Roman Inderst | Markus Opp

Socially Optimal Sustainability Standards with Non-Consequentialist ("Warm Glow") Investors

Leibniz Institute for Financial Research SAFE

Sustainable Architecture for Finance in Europe

Socially Optimal Sustainability Standards with Non-Consequentialist ("Warm Glow") Investors

Roman Inderst[*]   Markus Opp[†]

February 2022

Abstract

Agencies around the world are in the process of developing taxonomies and standards for sustainable (or ESG) investment products. A key assumption in our model is that of non-consequentialist private investors (households) who derive a "warm glow" decisional utility when purchasing an investment product that is labelled as sustainable. We ask when such labelling is socially beneficial even when the social planner can impose a minimum standard on investment and production. In a model of financial constraints (Holmström and Tirole 1997), which we close to include consumer surplus, we also determine the optimal labelling threshold and show how its stringency is affected by determinants such as the prevalence of warm-glow investor preferences, the presence of social network effects, or the relevance of financial constraints at the industry level.

Keywords: Sustainability; ESG; green financing; labelling;

---

[*]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 have 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, possibly even tied to bank supervision or monetary policy, the declared key purpose is to increase transparency for investors and thereby channel funding to more sustainable firms and projects.

In this paper, we put at center stage final investors (households) with possible preferences for more sustainable investment.[2] Though the evidence on this is still mixed (cf. below), we assume that these preferences make them accept a lower return for such sustainable investments. Another key assumption is that households' preferences are non-consequentialist, on which we comment in detail below. For each Euro invested, a household thus derives a (decisional) warm-glow utility that is dependent only on the respective labelling but not on the actual level of sustainability that the respective firms exhibit. In such an environment we analyze a social planner's optimal choice of the respective threshold that defines an investment as sustainable and we also ask whether such an instrument is valuable even when the social planner could set a minimum sustainability standard that all investment and production must satisfy. Our analysis is set in a corporate finance model where firms are subject to agency problems when they tap into external funding, for which we employ the workhorse model of Holmström and Tirole (1997). Our model is closed on the product market side, which generates additional effects.

One message of our analysis is that the socially optimal labelling standard, i.e., the sustainability threshold that must be met so that an investment can be marketed as sustainable, should be lower than what could be deduced from a simple cost-benefit trade-off. And it depends, amongst other determinants, on the prevalence and strength of investors' warm-glow preferences. Our main analysis thereby focuses on the long-term view where firms' level of sustainability is endogenous. Firms can choose a higher sustainability level by incurring at the same time higher per-unit costs of production. If warm-glow 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).

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

is (still) in relatively short supply, it does not affect the external margin of the industry, but the share of investment and production of more sustainable firms. The per-unit cost disadvantage of these firms is just compensated by the equilibrium differential in outside investors' required rate of return. Given the downward sloping supply curve of warm-glow capital, the social planner faces a trade-off when choosing the sustainability threshold that defines the respective label. As a consequence, the optimal threshold is strictly lower than what would be socially optimal otherwise, and it should increase as investors' preferences shift upwards (with respect to sustainability). When investors' warm-glow preferences depend, in addition, on social norms, which are shaped by the observed behavior of others, this further reduces the optimal threshold, as the social planner wants to benefit from such a positive feedback effect. We also note that given firms' agency problem, also more sustainable firms with access to cheaper warm-glow capital can not fully divest (entrepreneurial) own funds (to seek higher return elsewhere).

We briefly consider also the case where firm sustainability is exogenous (in the short term). Though then the labelling of investments can not induce changes in firm sustainability, it still has real effects. The access to cheaper warm-glow capital allows firms with such a label to expand their investment and output. As this decreases the product price, less sustainable firms reduce their investment and output.

The labelling of sustainable investment improves the social planner's objective even when she can also set a minimum standard. Importantly, this is not so as such labelling would allow to increase suitability while the respective costs were born by investors who in turn obtained a higher utility. In our analysis, the social planner's objective excludes the (decisional) warm-glow utility, so that there is no such "free lunch". When the industry is financially constrained, the minimum standard is chosen not too high as this would excessively shrink the industry to the detriment of consumers. The joint of use of a higher threshold for the sustainability label allows to increase sustainability at least for a fraction of firms without negatively affecting the industry's external margin.

We now defend our assumption on households' warm-glow preferences. Survey evidence seems to confirm that some investors exhibit moral preferences with regards to investments, at least in their statements (Riedl and Smeets 2017). There is also growing market evidence. For instance, 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. We acknowledge that there is also evidence of positive abnormal returns, e.g., Gibson and Krueger (2018) and Henke

3

(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). The literature has also studied implications of such preferences.[3] Heinkel et al. (2001) show that when a firm is boycotted by such moral investors, the higher concentration of other shareholders' holdings can make funding more expensive. In our model, the direct implications relate to firms who tap into such cheaper funding, albeit through the product market there is also a negative funding effect on less sustainable firms.

The theoretical literature has mainly focussed on consequentialist (institutional) investors, as in Chowdhry et al. (2014) and Oehmke and Opp (2020).[4] There, in contracting with the firm, consequentialist investors not only fully take into account the sustainability impact of an individual firm, but also the counterfactual (if, e.g., a given change is not made). In our model, households are non-consequentialist. The key implication of this is that their willingness-to-pay for sustainability (i.e., to accept a lower return) relates to the presence of the respective label, but not to the underlying stringency and thereby the actual sustainability of the respective firms. While such an assumption could also be motivated by lack of information or the excessive complexity associated with such an assessment, studies that elicit preferences for so-called non-use values, including for sustainability, confirm even for simpler issues a tendency towards such non-consequentialism. For example, a large literature in environmental and resource economic has pointed out that elicited willingness-to-pay for a sustainability attribute may fail a so-called "scope test" or "adding-up test": Subjects' willingness-to-pay is relatively insensitive to the actual size and impact of the respective scenario change. In our context, individuals may exhibit a high willingness-to-pay even when the environmental impact is miniscule, whereas their willingness-to-pay does not increase substantially when the impact, i.e., the "scope", is much larger. With respect to contingent valuation analysis, prominent studies that document these effects are Boyle et al. (1994) and more recently Desvousges et al. (2012). In other words, 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. In our model we treat this as a mere decisional utility that, other than the real return experienced by investors, does not enter the (consequentialist) social planner's

---

[3]Landier and Lovo (2020) analyze when sustainable institutional investors can have an impact even though final investors do not share such preferences.

