Asset Prices and Portfolios with Externalities*

Steven D. Baker[†] Burton Hollifield[‡] Emilio Osambela[§]

September 19, 2022

Abstract

Elementary portfolio theory implies that environmentalists optimally hold more shares of polluting firms than non-environmentalists, and that polluting firms attract more investment capital than otherwise identical non-polluting firms through a hedging channel. Pigouvian taxation can reverse the aggregate investment results, but environmentalists still overweight polluters. We introduce countervailing motives for environmentalists to underweight polluters, comparing the implications when environmentalists coordinate to internalize pollution, or have nonpecuniary disutility from holding polluter stock. With nonpecuniary disutility, introducing a green derivative may dramatically alter who invests most in polluters, but has no impact on aggregate pollution.

*We thank the editor, Martin Kacperczyk, the editorial team at Review of Finance and the referee for their helpful and constructive comments. We also than Aytek Malkhozov, Edward Van Wesep, Steffen Hitzemann, Mete Kilic, Marcus Opp, participants at meetings of the CEPR ESSFM (2018), NFA (2018), Paris Conference on Commodities, Volatility, and Risk Management (2019), UVA/Richmond Fed Jamboree (2019), FIRS (2019), SED (2019), CEMA (2019), and WFA(2020), and seminar participants at Carnegie Mellon University, the Board of Governors of the Federal Reserve System, INSEAD, the University of Georgia, the University of Luxembourg, and WU (Vienna) for comments and suggestions. Steven Baker gratefully acknowledges past support from the McIntire School of Commerce at the University of Virginia. The views expressed in this article are solely those of the authors. They do not necessarily reflect the views of the Federal Reserve Bank of Richmond or the Federal Reserve System.

†Federal Reserve Bank of Richmond, steven.baker@rich.frb.org

‡Tepper School of Business, Carnegie Mellon University, burtonh@cmu.edu

§Board of Governors of the Federal Reserve System, emilio.osambela@frb.gov

1 Introduction

A key obstacle to addressing global warming is a source of market failure recognized since at least Pigou (1920): greenhouse gas emissions create a negative externality. Producers and consumers of carbon intensive products typically make small individual contributions to the global stock of greenhouse gases, and carbon dioxide (CO_2) emissions would rarely have immediate or localized negative impacts. Hence such emissions are a particularly difficult externality to address, e.g., through Pigouvian taxation, because of their global scope and cumulative, long term, negative effects.

Attention has shifted to private initiatives, including the idea that capital markets might naturally address global warming even if government support is limited. To better understand the conditions for financial market solutions to emerge, we embed a negative externality in a canonical portfolio choice model in which some firms use a dirty technology and others use a clean technology. A fraction of investors we call environmentalists is more sensitive to the externality than other investors. We characterize the investment share of clean and dirty firms, the portfolios of the two investor types, and the sensitivity of equilibrium to the population share of environmentalists.

Our initial results cut against the idea that financial markets can help to mitigate externalities. Serious negative externalities of uncertain magnitude such as global warming represent undiversifiable risk. Assets that hedge such risk will trade at a premium. Stock in greenhouse gas emitting industries offers such a hedge. Environmentalists suffering the greatest utility loss have the strongest hedging motive, and should load up on polluters. If society in general has heightened sensitivity to environmental damage, then the cost of capital to polluters falls, increasing investment in polluting industries and environmental damages. The private hedging benefit creates the public cost of increased global warming.

We do not suggest that the hedging motive is the main driver for how investors behave in practice. Rather our message is that asset pricing and portfolio theory mechanisms should be part of the analysis of firms with negative environmental externalities.

To understand if financial market mechanisms complement government action, we include Pigouvian taxation in a production economy. Pigouvian taxation on dirty investment has a positive indirect effect via the hedging channel: dirty investment declines not only because of reduced net-of-tax returns, but also because hedging demand falls with decreased pollution. The market response enhances government action. However environmentalists still invest relatively more in dirty firms than other investors, even with Pigouvian taxation.

Since our initial findings are counterintuitive, we consider several modifications and extensions to our baseline model that lead to different results. We begin by reviewing several key intuitive and common assumptions in the literature underpinning our results. Similar to Brav and Heaton (2021), we assume that emissions and returns to dirty investment are positively correlated, and that high emissions states are high marginal utility states — bad times — especially for environmentalists. Allowing for subsidies for those worst affected by emissions or decoupling emissions from dirty output and imposing large enough emissions taxes on polluters can reverse some of our results.

We follow the integrated climate assessment literature (e.g., Nordhaus (1991); Nordhaus et al. (1992); Nordhaus (2015)) in assuming a damage function converting emissions into lost consumption, which mirrors the cost-benefit analysis commonly employed by policy makers. Our results hold for general damage functions, but impose restrictions on investor preferences when environmental quality and wealth are distinct consumption goods.

Simple alterations to conventionally assumed investor behavior reconcile our model's predictions with recent financial industry developments such as pledges of coordinated action on the environment and the rise of socially or environmentally responsible funds. Chava (2014) notes that \$1 in every \$8 invested is under socially responsible investment (SRI) restrictions, and finds that firms engaged in environmentally damaging activities have a higher cost of capital in both equity and debt markets. He further argues that such increased costs are not plausibly attributable to increased risk of future regulation alone. We model two behavioral mechanisms countervailing the hedging motive.

In the first mechanism, environmentalists internalize their collective contribution to pollution, and coordinate their investment strategy, optimally reducing their investment in polluters. The strategy is effective in reducing aggregate investment in dirty production, and may produce a positive spillover: if non-environmentalists in the population care somewhat about pollution, it is optimal for them to also reduce investment in polluters, since disinvestment by environmentalists reduces everyone's hedging demand. Internalizing cannot explain unilateral disinvestment by environmentalists when they are a small share of the population. Because environmentalists who internalize are motivated by the aggregate impact of their actions, they take stronger action when they are a larger share of the population.

