $\frac{d\bar{I}}{d\theta_m} < 0$ the first term in (17) is strictly positive. This implies that when the economy remains financially unconstrained, the optimal standard satisfies $\theta_m^* > \theta_{int}$ (where θ_{int} is the level at which the internal margin is just zero, $c'(\theta_{int}) = 1$). Hence, to reduce aggregate economic activity, which generates negative externalities, the social planner wants to increase the minimum standard even beyond θ_{int}.

We next suppose instead that firms' internal funds A are sufficiently small. Then, financial constraints bind not only on the level of individual firms, but also for the industry in the aggregate. In this case, the external margin in (16) no longer reduces to the externalities, as $P(I^*) > c(\theta_m) + (1 + r_0)$. As long as the external margin remains, however, still negative, the preceding observations apply, so that $\theta_m^* > \theta_{int}$. This finding reverses, however, when the financial constraints become sufficiently important relative to the avoided externalities.

Proposition 4 (Optimal minimum standard) *Suppose the social planner can impose a minimum standard θ_m, but that she cannot introduce a classification of sustainable investment (through a respective higher threshold θ_s). When A is sufficiently high and the economy is thus sufficiently unconstrained in the aggregate, the optimal minimum standard satisfies $\theta_m^* > \theta_{int}$ (where $c'(\theta_{int}) = 1$). The converse with $\theta_m^* < \theta_{int}$ holds only when A is sufficiently low.*

Joint optimization. Recall now from Proposition 3 the characterization of the optimal classification of sustainable investments, i.e., of the respective standard θ_s. Taking there the minimum standard θ_m as a given, we found that when $\theta_m \geq \theta_{int}$, it is strictly suboptimal to introduce a higher classification standard $\theta_s > \theta_m$, while the opposite holds when $\theta_m < \theta_{int}$. This insight can now be combined with the characterization of the optimal minimum standard in the absence of an additional instrument (Proposition 4):

Proposition 5 (Joint optimization of instruments) *Suppose the social planner can impose a minimum standard θ_m. Then the introduction of a higher standard for the classification of sustainable investment is only optimal when the economy is sufficiently financially constrained in the aggregate, in which case the optimal choices satisfy $\theta_m^* < \theta_s^* < \theta_{int}$. In this case, the minimum sustainability standard is optimally chosen higher than without the additional instrument of a sustainable classification. Otherwise, when financial constraints are sufficiently unimportant (or even absent) in the aggregate, it is strictly suboptimal to introduce such a classification for sustainable investment.*

If the minimum sustainability standard is chosen optimally from the social planner's perspective, absent financial frictions in the aggregate there is no case for introducing a classification for sustainable investments and to thereby lower the capital costs for firms that choose an even higher sustainability standard. What is more, such an additional instrument would be *strictly* suboptimal. This follows as in the absence of financial constraints, the optimal minimum standard satisfies $\theta_m^* > \theta_{int}$: So as to also adjust downwards to size of the respective economic activity, which creates a harmful externality, the standard is strictly higher than the level of sustainability that would optimally govern the internal margin, θ_{int}. It is then immediate that it is strictly suboptimal to subsidize firms for choosing an even higher standard, while keeping the overall level of economic activity, I^*, unchanged. These observations still apply when the economy is (marginally) financially constrained, as long as $\theta_m^* \geq \theta_{int}$.

When the economy is however sufficiently constrained, the marginal economic activity has positive social value, as the thereby generated private surplus, comprising the benefits to consumers, exceeds the generated externality. In this case the introduction of a sustainable investment category and with it the subsidization of more sustainable firms adds value, as it allows to raise the sustainability level of some firms without compromising the size of the economy, i.e., while leaving I^* unchanged. As also shown in Proposition 5, it is then optimal to increase the minimum sustainability standard compared to the optimal choice when the instrument of a sustainable investment classification was not available. The intuition derives from Corollary 3, by which an increase in θ_m makes the more sustainable investment relatively more attractive for firms, thereby reducing the required capital cost differential and ultimately increasing sustainable investment I_s^* (even though total investment I^* decreases). In this sense, when it is optimally to use both instruments, the two instruments are complements: When a sustainable investment classification is introduced, the minimum sustainability standard, by which all other firms must operate, should optimally increase.

4.3 A tax on externalities

As noted above, in many instances and for various reasons negative externalities are not dealt with through (Pigouvian) taxes but through a minimum standard (such as limits on emissions or immissions). We show now that when such taxes are instead feasible, our key insights still apply.