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

welfare function. In an extension, we also allow for the warm-glow effect to depend on social norms, which gives rise to a feedback effect.

Finally, we note that there exists a large theoretical literature on (optimal) certification, some of which also uses a binary labelling framework (see Bizzotto and Harstad 2021 for a recent contribution and a detailed survey). This literature focuses, however, on consumer willingness-to-pay and mainly takes the perspective of profit-maximizing certifiers (or that of industry associations).

In Section 2 we introduce the modelling framework. Section 3 analyzes the market equilibrium, while Sections 4 and 6 solve for the optimal policy. Section 5 introduces social preferences. Section 7 considers the short-term case with fixed sustainability types. Section 8 concludes. Proofs are relegated to the Appendix.

2 Model Framework

2.1 Set-up

We consider a single sector (or industry) of the economy. For this we model the technology choice and investment decision of a continuum of firms that produce a homogeneous output. Firms (owner-managers) are endowed with internal funds that are essential to mitigate a moral-hazard problem vis-a-vis outside investors. Outside investors are households that have otherwise access to an alternative investment opportunity and that have heterogeneous "warm-glow" 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. Firms with more sustainable investment and production can not secure a premium on the output market, which allows to focus exclusively on the financing side. Investors' warm-glow preferences depend on whether a firm has been labelled sustainable or not, and the design of the underlying threshold is our main policy variable. We refer to the Introduction for a motivation of this binary structure and investors' respective (non-consequentialist) preferences. We also do not consider a public tax or subsidy that would depend more flexibly on the sustainability dimension, as the use of such instruments is indeed only very restrictive, notably outside the case of environmental pollutants.

We next describe more formally the various ingredients of the model. We then define and characterize an 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 with warm-glow preferences.

Firms and industry. 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).[5] We index firms by i. Their investment technology is described below. Individual investment size is denoted by I_i so that total 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 also through a required rate of return r_i.[6] 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 it is monitored by the owner-manager. If she unobservably shirks, she obtains a private benefit per unit of investment $b > 0$, but with probability $q > 0$ the technology fails and returns no output. Output in terms of produced quantity is denoted by x_i, where for simplicity we set $x_i = I_i$ in case of success, while $x_i = 0$ holds in case of failure. Firms' output is sold in the same market and thus aggregated by $x = \int_0^1 x_i di$, with the price given by the inverse demand $P(x)$ with $P' < 0$.[7]

Sustainability and supply of "warm-glow" financing. The degree of sustainability of investment and output is captured by a single, real-valued measure θ. As we focus on a longer-term perspective, the sustainability level is endogenous and thus determined by each firm's choice. When a firm chooses the level θ_i this comes with per-unit costs $c(\theta_i)$, where $c' > 0$ and $c'' > 0$. Government policy describes a threshold θ_s so that firms with $\theta_i \geq \theta_s$ can offer to outside investors sustainable investment opportunities. Subsequently, we enrich the social planner's instruments to include a minimum standard θ_m. Presently, this is fixed. We specify below how θ enters the social planner's objective function.

We suppose that there is a sufficiently large potential supply of funds from atomistic households. We denote the aggregate by K. In line with our partial (one industry) analysis,

---

[5]For our analysis it is inconsequential whether these are going concerns with internal funds or whether these are potential entrepreneurs with respective own wealth.

[6]We note that we need not distinguish between different rates of returns promised to different classes of investors (according to their preferences).

[7]As each firm has mass zero, its own output does not affect the market price so that there are no incentives to withhold output. Recall that we fully abstract from potential differentiation on the product market.

Electronic copy available at: https://ssrn.com/abstract=4065630

households have access to an alternative investment opportunity that generates the fixed return r_0. We now bring in the warm-glow effect. An individual household, say j, derives from each unit of sustainable investment that she makes the warm-glow effect w_j (which is thus equivalent to an additional per-unit return). In the aggregate, households' preferences are characterized by the CDF $G(w)$ over $[0, \overline{w}]$, where we suppose that $G(0) > 0$ is large (so that a large fraction of households does not experience such a warm glow). We note again that with this specification households' preferences and with it the required rate of return are not dependent on the actual sustainability level of a given firm, but on the presence or absence of the binary sustainability label.

Timing. We consider the following timing of events. In $t = 0$ the social planner chooses the threshold θ_s. In $t = 1$ the for now initially homogeneous firms simultaneously choose their sustainability level θ_i. In $t = 2$ firms choose outside financing L_i and total investment I_i.[8] At this stage also households allocate their funds. Together this will give rise to an equilibrium funding condition for each firm, captured by r_i. In $t = 3$, at marginal cost c_i firms produce output, which is $x_i = I_i$ if there is no shirking, and realize a product price $P(I)$. At this stage they also honor their promises D_i to outside investors.