In the second mechanism, environmentalists suffer nonpecuniary disutility from investment in polluters. This explanation is isomorphic with environmentalists disagreeing with non-environmentalists about the expected returns on polluter stock. In this case, environmentalists substantially reduce their individual investment in polluters even if their action has small impact on the aggregate externality, as for example when they are a small share of the population. As a consequence, aggregate investment in polluters decreases approximately linearly with the environmentalists' share of the population.

Nonpecuniary disutility presents an opportunity for financial innovation: investments that are close statistical substitutes could offer different nonpecuniary utility. For example, green bonds might offer environmentalists positive nonpecuniary utility, even if green bonds are statistically similar to conventional bonds issued by non-polluting companies. Alternatively, purchasing carbon emissions allowance futures could avoid disutility even if such futures have high realized returns when pollution is high. Virtuous emissions allowance prices and sinful dirty investment are statistically intertwined.

We model a green alternative to polluter stock, a pure financial innovation styled as a carbon emissions forward contract that doesn't change the underlying productive technologies. Environmentalists shift away from direct investment in polluters, preferring the forward contract instead, whereas non-environmentalists short the forward and increase their direct investment in polluters: essentially, they perform statistical arbitrage. Strikingly, aggregate investment is unchanged by introducing the green alternative, despite dramatically altering direct investment in polluters by individual market participants.

Heinkel et al. (2001) model green investors as constrained to hold only clean firms, finding that sufficient divestment of polluter's stock by green investors can raise the polluter's cost of capital. In much of the emerging theory work on SRI, the preferences underpinning divestment are not explicitly modeled: In Heinkel et al. (2001) investors who care about externalities exogenously overweight clean firms, whereas in Oehmke and Opp (2020) and Landier and Lovo (2021) SRI investors coordinate to alter corporate policy but are risk-neutral, and so do not weigh SRI objectives relative to hedging or diversification benefits. We instead focus on understanding how assumptions regarding preferences can lead to different portfolio choices and equilibrium outcomes, showing that the hedging motive leads investors to underweight clean firms. We study the incentives – and disincentives – for moral or environmentally conscious investors to favor SRI, taking the externalities produced by some firms as intrinsic to their production technologies.

Roth Tran (2019) considers the portfolio choice problem of a philanthropy in partial equilibrium. Brav and Heaton (2021) show that expected returns will be lower for dirty firms when their returns act like a hedge in a two-state partial equilibrium model. We study general equilibrium endowment and production economies with a heterogeneous general investor population, and solve for endogenous prices, investment, and externality in general equilibrium. Our work provides a structure for interpreting the growing literature on SRI, environmental risk and asset prices (e.g., Baker et al. (2018); Barber et al. (2021); Engle et al. (2020); Hsu et al. (2022); Humphrey et al. (2020); Pástor et al. (2021); Zerbib (2019)).

2 Endowment Economy

To convey the theoretical intuition in a simple setting, we start with a static one-period portfolio choice model with fixed capital. Section 3 allows for flexible capital. While greenhouse gas emissions provide a motivating example, our model applies to externalities generally.

There are two dates, $t \in \{0, 1\}$. Each of two firms, $i \in \{C, D\}$, issue one unit of stock at price P_i in period zero, and produce a period 1 dividend $\tilde{Y}_i$. The dividends are normally distributed, with identical mean μ and variance σ^2, and correlation $\rho < 1$.[1] The distinguishing feature is that firm C uses a clean technology that produces no externality, whereas firm D uses a dirty technology that produces a negative externality $\tilde{X}$. We have in mind, for example, two electrical power generating companies, one of which uses wind, the other coal. We interpret the dividend of each firm as reflecting the quantity of electricity supplied, such that the negative externality is proportional to dividends of firm D: $\tilde{X} = \tilde{Y}_D$.[2] The government may levy a tax τ on dividends of dirty firm D, such that after tax dividends are $(1-\tau)\tilde{Y}_D$. Tax revenues $\tilde{T} = \tau\tilde{Y}_D$ are redistributed equally amongst investors.

There are two types of atomistic investor, $j \in \{L, M\}$, with population mass $\eta \in (0, 1)$ of type M and mass $1 - \eta$ of type L. Investors are distinguished by their sensitivity to environmental damage, or pollution. Type L suffers relatively less from pollution $\tilde{X}$, whereas type M suffers more from $\tilde{X}$.

Investors derive utility from composite consumption

$$\tilde{W}_j - \lambda_j \tilde{X} + \tilde{T}, \tag{1}$$

where $\tilde{W}_j$ is terminal wealth, $-\lambda_j \tilde{X}$ captures disutility from pollution, with $0 \leq \lambda_L < \lambda_M$, and $\tilde{T}$ is government transfers. Our treatment of disutility from pollution follows Nordhaus (2015), who models climate change damage as a linear function of global emissions, with

[1] $\rho = 1$ implies that the dividends of clean firms are also perfectly correlated with the externality.

[2] We assume one-to-one proportions without loss of generality.

heterogeneous slope coefficients capturing different regional exposures.[3]

Investors solve a portfolio optimization problem over terminal wealth, trading shares in each firm $\{\theta_{C,j}, \theta_{D,j}\}$, and a riskfree asset. Without loss of generality, we assume that the risk-free rate is zero, and that investors have zero initial wealth so terminal wealth is

$$\tilde{W}_j = \theta_{C,j}(\tilde{Y}_C - P_C) + \theta_{D,j}((1-\tau)\tilde{Y}_D - P_D). \tag{2}$$

Investors of each type $j \in \{L, M\}$ maximize CARA utility over terminal consumption,

$$\max_{\theta_{C,j},\theta_{D,j}} \mathbb{E}[-\exp\{-\alpha(\tilde{W}_j - \lambda_j \tilde{X} + \tilde{T})\}] \equiv \max_{\theta_{C,j},\theta_{D,j}} \mathbb{E}[-\exp\{-\alpha(\tilde{W}_j - \lambda_j \tilde{Y}_D + \tau \tilde{Y}_D)\}]. \tag{3}$$

All investors have identical coefficient of absolute risk aversion $\alpha > 0$, and differ only in their sensitivity to pollution λ_j. Equation (3) explicitly assumes investors are aware that pollution is worse when the dirty technology produces more electricity, which is also when the dirty firm is able to pay a higher dividend $\tilde{Y}_D$. Investors are also aware that government revenues and transfers are dependent upon taxation of the dirty firm. Atomistic investors are powerless to alter the dynamics of either the externality or transfers, and simply take them as given. Individual investors can, however, choose portfolio allocations that hedge risk of environmental damage.