Set-up and overview of results. Now, the social planner can impose a tax τ on externalities. Recall that we simplified the exposition by assuming that the marginal welfare cost associated with externalities was exactly one, so that a Pigouvian tax would stipulate $\tau = 1$. In a slightly informal analysis, we now ignore for a moment financial constraints. Ignoring also the instrument of a sustainable investment classification, with $\tau = 1$ a firm would optimally choose θ_i to minimize $c(\theta_i) + \tau(\bar{\phi} - \theta)$, so that when $\tau = 1$ we have $\theta_i^* = \theta_{int}$ (solving $c'(\theta_{int}) = 1$). With unconstrained financing the equilibrium size of the economy would be pinned down by the zero-profit condition

$$P(I^*) = (1 + r_0) + c(\theta_i^*) + \tau(\bar{\phi} - \theta_i^*),$$

which for $\tau = 1$ thus equates the marginal consumer surplus with marginal costs including the true cost of the externalities. Hence, abstracting from financial constraints a tax set at the Pigouvian level of $\tau = 1$ would lead to the welfare optimum, where both the internal and the external margin are set efficiently.[22] In this case the introduction of a sustainable investment classification would not be optimal. As the economy becomes, however, financially constrained, as we show formally in what follows, the tax is optimally reduced so as to alleviate the financial burden and thus increase the equilibrium size I^*, implying that $\theta_i^* < \theta_{int}$. As the resulting sustainability level thus falls below the level that would be efficient from the perspective of the internal margin alone, we know that the introduction of a higher sustainable investment threshold becomes strictly optimal. This confirms our qualitative insights from the preceding analysis with a minimum sustainability standard, albeit now the investment classification becomes strictly optimal if and only if financial constraints bind in the aggregate.

There is an additional difference between the imposition of a tax and the imposition of a minimum standard. The minimum standard imposes on firms a financial burden only

[22] It is immediate that this choice maximizes the social planner's objective function (15).

through the respective costs $c(\theta_m)$. Instead, when a tax τ is imposed so as to thereby incentivize a sustainability level θ_i, firms incur both the respective costs $c(\theta_i)$ and the tax burden $\tau(\bar{\phi} - \theta_i)$. In a world without financial constraints (and where distributive considerations could be ignored) the additional tax burden would be inconsequential. Instead, when financial constraints bind in the aggregate this reduces the equilibrium size of the economy. Thereby, the tension between the internal margin, i.e., the objective to choose a sufficiently high sustainability level for any given unit of investment and output, and the external margin, i.e., not to inefficiently shrink the size of the respective economic activity, is increased. This generates further scope for subsidizing a higher sustainability level for some firms through the respective investment classification, without negatively affecting aggregate output.[23]

In what follows we formalize these insights. For this we need to modify our respective expressions and analysis accordingly.

Preliminary analysis. Denote now the costs of production including taxes on externalities by $\widetilde{c}(\theta_i) = c(\theta_i) + \tau(\bar{\phi} - \theta_i)$. Recall that these costs are incurred only when there is success (which is the case with probability one if the owner-manager does not shirk). Thus, when an owner-manager does not shirk, she now realizes $I_i[P(I) - \widetilde{c}\,(\theta_i)] - D_i$. From this the incentive constraint (IC) remains unchanged once we replace c by $\widetilde{c}$. Observe next that the investors' participation constraint (IR) does not change. Combining the two constraints again gives rise to the cap on investment $I_i \leq k_i A_i$, where the investment-capacity multiplier k_i is obtained again by replacing c with $\widetilde{c}$. It is next useful to write out the objective of the owner-manager

$$U_i = I_i \left[P(I^*) - \widetilde{c}(\theta_i)\right] - (I_i - A)(1 + r_i).$$

Now if the choice of θ_i does not affect r_i, there is a single value θ_i that both maximizes k_i (and thus I_i) and $P(I^*) - \widetilde{c}(\theta_i)$. So as not to overburden the exposition with additional notation, we denote this (lower) level of sustainability again by θ_m, satisfying $\widetilde{c}'(\theta_m) = 0$ (or likewise $c'(\theta_m) = \tau$). When the economy is not financially constrained in the aggregate, the respective size $I^* = \bar{I}$ thus solves $P(\bar{I}) = 1 + r_0 + \widetilde{c}(\theta_m)$. If constrained, we have $k^* := k\,(I^*)$, where the multiplier again adjusts by use of $\widetilde{c}$. Hence, Proposition 1 for the characterization of I^* again applies, with the respective adjustment term $\tau(\bar{\phi} - \theta_m)$ in the

[23] In a cap-and-trade scheme, Heider and Inderst (2021) consider the additional instrument of allocating emission permits for free, thereby reducing the respective financing burden, albeit they do not consider the instrument of a sustainable investment classification.

definition of $\bar{I}$ and k^*. Turning to the characterization of I_s^* in Proposition 2, again the unique adjustment that is necessary relates to costs. That is, the cost differential in the definition of Δr^* now includes tax payments, so that $\Delta c = \widetilde{c}(\theta_s) - \widetilde{c}(\theta_m)$, where we note that $\Delta c > 0$ holds from $\widetilde{c}'(\theta_m) = 0$.