2.2 Firm financial constraints

We will invoke parameter restrictions so that we can safely restrict consideration to equilibria where owner-managers do not shirk. We denote aggregate equilibrium investment by I, which from $x = I$ pins down the equilibrium product price $P(I)$. The owner-manager realizes $I_i[P(I) - c_i] - D_i$ if the firm is successful, where c_i depends on the chosen sustainability level θ_i. In case of shirking, the firm fails with probability q but she realizes private benefits bI_i. Incentive compatibility thus requires that $I_i[P(I) - c_i] - D_i$ exceeds the sum of $(1 - q)[I_i[P(I) - c_i] - D_i]$ and bI_i, which transforms to

$$D_i \leq I_i \left[P(I) - c_i - b/q \right]. \tag{1}$$

This makes transparent the restricted external financing capacity. We note again that we need not distinguish between different classes of investors and their required rate of return, as in our model firms are of negligible size compared to the capital market and as

---

[8]We note that as we consider a continuum of ("mass zero") agents, there will not be strategic interactions. It would thus be inconsequential to collapse $t = 1$ and $t = 2$ into a single period.

households allocate funds based only on the promised return and the firm's sustainability label.[9]

We next set up outside investors' break-even constraint. Investors require the return r_i, which will be endogenized below. Under the expectation that the owner-manager does not shirk, investors thus just break even when they are promised $D_i = L_i(1 + r_i)$. Substituting this into (1) and noting that $L_i = I_i - A_i$, the incentive constraint transforms to

$$(I_i - A)(1 + r_i) \leq I_i \left[P(I) - c_i - b/q \right],$$

which imposes the following constraint on the firm's total investment and output:

$$I_i \leq \frac{1 + r_i}{b/q - [P(I) - c_i - (1 + r_i)]} A_i. \tag{2}$$

It is useful to denote the respective scaling factor on the right-hand side by

$$f_i = \frac{1 + r_i}{b/q - [P(I) - c_i - (1 + r_i)]}.$$

If the denominator on the right-hand is positive, which will always be the case in equilibrium, the size of firms' internal funds constrains investment. This does not yet imply that the industry is financially constrained in the aggregate. We return to this issue below. Also, it will depend on parameters whether a firm strictly prefers to leverage up to (2).

If the owner-manager invests all of her own funds A_i, her expected payoff is given by $U_i = I_i \left[P(I^*) - c_i \right] - D_i$, which after substitution of the break-even condition becomes

$$U_i = I_i \left[P(I^*) - c_i \right] - (I_i - A_i)(1 + r_i). \tag{3}$$

Parameter restrictions. Note again that presently the minimum sustainability standard is taken as given. The subsequent analysis is restricted to the case where even without warm-glow capital, production is profitable at this standard, as it holds for the respective cost $c_{ns} = c(\theta_m)$ that $P(0) > c_{ns} + (1 + r_0)$. Individual firms are financially constrained by the agency problem, as in Holmström and Tirole (1997). A sufficient condition is that $P(0) - c_{ns} - (1 + r_0) < b/q$.[10] Our final parameter restriction relates to the importance of monitoring. After all, a financial contract with outside investors could also neglect the

---

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

[10]We note that this is only a sufficient condition, as it is evaluated at the marginal profit from an investment when the industry size is zero, $I = 0$, so that the product price $P(I)$ is highest.

owner-manager's incentive constraint, in which case she shirks. To simplify the exposition, we stipulate that the project would then realize negative financial return, even at the highest possible price: $(1-q)(P(0) - c_{ns}) < 1$.

3 Market equilibrium

In equilibrium, firms and households make optimal decisions. Recall that for households only the distinction whether a firm is labelled sustainable or not is relevant. Therefore, we need to distinguish between only two rates of return. The required rate of return for investments into non-sustainable firms, r_{ns}, is equal to that obtained from the alternative investment opportunity, $r_{ns} = r_0$. We denote the required rate of return for sustainable investment by r_s. Our first equilibrium condition ("*investor optimality*") is that a household with warm-glow preferences $w_j \geq 0$ chooses to invest in a sustainable firm if $w + r_s \geq r_{ns}$ (and strictly prefers so if this holds strictly). This determines the supply of outside funds.

We now turn to our second equilibrium condition ("*firm optimality*"). An individual firm cannot affect market conditions of outside financing, r_{ns} and r_s, and this applies also to the product price $P(I)$. We take first the choice of a firm's sustainability level θ_i. Given the binary nature of the return required by outside investors and as $c(\theta_i)$ is strictly increasing, a firm will optimally choose either the minimum standard θ_m or θ_s. We already denoted the respective per-unit costs by $c_{ns} = c(\theta_m)$ and now denote also $c_s = c(\theta_s)$. The firm makes this choice and that of L_i and thus I_i so as to maximize U_i in (3) subject to the outside financing (incentive) constraint (2).[11]

Our third equilibrium condition relates to the market for warm-glow funding ("*warm-glow market clearing*"). We aggregate all L_i of sustainable firms to $L_s = \int_{\theta_i \geq \theta_s} L_i di$. Recall now that total household funds are given by K and that warm-glow preferences are distributed according to $H(w)$. From household optimality, i.e., the cutoff criterion $w + r_s \geq r_0$ where we used $r_{ns} = r_0$, the supply of respective funds is thus $K[1 - H(w^*)]$ with $w^* = r_0 - r_s$. The market-clearing condition is then

$$L_s = K[1 - H(w^*)].$$

Note that the marginal (warm glow) investor type w^* represents also the difference in investors' (equilibrium) required return for sustainable firms.

---

[11]As noted previously, we can shortcut our exposition by supposing that an active firm always invests all of its internal funds. We confirm this below.

In what follows, we need to distinguish between two cases. In the first case the industry is financially constrained in two respects: First, the industry is constrained in the aggregate; second, also the supply of warm-glow capital is sufficiently constrained. The first condition holds when internal funds A are not too large. The second condition holds when, compared to the size of the considered industry, investors with warm-glow preferences are sufficiently scarce. We provide precise conditions below. We first solve for this financially constrained case. Subsequently, we turn to the case where the industry is no longer constrained in the aggregate.

Characterization. When warm-glow capital is in short supply, not all investment in the industry can be financed by it. Hence, in equilibrium both firms with characteristics (θ_m, c_{ns}) and firms with characteristics (θ_s, c_s) are active. As presently the industry is financially constrained in the aggregate, each firm levers up maximally.