After solving Equation (3) from its first order conditions, the portfolio rule for each investor type $j \in \{L, M\}$ is

$$\theta_{C,j} = \frac{\mu - P_C}{\alpha\sigma^2} - \rho \frac{\mu - \rho(\mu - P_C) - \frac{P_D}{1-\tau}}{\alpha\sigma^2(1-\rho^2)}, \tag{4}$$

$$\theta_{D,j} = \frac{1}{1-\tau}\left(\frac{\mu - \frac{P_D}{1-\tau}}{\alpha\sigma^2} - \tau + \lambda\right) - \rho\frac{1}{1-\tau}\frac{\mu - P_C - \rho\left(\mu - \frac{P_D}{1-\tau}\right)}{\alpha\sigma^2(1-\rho^2)}. \tag{5}$$

The allocation to the clean firm $\theta_{C,j}$ is the same across types, because $\tilde{Y}_C$ is unrelated to the

[3] Appendix A shows the robustness of our results to nonlinear damage functions and more general utility functions.

externality. Examining the two types' allocations to the dirty firm provides our first result.

Proposition 1 *Investors of type M, who are hurt most by pollution, choose to hold more shares of polluting firm D than do investors of type L, who are hurt less by pollution. The difference in shareholdings, $\theta_{D,M} - \theta_{D,L}$, increases as firm D is more heavily taxed.*

Proof. From Equation (5) and $\lambda_L < \lambda_M$.

$$\theta_{D,M} - \theta_{D,L} = \frac{\lambda_M - \lambda_L}{1-\tau} > 0. \tag{6}$$

■

The intuition behind Proposition 1 is simple: investors wish to hedge environmental risk, by substituting material consumption for lost environmental consumption. To do so, they buy stock in firm D, which pays off when pollution is high. Because type M suffers most from pollution, he has the strongest hedging motive, and so buys more of stock D than does type L. A tax increase scales down dividends, so investors scale up their trade to maintain the same after-tax hedge.[4]

Imposing market clearing allows us to solve for stock prices. Substituting optimal portfolios into the market clearing conditions and solving for equilibrium stock prices,

$$P_C = \mu - \alpha\sigma^2 - \rho\alpha\sigma^2(1-\bar{\lambda}), \tag{7}$$

$$P_D = (1-\tau)\left(\mu - \alpha\sigma^2(1-\bar{\lambda})\right) - \rho(1-\tau)\alpha\sigma^2, \tag{8}$$

where

$$\bar{\lambda} = \eta\lambda_M + (1-\eta)\lambda_L$$

is the average population sensitivity. That stock in polluting firm D hedges environmental risk leads to a counterintuitive result.

[4]With fixed capital, the tax does not reduce the externality.

Proposition 2 *Suppose $\tau = 0$. Then $P_D > P_C$: stock in the polluting firm is worth more than stock in an otherwise identical nonpolluting firm.*

Proof.

$$P_D - P_C = \alpha\sigma^2(1-\rho)\bar{\lambda} > 0.$$

■

By continuity, the proposition continues to hold when τ and ρ are sufficiently small.

3 Production Economy

The previous section illustrated the main theoretical intuition in a simple and familiar portfolio choice framework with exogenous dividends. Analogous results hold when capital may be frictionlessly allocated to dirty or clean production technologies. Capital adjustment frictions are important for the effects of taxes. With fixed capital, a tax on the polluting firm's dividends reduces the value of its stock but does nothing to reduce pollution. With frictionless capital adjustment, Pigouvian taxation works to reduce equilibrium pollution, and the value of each firm is always identical to its capital stock, i.e., Tobin's $q = 1$. In reality capital is neither fixed nor completely liquid; that our results hold at both extremes suggest the mechanism at work is general.

As before there are two competitive firms, C and D. Each firm has access to a risky technology that produces the same final consumption good. In period 0, investors of type $j \in \{L, M\}$ choose investment $I_{C,j}$ and $I_{D,j}$, which they finance by borrowing at a risk-free rate normalized to zero. Output from production, realized in period 1, is

$$\tilde{\mu}_i I_i, \ i \in \{C, D\}, \tag{9}$$

where per capita investment in each firm is a weighted average reflecting the share of less

(L) and more (M) environmentally sensitive investors in the population,

$$I_i = (1-\eta)I_{i,L} + \eta I_{i,M}, \tag{10}$$

and $\tilde{\mu}_i$ is stochastic productivity. Analogous to the previous section, $\tilde{\mu}_D$ and $\tilde{\mu}_C$ are exogenous normal random variables with mean μ, standard deviation σ, and correlation ρ.

In the production setting it is convenient to take investment $I_{i,j}$ as the choice variable for each investor. To compare results from the production setting to the endowment setting, view output as a period 1 dividend,

$$\tilde{Y}_i = \tilde{\mu}_i I_i. \tag{11}$$

Normalize the number of shares issued by firm i to 1. Constant returns to scale and no capital adjustment costs together imply the firm i's stock price is equal to the investment in firm i: $P_i = I_i$. The shareholdings of investors j in firm i are

$$\theta_{i,j} = \frac{I_{i,j}}{I_i}. \tag{12}$$

As before, the dirty firm's technology produces a negative externality,

$$\tilde{X} = \tilde{\mu}_D I_D. \tag{13}$$

One interpretation is that the consumption good is electricity and I_D determines how many coal power plants are built, whereas $\tilde{\mu}_D$ captures intensity of utilization and hence the rate at which each plant consumes coal. The externality is proportional to the amount of dirty electricity generated.[5]

Dividends of the dirty firm are subject to a Pigouvian tax, so the after tax dividend is $(1-\tau)\tilde{Y}_D$. It is not important whether the proportional tax is levied on gross sales of

[5]As in Section 2, the constant of proportionality is not important, since this is captured later by the choices of λ_j.

the final good, on investment, or on the dividend: the result is the same. Tax receipts are distributed equally among investors.