Proposition 6 (Characterization in case of a tax on externalities) *The characterization of the equilibrium with a minimum standard in Propositions 1 and 2 extends to the case with a tax on externalities, after augmenting the costs of production by the respective tax payments, i.e., replacing $c(\theta)$ by the expression $\widetilde{c}(\theta) = c(\theta) + \tau(\bar{\phi} - \theta)$.*

Optimal policy. As in the case with a minimum standard in Proposition 3, we first consider the choice of the optimal threshold for the sustainable investment classification for a given choice of the tax on externalities. Note first that the social planner's objective function is still given by (15), where θ_m is now a function of the potential instrument τ.[24]

Proposition 7 (Optimal sustainable investment classification under a tax regime) *Take now a tax $\tau \leq 1$ on externalities as given, which results in a standard $c'(\theta_m) = \tau$. When τ lies below the Pigouvian level of $\tau = 1$, it is strictly optimal to introduce a higher standard $\theta_m < \theta_s^* < \theta_{int}$ for the classification of sustainable investment. When $\tau = 1$ it is strictly suboptimal to introduce a higher standard θ_s.*

We turn next to the combination of the two instruments. Here, we refer to the intuition that we have already laid out above. As described there, now the key case distinction refers to whether the economy is financially constrained in the aggregate when the Pigouvian level $\tau = 1$ is chosen, taking into account the additional tax payments $\tau(\bar{\phi} - \theta_{int})$.

Proposition 8 *Suppose the social planner can both impose an optimal tax τ on externalities and introduce a classification of sustainable investment with threshold θ_m. When at the Pigouvian level of the tax the economy is not financially constrained in the aggregate, $\tau^* = 1$ is uniquely optimal and it is strictly suboptimal to introduce a classification of sustainable investment. Otherwise, the optimal tax is strictly lower, $\tau^* < 1$, and it is strictly optimal to introduce such classification. In this case, the tax on externalities is optimally chosen higher than without the additional instrument of a sustainable investment classification.*

[24] We thus treat the tax as a welfare neutral transfer.

5 Social network effects

Recall from above that we interpreted a comparative analysis in $G(w)$ also across time. Such a change in investor preferences may also be brought about by a change in social norms. We now allow for such a feedback mechanism as we suppose that individual preferences for sustainable investment depend on the observed or anticipated behavior of other citizen. We thus suppose that the willingness-to-pay w for a more sustainable investment, i.e., the discount that households are willing to tolerate relative to a non-sustainable investment, increases when more households make such investment or, likewise, when there are fewer households making non-sustainable investments.

We first provide some motivation for these preferences. At the heart is the acknowledgement, which we borrow from a large literature on environmental and resource economics, that preferences for sustainability derive from non-use benefits[25] and as such are particularly susceptible to social norms, which in turn can be shaped by the anticipated or observed behavior of others. Such a notion of how social norms change is widely shared in social sciences,[26] and also recognized in economics. There, the relevance of the behaviour of others has also been confirmed in experiments, such as in games of contributions to a public good.[27] Also various field experiments relate especially environmentally conscious behavior to that of others.[28] Endogenous preferences and changes in norms have also been incorporated in policy suggestions to fight climate change.[29] Our key notion is thus that households' investment preferences are not exogenously fixed but endogenous in that they depend on the behavior of others.[30]

In what follows, we thus assume that the distribution of preferences for sustainable investment depends on how widespread the respective investments are among households. For simplicity, we capture this via a dependency on the equilibrium cutoff $w^* = \Delta r^*$, writing $G(w \mid w^*)$. Here we stipulate that as w^* decreases, so that sustainable investments are more widespread, the distribution moves upwards in the sense of strict First-Order Stochastic Dominance (FOSD). Regarding our preceding derivations, we only need to

[25] On this concept see, for instance, Pearce et al. (2006).

[26] In fact, also there the relevance for public policy has been recognized early, e.g., by Cialdini et al. (1990).

[27] See Sudgen (1984) for an early example.

[28] For instance, individual recycling behavior has been found to correlate strongly with beliefs about such behaviour in the community; see Cialdini (2003) or the various studies quoted in Schultz (2002).

[29] Stiglitz (2019) dedicates a separate subsection to endogenous preferences.

[30] We acknowledge that an observed positive feedback effect may also have other reasons, e.g., learning or imitating, or supply-side changes resulting in an increased offering and promotion of such investments as demand increases.

modify expression (13), where now

$$I_s^* = K[1 - G(w^* \mid w^*)]\frac{k^*}{k^* - 1} \text{ with } w^* = \Delta r^*. \tag{18}$$

We are interested in how the normative implications change, given that in $dI_s^*/d\theta_s$ we now need to take the total derivative of $G(w^* \mid w^*)$, where we use the shift according to FOSD as $G_2(w^* \mid w^*) > 0$. When the social planner chooses the optimal threshold θ_s, this feedback effect affects her decision as follows. When she reduces θ_s and thereby also w^*, this has both a direct positive effect on $1 - G(w^* \mid w^*)$ and an indirect positive effect through the (social norm) feedback. This makes it optimal to choose a lower threshold when such a feedback effect is present.