With maximum leverage, we now determine the total payoff U_i realized by the owner-manager of firm i with marginal costs c_i and a return r_i as required by outside investors. For this we substitute the binding constraint (2) into the payoff (3) to obtain[12]

$$U_i = A_i \frac{b}{q} f_i.$$

The presence of both sustainable and non-sustainable firms requires that firms are indifferent between the sustainable and non-sustainable choice. This is the case if the respective values for U_i are equal. With a slight abuse of notation, we refer to the respective scaling factors as f_s and f_{ns}, so that $U_s = U_{ns}$ implies that $f_s = f_{ns} = f$. Substituting now $r_i = r_0$ when $c_i = c_{ns}$ and $r_i = r_s$ when $c_i = c_s$ and using furthermore $w^* = r_0 - r_s$, we obtain from this for the interest rate differential (and the marginal warm-glow investor type) the equilibrium requirement[13]

$$w^* = r_0 - r_s = (c_s - c_{ns}) \frac{1 + r_0}{P(I^*) - c_{ns} - b/q}. \tag{4}$$

Note that for future reference we use I^* for total industry size in equilibrium. Hence, in the financially constrained case the equilibrium interest rate for the supply of warm-glow capital is determined by the requirement that the return on firms' internal funds is the same for the sustainable and the non-sustainable choice. Intuitively, w^* as given in (4) is

---

[12]Proceeding stepwise, we have with $D_i = I_i[P(I) - c_i - b/q]$ and by substituting the binding (2) $U_i = I_i(P(I) - c_i) - D_i = \frac{b}{q} I_i = \frac{b}{q} f_i A_i$.

[13]Inspection of $f_s = f_{ns}$ shows that this holds if $\frac{P(I) - c_s - b/q}{1 + r_s} = \frac{P(I) - c_{ns} - b/q}{1 + r_0}$, which transforms to (4).

strictly increasing in the per-unit cost differential $c_s - c_{ns}$, where $c_s = c(\theta_s)$ increases with the threshold that must be met so as to qualify as a sustainable investment.

As from firm indifference the respective scaling factors in f_s and f_{ns} must be the same, total industry investment can be obtained simply by using the scaling factor f_{ns} for a non-sustainable firm, so that $I^* = f_{ns}A$. After again writing out f_{ns}, this becomes

$$I^* = \frac{1 + r_0}{b/q - [P(I^*) - c_{ns} - (1 + r_0)]}A. \tag{5}$$

Equation (5) represents a fix point problem with a unique solution for I^* that depends only on primitives. We note that with given I^*, we have also determined w^* from (4) and with this r_s. The industry is indeed financially constrained in the aggregate when the marginal return at I^* is non-negative: $P(I^*) - c_{ns} - (1 + r_0) > 0$.

While this pins down aggregate investment and financing conditions, we still need to determine which fraction is accounted for by sustainable and which by non-sustainable firms. We denote the respective investments by I_s^* and I_{ns}^* (and also use the notation L_s^* and L_{ns}^* for equilibrium outside investment). Recall that the market clearing condition for warm-glow capital requires that $L_s^* = K[1 - H(w^*)]$ with $w^* = r_0 - r_s$. Using the break-even condition $D_i = L_i(1 + r_i)$, the incentive constraint $D_i \leq I_i[P(I^*) - c_i - b/q]$, and indifference for firms, we obtain (cf. the subsequent proof)

$$I_s^* = K[1 - H(w^*)]\frac{1 + r_0}{P(I^*) - c_{ns} - b/q}. \tag{6}$$

Warm-glow capital is indeed in short supply if the respective value satisfies $I_s^* < I^*$.

Proposition 1 *When both internal funds and warm glow capital are sufficiently scarce, the equilibrium is characterized as follows: Total industry size I^* is obtained uniquely from (5) and independendent of the sustainability label and warm-glow financing. Both sustainable and non-sustainable firms are present, where the interest rate differential of sustainable funding $w^* = r_0 - r_s$ is given by (4). Sustainable investment I_s^* makes up the fraction (6). This case applies when, at these values, it holds that $P(I^*) - c_{ns} - (1 + r_0) > 0$ and $I_s^* < I^*$.*

Below we make use of this characterization to determine the optimal labelling policy. For now, we obtain the following immediate comparative analysis of (4) and (6), on which we comment next:

11

Corollary 1 *In the financially constrained case, as warm-glow preferences increase (by way of a First-Order Stochastic Dominance shift in $H(w)$), the share of sustainable firms and that of sustainable investment increase, but the respective funding conditions r_s and with it the interest rate differential w^* remain unchanged. As the threshold θ_s for the sustainability standard increases, the share of sustainable firms and investment decreases and the interest rate differential w^* increases.*

These comparative statics results hinge crucially on the endogeneity of firms' sustainability types. Firm optimality alone then pins down the required interest rate differential w^*, irrespective of the availability of warm-glow financing. A corresponding change in investor preferences affects, however, the share of firms and with it that of investment that become sustainable. Decreasing the sustainability threshold that is necessary to obtain the respective label pushes down the equilibrium interest rate demanded by warm-glow investors, r_s, while shrinking the share of such investment. We note that this is 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. The change in the prevailing interest rate (differential) is an equilibrium effect as it just compensates firms for the higher cost of achieving the required sustainability standard. We finally note that in our longer-term perspective, where firms' sustainability is endogenous, firm profits are unaffected by a change in the threshold θ_s. This is an immediate consequence of the fact that in equilibrium firms are indifferent between becoming sustainable or not.