Otherwise we assume the same setup as before, such that each investor solves

$$\max_{I_{C,j},I_{D,j}} \mathbb{E}[-\exp\{-\alpha(I_{C,j}(\tilde{\mu}_C - 1) + I_{D,j}((1-\tau)\tilde{\mu}_D - 1) - \lambda_j \tilde{X} + \tilde{T})\}]. \tag{14}$$

Solving from first order conditions for each investor $j \in \{L, M\}$,

$$I_{C,j} = \frac{\mu - 1}{\alpha\sigma^2} - \rho \frac{\mu(1-\rho) + \rho - \frac{1}{1-\tau}}{\alpha\sigma^2(1-\rho^2)}, \tag{15}$$

$$I_{D,j} = \frac{1}{1-\tau}\left(\frac{\mu - \frac{1}{1-\tau}}{\alpha\sigma^2} - (\tau - \lambda_j)I_D\right) - \rho \frac{(\mu(1-\tau)-1)(1-\rho)+\tau}{\alpha\sigma^2(1-\rho^2)(1-\tau)^2}. \tag{16}$$

Each type of investor allocates the same capital to the clean firm. Notice that $I_{D,j}$ depends on per-capita investment in the dirty firm I_D, which individual investors take as given when solving their optimization problems, as well as on sensitivity to pollution λ_j. Employing equilibrium condition Equation (10),

$$\begin{aligned} I_{D,L} =& \frac{(1-\tau+\eta(\lambda_L - \lambda_M))\left(\mu - \frac{1}{1-\tau}\right)}{(1-\bar{\lambda})\alpha\sigma^2\,(1-\tau)} \\ &+ \rho \frac{(1-\tau+\eta(\lambda_L-\lambda_M))((1-\rho)(1-\mu(1-\tau))-\tau)}{(1-\bar{\lambda})\alpha\,\sigma^2(1-\rho^2)(1-\tau)^2}, \end{aligned} \tag{17}$$

$$\begin{aligned} I_{D,M} =& \frac{(1-\tau-(1-\eta)(\lambda_L - \lambda_M))\left(\mu - \frac{1}{1-\tau}\right)}{(1-\bar{\lambda})\alpha\sigma^2(1-\tau)} \\ &+ \rho \frac{(1-\tau-(1-\eta)(\lambda_L-\lambda_M))((1-\rho)(1-\mu(1-\tau))-\tau)}{(1-\bar{\lambda})\alpha\sigma^2(1-\rho^2)(1-\tau)^2}. \end{aligned} \tag{18}$$

We can now state propositions analogous to those in the previous section. We assume parameter values such that equilibrium investment in both firms is strictly positive.

Proposition 3 *Investors of type M, who are hurt most by pollution, invest more in polluting firm D than do investors of type L, who are hurt less by pollution.*

Proof. From Equation (17) and Equation (18),

$$\frac{I_{D,L}}{I_{D,M}} = \frac{1-\tau+\eta(\lambda_L-\lambda_M)}{1-\tau-(1-\eta)(\lambda_L-\lambda_M)} < 1,$$

since $\lambda_L - \lambda_M < 0$. Given parameters such that investment in firm D is strictly positive, the result follows. ■

Unlike the case with fixed capital, increasing the Pigouvian tax τ will decrease investment in the dirty firm. This will also cause the absolute difference $I_{D,M} - I_{D,L}$ to decrease, although the relative difference $I_{D,M}/I_{D,L}$ is increasing in τ. Equivalently, raising Pigouvian taxes will cause the environmentalist to hold an increasing fraction of the shares in the polluting firm, even though the value of those shares will decrease.

Combining Equations (10, 17, 18),

$$I_C = \frac{\mu-1}{\alpha\sigma^2} - \rho\frac{\mu-\rho(\mu-1)-\frac{1}{1-\tau}}{\alpha\sigma^2\left(1-\rho^2\right)}, \tag{19}$$

$$I_D = \frac{1}{1-\bar{\lambda}}\frac{\mu-\frac{1}{1-\tau}}{\alpha\sigma^2} - \rho\frac{1}{1-\bar{\lambda}}\frac{\mu-1-\rho\left(\mu-\frac{1}{1-\tau}\right)}{\alpha\sigma^2\left(1-\rho^2\right)}. \tag{20}$$

This leads to a proposition for investment analogous to Proposition 2 for prices in the previous section.

Proposition 4 *Suppose $\tau = 0$ and $\rho < 1$. Then $I_D > I_C$: investment in the polluting firm is greater than investment in an otherwise identical firm that does not pollute.*

Proof. For $\tau = 0$,

$$\frac{I_D}{I_C} = \frac{1}{1-(1-\eta)\lambda_L-\eta\lambda_M} > 1,$$

since $\eta\lambda_M > 0$, $(1-\eta)\lambda_L \geq 0$, and we assume parameters such that investment is greater than zero in both firms. ■

In summary, our main results hold for either fixed or flexible capital allocation across firms. Because the model with flexible capital endogenizes the amount of the externality, our extensions build on this foundation.

4 Discussion and Extensions

Before proceeding with variants and extensions of the basic model, we note several assumptions that can be relaxed without altering our qualitative results and some that cannot be relaxed. Our use of a damage function to convert environmental harm into pecuniary losses is standard in the integrated climate assessment modeling literature; see, e.g., Nordhaus (1991), Nordhaus et al. (1992), and Nordhaus (2015). A common criticism of such analysis, exemplified by Pindyck (2013), is that inadequate attention is paid to tail risk, or to preferences with strong aversion to tail risk. In Appendix A we prove analogues of Propositions 1 and 2 for generalized preferences and damage functions, including nonlinear damage functions that would produce fat tails in the distribution of environmental damage.[6]

Complementarity between environmental quality and general consumption or wealth in investor preferences can reverse our results. Appendix A.1 provides examples of preferences consistent with the environmental economics literature in which low environmental quality states are also low marginal utility states. An example of such preferences is $U(W,X) = \frac{(WX^{-\beta})^{1-\gamma}}{1-\gamma}$, with $\beta > 0$ and $\gamma < 1$. Interpret $WX^{-\beta}$ as a composite good strictly increasing in wealth and strictly decreasing in the externality. The investor has constant relative risk averse preferences over the composite good. A coefficient of relative risk aversion $\gamma < 1$ reverses our results while a coefficient of relative risk aversion $\gamma > 1$ maintains our results.