Proposition 9 *Suppose households' preferences are subject to the described social-norm feedback. Then the social planner should optimally choose a strictly lower threshold θ_s^* for the classification of sustainable investment (provided it is optimal to introduce such a classification).*

6 Concluding remarks

In this paper, we introduce investors with sustainable preferences into an economy with financially constrained firms. When a taxonomy for sustainable investment is introduced, firms that meet the respective sustainability standard can thus tap into cheaper financing. We pose two questions: first, whether the introduction of such a sustainable investment classification is beneficial at all, notably when a social planner can optimally choose more direct policy instruments, such a s a minium standard that all investment and production must satisfy; second, if such a classification is beneficial, how stringent should the respective threshold be chosen.

The answer to the first question is rendered non-trivial as in our model the social planner only cares about real effects, which we motivate by stipulating that households merely derive a (decisional) warm glow from any Euro invested according to the chosen sustainable (ESG) taxonomy. Whether the indirect subsidy that arises from the sustainable classification is socially beneficial depends on two determinants: first, on whether a more direct instrument such as a minimum standard or a tax on externalities can be optimally chosen and, second, whether the economic activity is subject to financial constraints. As noted in the Introduction, the view that political failure may prevent a socially efficient choice of more direct instruments has recently gained traction.[31] From that perspective, if there is window of opportunity, such as with the introduction of a sustainable investment category, this should be used to ratchet up the regulatory standard. But even if it is possible to choose the socially optimal level of the direct policy instrument, we show that the indirect subsidy arising from the sustainable investment classification adds a valuable second instrument that mitigates the trade off between achieving higher sustainability for given investment and production (internal margin) and avoiding an inefficient contraction of the considered economic activity (external margin). This echoes the recent discussion of Joseph Stiglitz (Stiglitz 2019), who has forcefully argued for an "alternative approach" to environmental regulation, which uses various instruments so as to deal with real-world complexities, including distributional issues, imperfect markets, or risk and uncertainty.[32] We also analyze potential determinants of the optimal design of the classification so as to thereby maximize the obtained indirect subsidy, given a downward-sloping supply of sustainable financing.

If none of these cases apply, the (additional) introduction of a classification for sustainable investment is socially strictly suboptimal. This is an immediate consequence of our assumption that households' warm-glow preferences are not part of the social planner's objective function. In the Introduction we offered some support for this assumption. While the introduction of such a classification would then meet with even potentially high demand by households, in our model there are thus no immediate benefits from "completing the market" by introduction of such a category.

Our analysis is restricted to a combination of at most a direct instrument, such as a minimum standard, and a sustainable investment classification, where the latter provides an indirect subsidy for more sustainable investments. When the social planner could provide such subsidies also directly, again this would erode the role for such classification. While raising the necessary financing for a direct subsidy through taxes should lead to additional distortions, in practice also the introduction of the sustainable taxonomy and its monitoring are not free of transaction costs. What is more, the eagerness with which sellers of financial products develop and advertise such products may also darken the

[31] This was potentially also fuelled by a CO_2 price that was viewed as too low, though recently this has more than doubled.

[32] That a single instrument, such as a tax on externalities, can be insufficient has been recognized also earlier in the literature, notably in imperfect oligopolies and endogenous market structure (cf. Carlton and Loury 1980 and the cited references). There exists also a large literature that analyzes the interaction of externality-related taxes on consumption with other distortionary taxes (cf. the overview in Bovenberg and Goulder 2002).

overall picture. Potential benefits may be frittered away by a revival of actively managed funds with high fees and high (marketing) expenditures, and the packaging and advertising of such funds may also involve a great deal of greenwashing.[33] While such considerations are outside our model, still our analysis cautions the frequently encountered optimism regarding more sustainable (ESG) investment.

[33] Even of fraudulent nature; cf. the recent actions against DWS, the in-house provider of investment funds of Deutsche Bank

References

Andersson, Mats, Patrick Bolton, and Frédéric Samama. 2016. Hedging Climate Risk. Financial Analysts Journal 72, 13-32.

Baker, Malcolm, Daniel Bergstresser, George Serafeim, and Jeffrey Wurgler. 2018. Financing the Response to Climate Change: The Pricing and Ownership of US Green Bonds. NBER.

Bizzotto, Jacopo and Bard Harstad. 2021. The Choice of Certifier in Endogenous Markets. Mimeo.

Bovenberg, Lans and Lawrence Goulder. 2002. Environmental Taxation and Tegulation. In: Auerbach, Alan and Martin Feldstein (ed.), Handbook of Public Economics, Chapter 23, 1471–1545. Elsevier.

Boyle, Kevin, William H. Desvousges, and F. Reed Johnson. 1994. An Investigation of Part-Whole Biases in Contingent-Valuation Studies. Journal of Environmental Economics Management 27, 64–83.

Chowdhry, Bhagwan, Shaun W. Davies, and Brian Waters. 2014. Incentivizing Impact Investing. In Conference on the Impact of Responsible and Sustainable Investing. Honk Kong: Hong Kong University of Science and Technology.