4 Optimal policy

To determine the social planner's optimal choice of θ_s we need to pin down the objective function. Integrating over all firms' choices, this is given by

$$\Omega = \int_0^{I^*} P(I)dI - \int I_i[c_i - \theta_i - (1 + r_0)]di. \tag{7}$$

We note that it is for our insights immaterial whether θ_i enters directly or, instead, with some multiplier. The social planner has consequentialist preferences, and the warm glow experienced by some households does not enter the objective function. We now comment on this choice. In the Introduction we described the warm-glow effect that households may experience when buying investment products that are labelled sustainable and referred to it as a "decision utility", and we already noted that we do not take this into account

in the social planner's objective. We also note that while θ_i may capture the impact on citizen welfare, the social planner's objective need not coincide with aggregate welfare of, in particular, the current cohort of citizen. This is the case, in particular, when the political process has lead to commitments that must be honored.[14]

Note next that presently, in the financially constrained case, total investment and output is independent of θ_s. Consequently, we can focus on the (partial) objective function $I_s^*(\theta_s - c_s) + (I^* - I_s^*)(\theta_m - c_{ns})$, so that the social planner's problem is equivalent to maximizing

$$I_s^* \left[(\theta_s - \theta_m) - (c_s - c_{ns}) \right]. \tag{8}$$

For any given unit of investment and production, the social planner would want to implement the level of sustainability θ_s so that $c'(\theta_s) = 1$. Note here that we stipulate that the presently fixed minimum standard θ_m is sufficiently low (so that at this level it holds that $c'(\theta_m) < 1$). Recall however that the social planner can not dictate such a level of sustainability. The only tool that she presently has at her disposal is to set a sufficiently high standard of sustainability so that firms that want to access cheaper warm-glow financing have to satisfy this standard. With the constrained and downward sloping supply of warm-glow financing, compensating firms for the respective higher costs c_s when θ_s is raised necessarily implies that the interest rate differential w^* must increase. But then the equilibrium supply of sustainable funding and thus ultimately I_s^* must decrease. In short, the social planner faces a trade-off as I_s^* decreases when the standard θ_s is set higher. This trade-off becomes attenuated when the supply of warm-glow financing increases. For an unambiguous comparative analysis, in the subsequent Proposition, we consider an increase in $H(w)$ in the sense of the Monotone Likelihood Ratio Property (MLRP).

Proposition 2 *In the financially constrained case, the social planner sets a sustainability standard θ_s for the respective label that is strictly below the level that would optimally trade off per-unit costs $c(\theta)$ with sustainability benefits θ. As investor preferences shift upwards in the sense of a MLRP-shift in $H(w)$, the optimal threshold becomes strictly higher.*

Proposition 2 is our first normative result. The presently single instrument of the social planner, the labelling of sustainable investment, allows to hENWAA the warm-glow preferences of (some) households. This gives the respective firms with a chosen

---

[14]These may also derive from decisions of courts. Here, 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 commission.

higher sustainability level access to cheaper financing. In equilibrium, the interest rate differential just covers the respective cost difference. The higher the required differential w^*, the smaller will however be the set of households whose warm-glow preferences are sufficiently strong. Proposition 2 shows how the social planner must bear this downward sloping supply of warm-glow capital in mind when she wants to optimally harness it to incentivize sustainable investment. Rather than "challenging" the market with a high initial threshold, the social planner should optimally lower the threshold when warm-glow preferences are still less prevalent. Proposition 2 provides also a rationale for increasing the standard over time as such preferences become more widespread.

Sector not financially constrained in the aggregate. Before we proceed to the introduction of an additional policy instrument, the minimum standard θ_m, we briefly extend the analysis to the case where the industry is not financially constrained in the aggregate, as there are sufficient internal funds. Then, the equilibrium size of the industry I^* is pinned down by a condition of zero marginal return:

$$P(I^*) - c_{ns} = 1 + r_0. \tag{9}$$

Again, the choice of θ_s affects the internal margin, i.e., the fraction of sustainable investment I_s^*. In this respect, we can show that the positive (comparative) insights from Corollary 1 and the normative insights from Proposition 2 continue to apply.

An interesting additional insight is now that firms choosing the higher sustainability level and thereby enjoying cheaper financing will always lever up maximally. It is also only when they fully exploit their maximum leverage, that firm owners enjoy the same return on internal funds as non-sustainable firms, which is just r_0 in the presently analyzed case where the industry expands until (9) holds.

We note that as firm-owners do not share warm-glow preferences, without the agency problem they would not invest any of their own funds. Instead, taking a somewhat larger picture, they would try to divest all of their own funds, e.g., through paying out a large dividend, when turning more sustainable and thereby tapping into cheaper warm-glow funding. The agency problem in our model renders this impossible, thereby adding realism.

Proposition 3 *When the industry is not financially constrained in the aggregate, though individual firms are, the choice of θ_m affects the share of sustainable production, but not that of total industry size. The positive results from Corollary 1 and the normative results from Proposition 2 continue to hold.*

14

5 Social network effects

In this section, we briefly extend the analysis to social preferences in the following way. We now allow for a feedback mechanism as we suppose that individual (relative) 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 are essentially non-use benefits[15] and as such particularly susceptible to social norms, which in turn can be shaped by the anticipated or observed behavior of others. While the notion of such norm formation is widely spread in social sciences,[16] it is 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. In an early contribution, Robert Sugden hypothesizes that individuals follow a conditional moral rule of "*contributing of what I wish others to contribute, but not needing to contribute more than the person who contributes the least*".[17] Also various field experiments relate especially environmentally conscious behavior to that of others.[18] Endogenous preferences and changes in norms have also been incorporated in policy suggestions to fight climate change.[19] Our key notion is thus that households' investment preferences are not exogenously fixed but endogenous in that they depend on the behavior of others.[20]

In what follows, we thus assume that the distribution of warm-glow preferences depends on how widespread the respective investments are among households. For simplicity, we capture this via a dependency on the cutoff w^*, writing $H(w \mid w^*)$. Here we stipulate that

---

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

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

[17] Sudgen (1984).