We assume, similar to Brav and Heaton (2021), that the returns of firm D are more positively correlated with the externality than those of firm C.[7] The assumption chiefly reflects our focus on the risk of environmental damage itself, rather than on policy risk

[6]Pindyck (2013) is principally concerned with the important and realistic effects of tail risk on magnitudes, specifically of societal WTP to reduce climate change, whereas our focus is on qualitative directional results for portfolio choice and risk premia. Our qualitative results are robust to fat tailed damage distributions.

[7]Pástor et al. (2021) suggest that shifting investor or consumer tastes might correlate with environmental damage, making firm D stock less positively correlated with the externality. We discuss a similar mechanism in a static setting in Section 4.4.

associated with environmental damage, as emphasized, for example, in Hsu et al. (2022). The government might tax dirty output at a predictable rate but redistribute tax revenues to citizens according to their type, e.g., offering greater help to those regions most adversely affected by environmental damage. If the government overcompensates those in damaged areas, then Proposition 3 is reversed, as we show in Section 4.1. Alternatively, the government might levy high taxes in high emissions states that are not offset by proportionally high revenues from output. Proposition 3 holds when Pigouvian taxation is a constant fraction of investment or output, but it may fail when taxation introduces risk, as we show in Section 4.2.

Although policy risk can overturn the hedging motive without departing from conventional preferences, SRI often presumes an inadequate governmental response to externalities. Therefore we also consider two departures from conventional behavioral assumptions that are directly motivated by SRI: that the more sensitive (M) investors internalize the public impact of their collective investment decisions (Section 4.3), or that the more sensitive investors get nonpecuniary disutility from investing in the dirty firm (Section 4.4).

These explanations differ in two key respects. First, coordinating investors are concerned with the efficacy of collective actions to reduce the externality, whereas nonpecuniary utility does not require that investor actions have any effect on the externality. The second distinction is intrinsic to the idea of coordination: it relaxes the assumption of fully competitive investor behavior, whereas that assumption is maintained with nonpecuniary utility. Empirically, coordination and nonpecuniary utility may be combined: for example, the social norms that motivate Hong and Kacperczyk (2009) are not optimally coordinated decisions oriented towards a specific objective, but clearly they have aspects of group coordination. We differentiate the two explanations within our theoretical framework, and highlight their different implications.

Table 2 summarizes the extensions we consider. We assume $\rho = 0$ to simplify expressions.

4.1 Proportional Taxation with Redistribution by Type

The government may levy a uniform Pigouvian tax at rate τ proportional to emissions, but choose to redistribute tax revenues to citizens according to their type, e.g., offering greater help to those regions most adversely affected by warming.

Suppose the government can reliably identify investors by type. For example, type M may be residents of coastal or warm regions that are more adversely affected by global warming, whereas type L are residents of inland or cold regions. Assume moving is sufficiently costly that investors don't change type. Define type-specific tax rebate $g_j \tilde{\mu} I_D$, such that $\eta g_M + (1-\eta) g_L = \tau$, ensuring aggregate tax revenues still equal aggregate rebates.

With this change, each type's net exposure to the externality is now $(g_j - \lambda_j)\tilde{X}$, and optimal investment in the dirty firm is

$$I_{D,j} = \frac{1}{1-\tau} \left(\frac{\mu - \frac{1}{1-\tau}}{\alpha \sigma^2} - (g_j - \lambda_j) I_D \right). \tag{21}$$

For a sufficiently redistributive policy, $g_M - \lambda_M > g_L - \lambda_L$. Type M has lower net exposure to the externality than type L, such that Proposition 3 is reversed: type L invests most in the dirty firm. This happens when the government overcompensates type M for their losses due to the externality and undercompensates type L.

4.2 Taxation with Emissions Uncertainty

Suppose the dirty technology has an uncertain cost-benefit tradeoff, such that realized emissions may be high even when realized output is low. In this section only, we formalize this by redefining the externality as

$$\tilde{X} = \left[\rho_X (\tilde{\mu}_D - \mu) + \sqrt{1 - \rho_X^2}\,\tilde{\epsilon} + \mu \right] I_D, \tag{22}$$

such that $\rho_X = Corr(\tilde{X}, \tilde{\mu}_D) \in (0,1)$, whereas in the baseline model we assume $\rho_X = 1$. The mean and variance of $\tilde{X}$ remain μI_D and $\sigma^2 I_D^2$, respectively.

As I_D increases, both output and emissions rise on average, and since $\rho_X > 0$, dirty output is still positively correlated with the externality.

Pigouvian taxes levied on firm D are $\tau \tilde{X}$, i.e., they are proportional to the externality, but are not exactly proportional to output. The government holds the dirty firm responsible for all of the emissions it produces, whether or not such emissions are productive.

Net of taxes, returns to firm D investment are

$$\tilde{r}_D = \tilde{\mu}_D - \tau \frac{\tilde{X}}{I_D} - 1, \tag{23}$$

and the correlation of net dirty returns with the externality is

$$Corr(\tilde{r}_D, \tilde{X}) = \frac{(\rho_X - \tau)\sigma}{\hat{\sigma}}, \tag{24}$$

where $\hat{\sigma}$ is the standard deviation of net of tax returns.[8]

$Corr(\tilde{r}_D, \tilde{X})$ is positive if $\rho_X > \tau$, and negative if $\rho_X < \tau$. If dirty output is only weakly linked to emissions and Pigouvian taxes are high, net returns may become negatively correlated with emissions due to taxes.