Cialdini, Robert B. 2003. Crafting Normative Messages to Protect the Environment. Current Directions in Psychological Science 12, 105-109.

Cialdini, Robert B., Raymond R. Reno, and Carl A. Kallgren. 1990. A Focus Theory of Normative Conduct: Recycling the Concept of Norms to Reduce Littering in Public Places. Journal of Personality and Social Psychology 58, 1015-1026.

Desvousges, William H., Kristy Mathews, and Kenneth Train. 2012. Adequate Responsiveness to Scope in Contingent Valuation. Ecological Economics 84, 121-128.

Gibson, Rajna and Philipp Krueger. 2018. The Sustainability Footprint of Institutional Investors. Swiss Finance Institute Research Paper.

Hartzmark, Samuel M. and Abigail B. Sussman. 2019. Do Investors Value Sustainability? A Natural Experiment Examining Ranking and Fund Flows. The Journal of Finance 74, 2789-2837.

Heider, Florian and Roman Inderst. 2021. A Corporate Finance Perspective on Environmental Policy.

Heinkel, Robret, Alan Kraus, and Josef Zechner 2001. The Effect of Green Investment on Corporate Behavior. Journal of Financial and Quantitative Analysis 36, 431-449.

Henke, Hans-Martin. 2016. The Effect of Social Screening on Bond Mutual Fund Performance. Journal of Banking and Finance 67, 69-84.

Holmström, Bengt and Jean Tirole. 1997. Financial Intermediation, Loanable Funds, and the Real Sector. Quarterly Journal of Economics 112, 663–691.

Kahneman, David and Jack L. Knetsch. 1992. Valuing Public Goods: The Purchase of Moral Satisfaction. Journal of Environmental Economics and Management 22, 57-70.

Landier, Augustin and Stefano Lovo. 2020. ESG Investing: How to Optimize Impact?. HEC Paris Research Paper No. FIN-2020-1363.

Oehmke, Martin and Marcus Opp. 2020. A Theory of Socially Responsible Investment. CEPR Discussion Paper DP14351.

Pearce, David W., Giles Atkinson, and Susana Mourato. 2006. Cost-Benefit Analysis and the Environment: Recent Developments. Paris, France: OECD Publishing.

Riedl, Arno and Paul Smeets. 2017. Why Do Investors Hold Socially Responsible Mutual Funds? The Journal of Finance 72, 2505-2550.

Schultz, P. Wesley. 2002. Knowledge, Information, and Household Recycling: Examining the Knowledge-Deficit Model of Behavior Change. In: T. Dietz and P. Stern (Eds.) Education, Information, and Voluntary Measures in Environmental Protection, 67-82. Washington, DC: National Academy of Sciences.

Stiglitz, Joseph E. 2019. Addressing Climate Change through Price and Non-price interventions. European Economic Review 119, 594-612.

Sudgen, Robert. 1984. Reciprocity: The Supply of Public Goods Through Voluntary Contributions. The Economic Journal 94, 772-787.

Tirole, J. 2012. Some Political Economy of Global Warming. Economics of Energy & Environmental Policy 1, 121-132.

Zerbib, Oliver D. 2019. The Effect of Pro-Environmental Preferences on Bond Prices: Evidence from Green Bonds. Journal of Banking and Finance 98, 39-60.

Proofs

Proof of Proposition 1. We can restrict attention to equilibria with $I^* \leq \bar{I}$ since for any $I > \bar{I}$ firms would earn a lower return than their outside option r_0. We now turn to the question whether $\bar{I}$ is feasible. Suppose that aggregate investment is at $\bar{I}$, then the associated investment capacity multiplier is $k\left(\bar{I}\right) = \frac{1+r_0}{B/\Delta p}$. Each individual firm takes this multiplier as given. We now distinguish between two cases.

If firm assets are below $\bar{I}\frac{B/\Delta p}{1+r_0}$, then even if all firms were to lever up to the maximum (using the candidate multiplier $k\left(\bar{I}\right) = \frac{1+r_0}{B/\Delta p}$), aggregate investment of $\bar{I}$ would not be feasible. Then, I^* is simply the unique solution to (11), where uniqueness follows from the fact that $k\left(I\right)$ is strictly decreasing and continuous in I and $Ak\left(0\right) > 0$. Conversely, if firm assets are larger than $\bar{I}\frac{B/\Delta p}{1+r_0}$, it is feasible to have aggregate investment of $\bar{I}$. In this case, firms are indifferent between investing and the storage technology.