[18] For instance, individual recycling behavior has been found to strongly correlate with beliefs about such behaviour in the community; see Cialdini (2003, 2005) or the various studies quoted in Schultz (2002). Other studies analyze the influence on more sustainable clothing (e.g., Kim et al. 2012).

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

[20] 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.

Electronic copy available at: https://ssrn.com/abstract=4065630

as w^* decreases, so that sustainable investments are more widespread, the distribution moves upwards in the sense of strict First-Order Stochastic Dominance.

Regarding our preceding derivations, we only need to modify expression (6) accordingly, which now reads

$$I_s^* = K[1 - H(w^* \mid w^*)]\frac{1 + r_0}{P(I^*) - c_{ns} - b/q}.$$

We are interested in how the normative implications change. Here, it is important that in $dI_s^*/d\theta_s$ we now need to take the total derivative of $H(w^* \mid w^*)$, where $H_2(w^* \mid w^*) > 0$. 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 the fraction $1 - H(w^* \mid w^*)$ 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 4 *Suppose households' warm-glow preferences are subject to the described social-norm feedback. Then it is optimal for the social planner to choose a strictly lower threshold for the sustainability label.*

6 Choosing simultaneously the minimum standard

So far we have taken the minimum standard θ_m as given. As we discussed in the Introduction, in many cases notably outside environmental sustainability there may not exist such a legally imposed requirement, or it may not be very stringent.[21]

We now consider applications where a social planner could set such a minimum standard. Our key question is whether in this case there is still scope for the envisaged suitability labelling. After all, the social planner could now directly target a particular value of $\theta_i = \theta_m$, irrespective of the supply of warm-glow capital.

A first response to this question could be that such labelling allows to increase suitability while the respective costs are born by investors, who in turn obtain a higher utility. We note, however, that our objective function takes into account the true cost of capital r_0, excluding the decisional warm-glow utility, so that there is no such "free lunch". Still, as we show, there is scope for a distinction between a minimum standard and a higher threshold for such labelling.

---

[21] A case in place is the recently enacted German supply chain legislation, which imposes compliance only with a set of fundamental rights and largely conditional on that the country in which a firm operates has adopted these rights in its legislation.

We first recall from a rewriting of (5) that total industry size is given by

$$I^* = \frac{1 + r_0}{b/q - [P(I^*) - c(\theta_m) - (1 + r_0)]} A,$$

which is strictly decreasing in θ_m. This is intuitive as with a higher minimum threshold the respective per-unit costs increase, which in the considered financially constrained case reduces maximum leverage. As the minimum threshold changes, this has two implications for the industry outcome. First, the sustainability per-unit of investment and production changes; second, the size of the industry changes as well. As we already observed, ceteris paribus the social planner would want to set θ_m so that marginal benefits equal marginal costs per unit of production. But this interacts now with industry size. Recall that this is different when the social planner uses, in addition, labelling and with it the threshold θ_s. Imposing such a higher labelling threshold is thus always beneficial if the social planner does not want to increase θ_m further, i.e., as the marginal social benefit from a larger industry size is positive. Whether the latter is the case depends on two parameters, as we discuss after stating the following Proposition:[22]

Proposition 5 *Suppose the social planner can set a minimum standard θ_m. If at the optimal minimum standard the industry is financially constrained from the perspective of the social planner's objective function, as $P(I^*) - (1 + r_0) + (\theta_m - c(\theta_m)) > 0$, it is optimal to introduce, in addition, a higher threshold θ_s and with it a sustainability label for investment.*

Proposition 6 thus provides a rationale for introducing sustainability labelling for investment products even when, first, the social planner can also set a binding minimum standard and, second, the realization of warm-glow preferences does not enter the social planner's objective. We now discuss briefly the condition when it is optimal to use both instruments. One determinant is arguably the relevance of financial constraints and with it firms' size of internal funds A. Ceteris paribus, a reduction of A reduces equilibrium investment and output and thereby increases the marginal social return of expanding the industry, as marginal consumer welfare $P(I^*)$ is then higher. The second determinant of the equilibrium marginal social surplus is the nature of θ. If production results in a negative externality irrespective of the chosen sustainability level, then θ is always negative, i.e., even at the minimum standard. For other sustainability goals, the opposite may be the case, for instance, if θ is a metric tied to the fairness of production.

---

[22]Given that each firm is financially constrained, the marginal profit is always strictly positive $P(I^*) - c_{ns} - (1 + r_0) > 0$.

17

7   A brief consideration of the short-term perspective

So far the choice of the sustainability type was fully endogenous, with firms being ex-ante homogenous. We interpreted this as the long-term view. In the shorter term, firms may however have intrinsic type differences. This may depend on their production technology or their location. We show now that even then, when firms' sustainability type is fixed but heterogeneous, the considered policy of sustainability labelling is effective.

We thus suppose now that firms' type θ_i is exogenous with CDF $G(\theta_i)$. Obviously, only types $\theta_i \geq \theta_m$ can be active. If there is no sustainability labelling, then all receive the same funding conditions and can lever up to the same investment size I_i. In what follows we focus on the financially constrained case, so that from (5) and with homogeneous c_0 we have for industry size

$$I^* = \frac{1 + r_0}{b/q - [P(I^*) - c_0 - (1 + r_0)]} A[1 - G(\theta_m)].$$

We now add a sustainability standard $\theta_s > \theta_m$. Again, this will make funding cheaper for types that satisfy this standard. While now this does not go hand-in-hand with a change of θ_i, this still has real effects as it increases the respective investment. Formally, we how have $I_i^* > I_j^*$ if $\theta_i \geq \theta_s$ but $\theta_j < \theta_s$.