Assume that firm C productivity is uncorrelated with the externality $\tilde{X}$ and firm D productivity, such that optimal clean investment is given by Equation (15) with $\rho = 0$.

Then optimal dirty investment by type j is

$$I_{D,j} = \frac{(1-\tau)\mu - 1 - \alpha(\tau - \lambda_j)(\rho_X - \tau)\sigma^2 I_D}{\alpha \hat{\sigma}^2}, \tag{25}$$

[8] $\hat{\sigma} = \sqrt{(1-\tau\rho_X)^2 + \tau^2(1-\rho_X^2)}\sigma$.

and the difference between the dirty investment of types M and L is

$$I_{D,M} - I_{D,L} = \frac{(\lambda_M - \lambda_L)(\rho_X - \tau)\sigma^2 I_D}{\hat{\sigma}^2}. \tag{26}$$

Since $\lambda_M > \lambda_L$, investor M invests less in the dirty firm than investor L when $\rho_X < \tau$, i.e., when net dirty returns are negatively correlated with the externality. In this case, Pigouvian taxation can overturn Proposition 3.

4.3 Internalizing the Externality via Coordinated Investment

Investors of type M internalize their contribution to the externality by thinking as follows: "if everyone of my type behaves as I do, how will that affect the externality?" We assume no Pigouvian taxation, $\tau = 0$, to simplify expressions.

Since the externality is

$$\tilde{X} = \tilde{\mu}_D I_D = \tilde{\mu}_D \left[(1-\eta) I_{D,L} + \eta I_{D,M}\right], \tag{27}$$

M can affect the externality through the choice of $I_{D,M}$. We assume he cannot coordinate with type L, and takes the investment decision of type L investors as given. Type M's investment problem is

$$\max_{I_{C,M}, I_{D,M}} \mathbb{E}[-\exp\{-\alpha(I_{D,M}((1-\eta\lambda_M)\tilde{\mu}_D - 1) + I_{C,M}(\tilde{\mu}_C - 1) - I_{D,L}(1-\eta)\lambda_M \tilde{\mu}_D)\}]. \tag{28}$$

The key aspect of the problem in Equation (28) is that it incorporates knowledge of M's population share, η.

Some shorthand notation simplifies expressions for equilibrium investment, and also has a natural interpretation. Investor M effectively taxes the polluting firm's output at rate $\tau_M = \eta\lambda_M$, reflecting his sensitivity to the externality and his type's population share. The analogous expression for investor L is $\tau_L = (1-\eta)\lambda_L$. We also define net-of-tax coefficients

$T_L = 1 - \tau_L$ and $T_M = 1 - \tau_M$.

Optimal investment for type M conditional on the investment choice of type L is

$$I_{C,M} = \frac{\mu - 1}{\alpha\sigma^2}, \tag{29}$$

$$I_{D,M} = \frac{1}{T_M}\left[\frac{\mu - \frac{1}{T_M}}{\alpha\sigma^2} + (1-\eta)\lambda_M I_{D,L}\right]. \tag{30}$$

Since internalizing is like privately taxing investment in firm D at rate τ_M, it is not surprising that M's investment strategy resembles that with Pigouvian taxation at rate $\tau = \tau_M$.

$I_{D,M}$ depends recursively on investor L's investment decision. Solving for equilibrium requires a further assumption: does L also internalize, coordinating with other type L investors, or not? Here we focus on the case where L does not internalize, with the case where both investors internalize in Appendix B.

If L does not internalize, equilibrium investment in the dirty firm is

$$I_{D,L} = \frac{(1 + (\lambda_L - \lambda_M)\eta)\mu - T_M - \frac{\lambda_L \eta}{T_M}}{(1 - \tau_L - \tau_M)\alpha\sigma^2}, \tag{31}$$

$$I_{D,M} = \frac{(1 + (\lambda_M - \lambda_L)(1-\eta))\mu - \frac{T_L}{T_M} - \lambda_M(1-\eta)}{(1 - \tau_L - \tau_M)\alpha\sigma^2}, \tag{32}$$

$$I_D = \frac{\mu - \frac{1-(1-\eta)\tau_L}{T_M}}{(1 - \tau_L - \tau_M)\alpha\sigma^2}. \tag{33}$$

Internalizing differs from the case with Pigouvian taxation for several reasons. The private tax τ_M only applies to type M, since type L does not internalize.[9] And while internalizing is conceptually similar to a tax, it generates no revenues to redistribute. Finally there is the equilibrium hedging effect: any reduction in I_D reduces the hedging motive, because it reduces the variance of $\tilde{X}$. Provided $\lambda_L > 0$, $I_{D,L}$ will decrease somewhat in response to internalizing by type M because type M's actions reduce the mean and variance of $\tilde{X}$.

Comparing expressions for equilibrium investments leads to intuitive results: when M

[9]While τ_L and τ_M have a natural interpretation as investor-specific private tax rates, these terms arise in the expressions for equilibrium investment even when investors do not internalize.

internalizes, $I_{D,L}$, $I_{D,M}$, and I_D all decrease relative to the cases without internalizing and without taxation, but the decrease is smaller than a Pigouvian tax of $\tau = \tau_M$ achieves.

To compare internalizing to Pigouvian taxation in magnitudes, we present a numerical example using parameter values in Table 1. From an individual perspective, investment in firm D offers the same expected return as firm C, and also hedges bad realizations of the externality. But aggregate investment in firm D has effectively 1% lower expected return from the standpoint of investor L, or 3% lower expected return from the standpoint of investor M, due to the negative externality.[10]

In the base case where both investors behave competitively and without Pigouvian taxation, Figure 1 shows that type M invests more in firm D than type L, and aggregate investment I_D rises with M's population share, in line with Proposition 3. This is the hedging motive at work. However the difference $I_{D,M} - I_{D,L}$ is small, and as a consequence I_D rises only slightly with η. Imposing a fixed Pigouvian tax of $\tau = \lambda_M = 3\%$ reduces $I_{D,L}$, $I_{D,M}$, and I_D substantially, by about 1/3.[11] Otherwise the introduction of taxation changes little: M still invests relatively more than L, and the difference is still small.