Recall finally from the main text that we suppose that in equilibrium all firms are incentivized not to shirk. A sufficient condition for this to be the case is that under shirking the return is negative even from the perspective of investors with the highest sustainability preferences, i.e., even when at $I = 0$ and using $\bar{w}$: $P(0) - c\left(\theta_m\right) < \frac{1+r_0-\bar{w}}{1-\Delta p}$. We note that this is implied by condition (5) iff $B/\Delta p < \frac{\Delta p}{1-\Delta p}(1 + r_0 - \bar{w})$. **Q.E.D.**

Proof of Proposition 2. Since aggregate investment is given by $I^* = \min\left\{\bar{I}, Ak^*\right\}$ (by Proposition 1) and total sustainable investment is given by (13), the ratio is given by (14). The condition on the short supply of sustainable capital requires that $\sigma^* < 1$, i.e., $K[1 - G\left(\Delta c\frac{k^*}{k^*-1}\right)]\frac{k^*}{k^*-1} < \min\left\{\bar{I}, Ak^*\right\}$, which is always satisfied if a sufficiently high fraction of investors, $G(0)$, cares only about financial returns. **Q.E.D.**

Proof of Corollary 1. Note that k^* is independent of the distribution $G(w)$, from which it is immediate that σ^* increases when $G(\Delta c\frac{k^*}{k^*-1})$ decreases from a First-Order Stochastic Dominance shift. The invariance of r_0 and Δr^* follows from (12). **Q.E.D.**

Proof of Corollary 3. Consider first the comparative statics of I^*, where when financial constraints do not bind in the aggregate we have

$$\frac{d\bar{I}}{d\theta_m} = \frac{c'(\theta_m)}{P'(\bar{I})} < 0,$$

while otherwise we have from implicit differentiation of $I^* - Ak(I^*) = 0$

$$\frac{dI^*}{d\theta_m} = -\frac{A(1 + r_0)c'(\theta_m)}{[B/\Delta p - [P(I^*) - c\left(\theta_m\right) - (1 + r_0)]]^2 - AP'(I^*)(1 + r_0)} < 0.$$

We note also that in the latter case $P(I^*) - c(\theta_m)$ decreases in θ_m (and with it k^*). We next show that

$$\frac{dI_s^*}{d\theta_m} = K \frac{d\left[[1 - G(\Delta r^*)] \frac{k^*}{k^*-1} \right]}{d\theta_m} > 0, \tag{19}$$

for which we note first that $\gamma = \frac{k^*}{k^*-1} = \frac{1+r_0}{P(I^*)-c(\theta_m)-B/\Delta p}$ is strictly increasing by monotonicity of $P(I^*) - c(\theta_m)$. The comparative result for I_s^* thus follows if, in addition, Δr^* decreases (such that $1 - G(\Delta r^*)$ increases). For this we rewrite

$$\Delta r^* = (1 + r_0) \frac{c(\theta_s) - c(\theta_m)}{P(I^*) - c(\theta_m) - B/\Delta p} \tag{20}$$

to obtain

$$\frac{d\Delta r^*}{d\theta_m} = -(1 + r_0) \frac{c'(\theta_m)(P(I^*) - B/\Delta p - c(\theta_s)) + \frac{dI^*}{d\theta_m} P'(I^*)(c(\theta_s) - c(\theta_m))}{[P(I^*) - c(\theta_m) - B/\Delta p]^2}.$$

To see that this is strictly negative note, first, that $P(I^*) > B/\Delta p + c(\theta_s)$ (cf. (IC)) and that, second, $c(\theta_s) > c(\theta_m)$, $\frac{dI^*}{d\theta_m} < 0$, and $P' < 0$. **Q.E.D.**

Proof of Proposition 3. Recall that the social planner's objective function reduces to that of maximizing

$$I_s^* \left[(\theta_s - \theta_m) - (c(\theta_s) - c(\theta_m)) \right]. \tag{21}$$

Now the first part of the claim follows immediately from the observation that the derivative of the second term in (21) is strictly positive for all $\theta_s < \theta_{int}$ and strictly negative for all $\theta_s > \theta_m$. When now $\theta_m < \theta_{int}$, the first-order condition for θ_s is given by

$$\frac{dI_s^*}{d\theta_s} \left[(\theta_s^* - \theta_m) - (c(\theta_s^*) - c(\theta_m)) \right] + I_s^* \left[1 - c'(\theta_s^*) \right] = 0 \tag{22}$$

with

$$\begin{aligned} \frac{dI_s^*}{d\theta_s} &= -K \frac{k^*}{k^* - 1} g(\Delta r^*) \frac{d\Delta r^*}{d\theta_s} \\ &= -K \left(\frac{k^*}{k^* - 1} \right)^2 g(\Delta r^*) c'(\theta_s) < 0. \end{aligned}$$

As long as $\theta_m < \theta_s < \theta_{int}$, the factor that is multiplied with $\frac{dI_s^*}{d\theta_s}$ is strictly positive, indicating that for all such levels an increase in I_s^* is strictly welfare optimal. This formalizes the trade-off described in the main text, from which $\theta_s^* < \theta_{int}$ holds strictly.