Interestingly, this wedge in the size of investment and production between firms above and below the standard has two reasons. First, as already noted these firms have different funding conditions, which for more sustainable firms relaxes the financial constraint. Second, as funding conditions for more sustainable firms improve and their investment and production increase, ceteris paribus this increases total production and thereby reduces the product price $P(I^*)$. Even though the introduction of the sustainability standard does not affect directly the funding conditions for firms below the standard, their size of investment and production shrinks with a lower product price $P(I^*)$. In essence, through the cheaper funding conditions firms with a type above the sustainability label crowd out investment and production of firms with a lower type. We note that this effect was absent in the previously analyzed (long-term) case, as, with ex-ante identical firms, the advantage from the better funding conditions was exactly compensated by the higher (per-unit) costs that were necessary to meet the standard.

With such heterogeneity, we can finally write the objective function of the social planner as

$$\Omega = \int_0^{I^*} [P(I) - c_0 - (1 + r_0)] dI + \int_{\theta_m}^{\bar{\theta}_0} \theta_i I_i^* dG(\theta_i).$$

18

The choice of the standard has now a more complex effect on the objective function through the described effect on I^* and thus $P(I^*)$, which again affects the leverage and investment of each individual firm, I_i^*. Moreover, recall now from the preceding discussion of the minimum standard the discussion of whether, from the social planner's perspective, the overall size of the industry is too large or too small. Despite financial constraints, the industry may be too large due to non-internalized externalities, as expressed by negative values of θ. We note however that the additional use of an investment label only raises investment and production for firms above the threshold and leads to a respective decrease for all firms below the standard. Now that a higher sustainability standard is not associated with higher costs, as in the preceding (long-term) case, the social planner would always want to use such an instrument in addition to that of a minimum standard, as it involves a shift of production to higher-sustainability firms.

Proposition 6 *Consider now with ex-ante heterogeneous firms the (short-run) case where the sustainability type is fixed. Then, introducing a sustainability label still has real effects as it shifts investment and production to more sustainable firms, precisely as their investment increases and, through a reduction of the equilibrium product price, that of non-sustainable firms decreases. When also the minimum standard θ_m can be optimally chosen, it is always strictly preferable to have, in addition, a sustainability label with threshold $\theta_s > \theta_m$.*

8 Concluding remarks

In this paper, we introduce investors with warm-glow preferences into an equilibrium model of financially constrained firms. When an investment label is introduced, firms that meet the respective sustainability standard can tap into such cheaper financing.

In a longer-term perspective, where the sustainability standard is endogenous, more sustainable firms must incur the respective higher per-unit costs. When choosing the threshold applied to the label the social planner must bear in mind these additional costs and the required higher interest rate differential so that firms find it optimal to meet such a standard. With an elastic supply of warm-glow financing, this results in a trade-off, and in equilibrium, the fraction of firms that will choose the respective standard becomes smaller when the standard increases. While the respective investment standard only affects the distribution of more or less sustainable firms in the sector, instead the requirement of a minimum standard for all active firms affects total industry size, given that it raises

costs for all firms. When the marginal social surplus is positive, which is notably the case when financial constraints are sufficiently binding, this implies that the social planner would always want to use both a minimum standard and a higher standard for labelling sustainable investments.

We also showed that when firms can not change their sustainability type in the short run but are heterogeneous, a sustainability label can still be effective. While in this case it does not incentivize firms' strategic choice of higher sustainability, it eases up the financing constraints for more sustainable firms and, through a product market effect, reduces investment and output for less sustainable firms.

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.

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. 2005. Don't Throw in the Towel: Use Social Influence Research. American Psychological Society Observer 18, 33-34.

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.

Inderst, Roman, Felix Rhiel and Stefan Thomas. 2021. Sustainability Agreements and Social Norms.

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

Kim, Hanna, Eun-Jung Lee and Won-Moo Hur. 2012. The Normative Social Influence on Eco-Friendly Consumer Behavior: The Moderating Effect of Environmental Marketing Claims. Clothing and Textiles Research Journal 30, 4-18.

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.

22

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

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

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.

9 Appendix: Omitted proofs

Proof of Proposition 1. Regarding the fix point problem in (5), we note first that all other parameters but I^* are exogenous. Next, the left-hand side is indeed strictly decreasing in I^* by the same property of $P(I)$. This establishes uniqueness. Existence follows from i) continuity of $P(I)$, ii) as at $I^* = 0$ we have from our parameter restriction on $P(0)$ that the right-hand side is strictly positive, and iii) as the left-hand side increases beyond bounds in I^*. We next determine I_s^*. As firms choose maximum leverage, we have from the incentive constraint that $D_i = I_i[P(I^*) - c_i - b/q]$. Together with the break-even constraint $D_i = L_i(1 + r_i)$ and after aggregation over all firms choosing $\theta_i = \theta_s$, we have

$$I_s^* = \frac{L_s(1 + r_s)}{P(I^*) - c_s - b/q} = K[1 - H(w^*)]\frac{1 + r_s}{P(I^*) - c_s - b/q}. \tag{10}$$

As a final step we can substitute from indifference $\frac{1 + r_s}{P(I^*) - c_s - b/q} = \frac{1 + r_0}{P(I^*) - c_{ns} - b/q}$, which yields (6).