When M internalizes, Figure 1 shows that $I_{D,M}$ and I_D become very sensitive to η. The two previously described reference cases bound investment when M internalizes. If $\eta = 0$, then $I_D = I_{D,L}$, and $I_{D,M}$ is the same as the case without Pigouvian taxation, because even as a group type M's investment does not contribute to the externality. When $\eta = 1$, then $I_D = I_{D,M}$, and $I_{D,M}$ is the same as the case with Pigouvian taxation at rate λ_M, because the externality is entirely attributable to group M's investment decision. Finally, internalizing produces a more rapid drop in I_D as η nears 1, due to the combined effects of M's increasing population share and the greater group incentive for type M to reduce investment per capita.

Include Figure 1 about here.

[10] Our production parameters are broadly consistent with stock market data. Damage parameters λ_L and λ_M would depend upon the externality being modeled, and are subject to considerable uncertainty; see for example Pindyck (2013). We discuss empirical linkages in Section 6. The qualitative results upon which we focus are relatively insensitive to parameter values.

[11] $\tau = \lambda_M$ is the largest tax that could be justified based on either investor's utility loss from the externality. So I_D in this case could be viewed as a lower bound for socially optimal investment in firm D.

In our example in Figure 1, internalizing cannot explain why environment investors (type M) would invest substantially less than others (type L) unless environmental investors are a large portion of the population. When η is small, say 5% of the population, type M still invests about the same as type L. Only when collective action would be very effective – large η – does M overcome his greater hedging motive and disinvest from the dirty firm.

4.4 Nonpecuniary Disutility from Investment

It is possible that socially responsible investors coordinate to achieve a common objective, thus overcoming their individual hedging motives. An even simpler explanation is that such investors have a visceral distaste for investment in polluting companies that overcomes pecuniary motives for doing so. Pástor et al. (2021) and Pedersen et al. (2021) use nonpecuniary utility from investment or investor tastes (Fama and French, 2007) to explain demand for socially responsible investment. Our framework allows us to explicitly model an underlying externality, noting how investors who coordinate to reduce the externality differ from those motivated by feelings about polluter stock ownership.

Similar behavior consistent with nonpecuniary utility is documented by Morewedge et al. (2018), who investigate why sports fans bet on their own teams to win, rather than hedging the disutility from a loss by betting on the opposing team. The hypothesized source of disutility is a conflict with the bettor's identity as a supporter of his team. Similarly, an investor who cares greatly for the environment would find investment in a polluting firm contrary to his identity as an environmentalist, even if such investment hedges dirty environmental outcomes that are beyond his control.

Suppose type M receives nonpecuniary disutility proportional to his personal investment in firm D. As in Section 4.3, our starting point is the investment economy without Pigouvian taxation, and we assume that type L's problem is unchanged from this baseline.

Incorporating the new disutility, investors of type M solve

$$\max_{I_{C,M},I_{D,M}} \mathbb{E}[-\exp\{-\alpha(I_{D,M}(\tilde{\mu}_D - 1) + I_{C,M}(\tilde{\mu}_C - 1) - \zeta_D I_{D,M} - \lambda_M \tilde{X})\}]. \tag{34}$$

We nest the new disutility within total utility, and make it linear in $I_{D,M}$, with slope $-\zeta_D < 0$. While many alternative approaches are justifiable, this parsimonious choice fits nicely within our framework, and allows for simple comparisons with results from previous sections. It highlights what we view as the key distinction relative to Sections 3 and 4.3: disutility is now proportional to $I_{D,M}$, not I_D or $\eta I_{D,M}$. Hence nonpecuniary disutility from investment, while an unusual ingredient in a portfolio choice problem, is within the scope of a conventional competitive investor utility maximization problem: it is not proportional to aggregate choice variables, and it does not require coordination.

While we have in mind that parameter ζ_D scales a nonpecuniary penalty, in our static framework it is equivalent to say type M perceives expected net return of $\mu - \zeta_D - 1$ on investment in firm D, whereas type L perceives expected net return of $\mu - 1$. It is as if the types have heterogeneous beliefs about expected stock returns.[12] Hong and Kostovetsky (2012) discuss nonpecuniary incentives for fund managers to pursue SRI – the perk of increased approval from Democratic peers – as distinct from pecuniary explanations – such as different models of stock market behavior. They acknowledge that the two motives may have similar implications. In our example, the motives are equivalent.

Incorporating the new disutility, investor M chooses investment

$$I_{C,M} = \frac{\mu - 1}{\alpha\sigma^2}, \tag{35}$$

$$I_{D,M} = \frac{\mu - \zeta_D - 1}{\alpha\sigma^2} + \lambda_M I_D. \tag{36}$$

The key difference relative to Section 4.3 is that ζ_D is not scaled by η: whether or not his type

[12] This equivalence would not carry over to a dynamic setting, because the investors would agree about wealth dynamics with nonpecuniary disutility, whereas they would disagree about wealth dynamics with heterogeneous beliefs.

contributes much to the externality, M feels badly about investing in firm D. Investment by investor L is unchanged from Equation (16) and Equation (15), conditional on a given level of aggregate investment I_D in firm D.

Investment in the dirty firm is

$$I_{D,L} = \frac{(1+(\lambda_L-\lambda_M)\eta)(\mu-1)-\lambda_L\eta\zeta_D}{(1-\tau_L-\tau_M)\alpha\sigma^2}, \tag{37}$$

$$I_{D,M} = \frac{(1+(\lambda_M-\lambda_L)(1-\eta))(\mu-1)-T_L\zeta_D}{(1-\tau_L-\tau_M)\alpha\sigma^2}, \tag{38}$$

$$I_D = \frac{\mu-1-\eta\zeta_D}{(1-\tau_L-\tau_M)\alpha\sigma^2}, \tag{39}$$

for types L and M, and on an average per-capita basis, respectively. Similar to the results when M internalizes, investor L decreases his investment $I_{D,L}$ somewhat, because the decrease in I_D reduces the variance of the externality $\tilde{X}$, which reduces hedging demand.