We finally turn to the comparative statics in θ_s^* when $\theta_m < \theta_{int}$. Substituting for I_s^* and $\frac{dI_s^*}{d\theta_s}$ into (22) and rearranging terms, we have that

$$\frac{g(\Delta r^*)}{1 - G(\Delta r^*)} = \frac{1 - c'(\theta_s^*)}{c'(\theta_s^*)} \frac{k^* - 1}{k^*} \frac{1}{(\theta_s^* - \theta_m) - (c(\theta_s^*) - c(\theta_m))}. \tag{23}$$

Note now that the right-hand side of (23) is strictly decreasing in θ_s^* as this holds for $1/c'(\theta_s^*)$ due to convexity of the cost function and as the denominator in the last term strictly increases from $1 - c'(\theta_s^*) > 0$ at the optimum (given $\theta_s^* < \theta_{int}$). As for any given Δr^* the left-hand side strictly decreases with a Monotone Likelihood Ratio Property shift in $G(w)$, which implies strict monotonicity of the hazard rate in the interior of the support, θ_s^* must indeed strictly increase. **Q.E.D.**

Proof of Proposition 4. When we ignore financial constraints in the aggregate, θ_m^* solves (17). If at this level of the minimum standard and respective $I^* = \bar{I}$ it holds that $A \geq \bar{I}\frac{B/\Delta p}{1+r_0}$, this indeed characterizes the optimal policy. For smaller values of A, it then follows equally that under the then optimal level of θ_m^* financial constraints must bind in the aggregate, with I^* given by the fixed-point problem (11) and the first-order condition for θ_m^* becoming

$$\frac{dI^*}{d\theta_m}\left[P(I^*) - \left(c\left(\theta_m^*\right) + (\bar{\phi} - \theta_m^*) + (1 + r_0)\right)\right] - I^*[1 - c'(\theta_m^*)] = 0. \tag{24}$$

Observe first that also then $\frac{dI^*}{d\theta_m} < 0$, as the higher costs reduce maximum leverage for each firm, and that, in this case, each firm levers up maximally. Precisely, from implicit differentiation of $I^* = Ak\left(I^*\right)$ we obtain

$$\frac{dI^*}{d\theta_m} = -\frac{c'(\theta_m)}{\left(B/\Delta p - \left[P(I) - c\left(\theta_m\right) - (1 + r_0)\right]\right)^2 - P'(I^*)A} < 0.$$

We note that the comparative statics of θ_m^* in A depend on higher-order derivatives of the functions P and c. To prove the claim in the Proposition, it is however sufficient to consider the case where $A \to 0$ and thus $I^* \to 0$. In this case the sign of the derivative $\frac{d\Omega}{d\theta_m}$ is determined by (the negative of) the expression $P(I^*) - \left(c\left(\theta_m\right) + (\bar{\phi} - \theta_m) + (1 + r_0)\right)$, which at $I^* = 0$ and $\theta_m = \theta_{int}$ is strictly positive from (2): When A is low and when $\theta_m = \theta_{int}$, the marginal social value of the considered economic activity is strictly positive. Turning back to the first-order-condition (24), the resulting optimal minimum standard must then satisfy $\theta_m^* < \theta_{int}$. **Q.E.D.**

Proof of Proposition 5. We can first rely on the argument in the proof of Proposition 3, from which $I_s^* = 0$ is uniquely optimal in case $\theta_m \geq \theta_{int}$, as then $\theta_s - \theta_m < c(\theta_s) - c(\theta_m)$, while when $\theta_m < \theta_{int}$, it is strictly optimal to choose a standard $\theta_m < \theta_s^* < \theta_{int}$ and with it $I_s^* > 0$. Also, for the jointly optimal choice of (θ_m^*, θ_s^*), as I^* does not depend on θ_s, still the first-order condition (22) applies. We thus turn finally to the respective condition for

θ_m^*. Here, by the preceding observations we can focus on the case where the economy is constrained in the aggregate and thus $I^* < \bar{I}$. Adjusting (24) in the proof of Proposition 4 to account for the presence of more sustainable investment I_s^*, the first-order condition changes to

$$\begin{aligned}\frac{d\Omega}{d\theta_m} &= \frac{dI^*}{d\theta_m}\left[P(I^*) - (1+r_0) + (\theta_m^* - c(\theta_m^*))\right] + (I^* - I_s^*)[1 - c'(\theta_m^*)] \\ &\quad + \frac{dI_s^*}{d\theta_m}\left[(\theta_s^* - \theta_m^*) - (c(\theta_s^*) - c(\theta_m^*))\right] \\ &= 0.\end{aligned}$$

Recall here the derivation of $\frac{dI_s^*}{d\theta_m} > 0$ following (19). From this and $\theta_m < \theta_s^* < \theta_{int}$ the additional term in the second line, compared to expression (24), is strictly positive. When Ω is strictly quasiconcave this implies that in the presence of a sustainable investment standard θ_s, the optimal minimum standard increases. **Q.E.D.**