Next, we establish when the industry is indeed financially constrained. For this define first I_0 by $P(I_0) - c_{ns} - (1 + r_0) = 0$. This (zero marginal profit) level of investment I_0 would be reached when firms were financially unconstrained. The industry is thus financially constrained when the solution to (5) is smaller than I_0. We can write this slightly different as follows: Setting $P(I^*) - c_{ns} - (1 + r_0)$ in (5), we have that the industry is financially constrained if $\frac{q}{b}(1 + r_0)A < I_0$, which transforms to

$$A < I_0 \frac{b}{q} \frac{1}{1 + r_0}. \tag{11}$$

We turn now to the determination of I_s^* in (6). For given parameter values, including the labelling threshold, warm-glow capital is indeed in short supply when $I_s^* < I^*$. Substituting all terms this can be written as

$$H\left((c_s - c_{ns})\frac{1 + r_0}{P(I^*) - c_{ns} - b/q}\right) > 1 - \frac{P(I^*) - c_{ns} - b/q}{b/q - [P(I^*) - c_{ns} - (1 + r_0)]}\frac{A}{K}. \tag{12}$$

We make the following observation. As the labelling threshold decreases and with it $c_s - c_{ns} \to 0$, the left-hand side converges to $H(0)$. Hence, warm-glow capital is indeed always in short supply (so that it does not determine the marginal investment in the industry) when there is a sufficiently large fraction of households with no warm-glow preferences, $H(0)$, exceeding the right-hand side in (12). **Q.E.D.**

24

Proof of Proposition 2. It remains to consider the comparative statics in $H(w)$. Recall that, in the presently analyzed case, the optimal choice of θ_s maximizes (8), with I_s^* given by (6) and w^* given by (4). The first-order condition is thus

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

with

$$\frac{dI_s^*}{d\theta_s} = -h(w^*)K\left(\frac{1 + r_0}{P(I^*) - c(\theta_m) - b/q}\right)^2 c'(\theta_s).$$

Substituting and transforming obtains for the first-order condition

$$\frac{h(w^*)}{1 - H(w^*)} = \frac{1/c_s' - 1}{(\theta_s - \theta_m) - (c_s - c_{ns})}\frac{P(I^*) - c_{ns} - b/q}{1 + r_0},$$

where we abbreviated cost expressions. We note now that the right-hand side is strictly decreasing in θ_s as this holds for $1/c_s'$ due to convexity of the cost function and as the denominator strictly increases from $1 - c_s' > 0$ at the optimum. As for any given w^* the left-hand side decreases by the MLRP of $H(w)$, which implies monotonicity of the hazard rate, the optimal θ_s must indeed increase. **Q.E.D.**

Proof of Proposition 3. When the industry is not financially constrained in the aggregate, recall that individual firms still are. This will now be of importance for sustainable firms, as it constrains the amount of outside financing and therefore also of warm-glow financing that they can raise individually.

Starting with non-sustainable firms, recall that when the industry is not financially constrained, the marginal return on investment is just zero, implying that $U_i = (1 + r_0)A$. This also pins down industry size from $P(I^*) - c_{ns} - (1 + r_0) = 0$.

As the sustainability type is endogenous, this must also hold for sustainable firms: Also their return on own assets must be r_0. Recall now that we have from the binding incentive and participation constraint that $U_i = \frac{b}{q}f_iA_i$, so that the indifference requirement, $U_s = U_{ns}$, is now that $(1 + r_0)A = \frac{b}{q}f_sA$, i.e.,

$$1 + r_0 = \frac{b}{q}\frac{1 + (r_0 - w^*)}{b/q - [P(I) - c_s - (1 + (r_0 - w^*))]}.$$

We can now substitute from $P(I^*) - c_{ns} - (1 + r_0) = 0$ so that $P(I) - c_s - (1 + (r_0 - w^*)) = w^* + (c_{ns} - c_s)$. After transformations this yields

$$w^* = (c_s - c_{ns})\frac{1 + r_0}{b/q - (1 + r_0)}.$$

Again, this is just a function of the primitives and $c_s = c(\theta_s)$. Finally, the fraction of sustainable investment is again obtained from the respective market-clearing condition, as in (10). On these expressions, we can finally conduct the analysis as for Corollary 1 and Proposition 2. **Q.E.D.**

Proof of Proposition 4. Proceeding as in Proposition 2 we thus have now

$$\frac{dI_s^*}{d\theta_s} = -\frac{dH(w^* \mid w^*)}{dw^*} K \left(\frac{1 + r_0}{P(I^*) - c(\theta_m) - b/q} \right)^2 c'(\theta_s)$$

and ultimately the first-order condition for the social planner

$$\frac{dH(w^* \mid w^*)}{dw^*} \frac{1}{1 - H(w^*)} = \frac{1/c_s' - 1}{(\theta_s - \theta_m) - (c_s - c_{ns})} \frac{P(I^*) - c_{ns} - b/q}{1 + r_0}. \tag{13}$$

Here, the left-hand side consists of two terms:

$$\frac{h(w^* \mid w^*)}{1 - H(w^* \mid w^*)} + \frac{H_2(w^* \mid w^*)}{1 - H(w^* \mid w^*)}.$$

The second term arises only when the discussed feedback effect is present and it is strictly positive. By the arguments in Proposition 2, i.e., as the right-hand side in (13) is strictly monotonic in θ_s, this implies that the optimal threshold is strictly lower in the presence of the feedback effect. **Q.E.D.**

Proof of Proposition 6. The objective function (7) can now be written as

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

As noted in the main text, we confine ourselves to a local analysis (thus assuming that the objective function is well behaved). We conduct the analysis at the point where initially no separate (higher) standard for the label is set. Focussing thus on the minimum standard, the first-order condition yields

$$\frac{d\Omega}{d\theta_m} = \frac{dI^*}{d\theta_m} [P(I^*) - (1 + r_0) + (\theta_m - c(\theta_m))] + I^*[1 - c'(\theta_m)] = 0. \tag{14}$$

When the marginal social surplus is strictly negative, we have from $\frac{dI^*}{d\theta_m} < 0$ that $c'(\theta_m) > 1$. We now consider the introduction of $\theta_s > \theta_m$. Starting from $\theta_s = \theta_m$, the local derivative is

$$\frac{d\Omega}{d\theta_s} = I_s^*[1 - c'(\theta_m)],$$

which is then strictly negative. When the marginal social surplus is strictly positive, the opposite holds, so that $\frac{d\Omega}{d\theta_s} > 0$ at $\theta_s = \theta_m$. **Q.E.D.**

26

Recent Issues