Proof of Proposition 7. Holding now the tax fixed, the social planner's objective is to maximize (21), where θ_m is obtained from $\widetilde{c}'(\theta_m) = 0$. This implies that optimally $I_s^* = 0$ when $\tau = 1$ (or even larger). Otherwise, the first-order condition is given by (22), with the only difference that now

$$\frac{dI_s^*}{d\theta_s} = -K\left(\frac{k^*}{k^*-1}\right)^2 g\left(\Delta r^*\right)\widetilde{c}'(\theta_s) < 0 \tag{25}$$

when $\theta_s < \theta_m$. With $\theta_m < \theta_{int}$ it is then again strictly optimal to set $\theta_m < \theta_s^* < \theta_{int}$ and with this $I_s^* > 0$. **Q.E.D.**

Proof of Proposition 8. The proof is analogous to that of Proposition 5. For the jointly optimal choice of (τ^*, θ_s^*), in case of binding financial constraints in the aggregate, recall that I^* does not depend on θ_s, so that also here the first-order condition (22) applies (with $\frac{dI_s^*}{d\theta_s}$ given by (25)). This leaves us with the determination of τ^* and thus θ_m^*. Here we need to take account of the additional financial burden imposed by a higher tax. For this it is useful to consider the total effect on I^* by writing $\frac{dI^*}{d\tau} = \frac{dI^*}{d\widetilde{c}}\frac{d\widetilde{c}}{d\tau}$, where $\frac{d\widetilde{c}}{d\tau} = \bar{\phi}$ and where $\frac{dI^*}{d\widetilde{c}} < 0$ follows immediately from $I^* - Ak\left(I^*\right)$. It is now instructive to first set $I_s^* = 0$, so that

$$\frac{d\Omega}{d\tau} = \frac{dI^*}{d\tau}\left[P(I^*) - c\left(\theta_m\right) - (\bar{\phi} - \theta_m) - (1+r_0)\right] + \frac{d\theta_m}{d\tau}I^*[1 - c'(\theta_m)]. \tag{26}$$

When the economy is not financially constrained, $\tau = 1$ results in both $P(I^*) - c\left(\theta_m\right) - (\bar{\phi} - \theta_m) - (1+r_0) = 0$ and $1 - c'(\theta_m) = 0$, which uniquely satisfies the first-order condition. With financial constraints and $\tau = 1$, the external margin is strictly positive, $P(I^*) - c(\theta_m) - (\bar{\phi} - \theta_m) - (1 + r_0) > 0$, from which we have $c'(\theta_m^*) > 1$ and thus $\theta_m^* < \theta_{int}$ and $\tau^* < 1$ (as claimed in the main text). Now, when we introduce θ_s, the derivative of Ω in (27) expands by

$$\frac{dI_s^*}{d\tau}\left[(\theta_s^* - \theta_m) - (c(\theta_s^*) - c(\theta_m))\right] - I_s^* \frac{d\theta_m}{d\tau}[1 - c'(\theta_m)], \tag{27}$$

where again $\frac{dI_s^*}{d\tau} = \frac{dI_s^*}{d\widetilde{c}} \frac{d\widetilde{c}}{d\tau} > 0$ (using $\widetilde{c} = \widetilde{c}(\theta_m)$)) with, in analogy to the derivation following (19), now $\frac{dI_s^*}{d\widetilde{c}} > 0$. Formally, the latter holds as k^* and $\gamma = \frac{k^*}{k^*-1}$ strictly decrease in $\widetilde{c}$, $P(I^*) - \widetilde{c}$ strictly decreases (again taking into account the dependency of I^*), and Δr^* strictly increases from

$$\frac{d\Delta r^*}{d\widetilde{c}} = (1 + r_0) \frac{(P(I^*) - B/\Delta p - \widetilde{c}(\theta_m)) - P'(I^*)\left(\widetilde{c}(\theta_s) - \widetilde{c}(\theta_m)\right)}{\left[P(I^*) - \widetilde{c}(\theta_m) - B/\Delta p\right]^2}.$$

Thus, the term (27) is again strictly positive when evaluated at $\theta_m < \theta_s^* < \theta_{int}$, so that $\tau^* < 1$ is indeed strictly higher when a sustainable classification is introduced. **Q.E.D.**

Proof of Proposition 9. Recall that when optimizing over θ_s, for given θ_m, the respective objective function reduces to (21). To prove the claim, using strict quasiconcavity, we only need to show that with the additional effect, ceteris paribus $\frac{dI_s^*}{d\theta_s} < 0$ is strictly lower (thus strictly higher in absolute value). To see this, note that k^* is independent of θ_s and that we obtain from (18)

$$\frac{dI_s^*}{d\theta_s} = -K \frac{k^*}{k^* - 1} \frac{d\Delta r^*}{d\theta_s} \frac{dG\left(\Delta r^* \mid \Delta r^*\right)]}{d\Delta r^*}$$

with now

$$\frac{dG\left(\Delta r^* \mid \Delta r^*\right)}{d\Delta r^*} = g\left(\Delta r^* \mid \Delta r^*\right) + G_2\left(\Delta r^* \mid \Delta r^*\right).$$

The assertion follows from $G_2 > 0$, i.e., as by assumption a decrease in $w^* = \Delta r^*$ (and with it an increase in I_s^*) leads to a FOSD shift. **Q.E.D.**