Continuing our numerical example introduced in the previous section, we use common parameters in Table 1, and choose disutility parameter $\zeta_D = \frac{\lambda_M}{1-\lambda_M} = 3.09\%$.[13] Figure 2 shows $I_{D,L}$, $I_{D,M}$, and I_D incorporating disutility. We again include the frictionless investment case without Pigouvian taxation ($\tau = 0$), and with Pigouvian taxation at rate $\tau = \lambda_M = 3\%$, to provide points of reference.

Include Figure 2 about here.

Similar to Figure 1, aggregate investment I_D decreases substantially as M's population share η increases. But in Figure 2 the rate of decrease is approximately linear in η, because $I_{D,M}$ and $I_{D,L}$ are nearly constant over the domain of η. Even though I_D is nearly 1/3 lower at $\eta = 1$ than at $\eta = 0$, such that the standard deviation of the externality $\tilde{X}$ is also about 1/3 lower, $I_{D,M}$ and $I_{D,L}$ change by only a percent or two.

The example in Figure 2 suggests that disutility from investment could easily overcome

[13] ζ_D is chosen such that $I_{D,M}$ has the same value when $\eta = 1$ as the case where M internalizes, or there is Pigouvian taxation at rate λ_M. Our model is not oriented towards quantitative application, but our example parameters fall within a reasonable range. For example Barber et al. (2021) find that impact investors are willing to forgo returns of 2.5% to 3.7%.

the hedging motive, and that it could explain environmentalists' preference for socially responsible investment provided that nonpecuniary disutility from investment is of approximately the same magnitude as disutility from pollution. The example also suggests that, contrary to the case where environmentalists internalize, the population share of environmentalists should have little impact on how they invest.

5 Green Financial Innovation: Derivatives

Nonpecuniary utility from investment decouples feelings from tangible results: investors may forgo ownership of polluter stocks even if their decision has little impact on pollution. Such decoupling can generate demand for new securities differentiated by their nonpecuniary utility even if they do not expand the set of productive technologies.

Suppose there is a green derivative contract offering a similar payoff to investment in firm D, but without the disutility. Consider a forward contract with payoff

$$\tilde{S} - F, \tag{40}$$

where F is the equilibrium forward price with $\tilde{S}$ the underlying spot price, which we take as exogenous with distribution $\tilde{S} \sim \mathcal{N}(\mu, \sigma)$. $\tilde{S}$ has correlation $\rho_S \neq 0$ with the externality $\tilde{X} = \tilde{\mu}_D I_D$ or, equivalently, with firm D returns, but it is uncorrelated with $\tilde{\mu}_C$. For now we leave $\tilde{S}$ abstract: it is statistically related to $\tilde{\mu}_D$, but type M can trade the forward without disutility.

Define f_j, the forward position of investor j. In the economy with nonpecuniary disutility and forwards, investors L and M solve

$$\max_{I_{C,L},I_{D,L},f_L} \mathbb{E}[-\exp\{-\alpha(I_{D,L}(\tilde{\mu}_D - 1) + I_{C,L}(\tilde{\mu}_C - 1) + f_L(\tilde{S} - F) - \lambda_L\tilde{X})\}] \text{ and} \tag{41}$$

$$\begin{aligned}\max_{I_{C,M},I_{D,M},f_M} \mathbb{E}[-\exp\{-\alpha(I_{D,M}(\tilde{\mu}_D - 1) + I_{C,M}(\tilde{\mu}_C - 1) + f_M(\tilde{S} - F) \\ - \lambda_M\tilde{X} - \zeta_D I_{D,M})\}].\end{aligned} \tag{42}$$

Investor $j \in \{L, M\}$ optimally takes forward position

$$f_j = \frac{\mu - F}{\alpha\sigma^2} - \rho_S(I_{D,j} - I_D\lambda_j). \tag{43}$$

Forwards are in zero net supply, with market-clearing forward price

$$F = \mu - \rho_S(\mu - 1 - \eta\zeta_D). \tag{44}$$

F depends on investor M's disutility parameter ζ_D but not on sensitivity to the externality λ_L or λ_M. The motive for trade in forwards is not that L is less sensitive to the externality than M, but rather that M has nonpecuniary disutility from investment in firm D. Without such disutility, forwards would not trade, because the investors can already adjust their exposure to $\tilde{\mu}_D$ by adjusting their investment in firm D.

Investment by each investor and average per-capita investment are

$$I_{D,L} = \frac{(1 + (\lambda_L - \lambda_M)\eta)(\mu - 1 - \rho_S(\mu - F)) - \lambda_L\eta\zeta_D}{(1 - \tau_L - \tau_M)\alpha\sigma^2(1 - \rho_S^2)}, \tag{45}$$

$$I_{D,M} = \frac{(1 + (\lambda_M - \lambda_L)(1 - \eta))(\mu - 1 - \rho_S(\mu - F)) - T_L\zeta_D}{(1 - \tau_L - \tau_M)\alpha\sigma^2(1 - \rho_S^2)}, \tag{46}$$

$$I_D = \frac{\mu - 1 - \eta\zeta_D}{(1 - \tau_L - \tau_M)\alpha\sigma^2}. \tag{47}$$

With forwards, individual portfolios change, but aggregate dirty investment stays the same.

Proposition 5 *For a given $\eta \in (0, 1)$, when investor M has nonpecuniary disutility from* *firm D investment, introducing the forward contract decreases $I_{D,M}$ relative to the case without the forward, and increases $I_{D,L}$ relative to the case without the forward. Both differences increase in η. Introducing the forward does not change average per-capita investment I_D.*

Proof. The difference between $I_{D,M}$ with versus without forwards is Equation (46) minus Equation (38), which reduces to

$$\frac{-\rho_S^2 \zeta_D (1-\eta)}{\alpha\sigma^2(1-\rho_S^2)} < 0,$$

given $\rho_S \neq 0$, and $\eta \in (0,1)$. The expression is linearly increasing in η. The difference between $I_{D,L}$ with versus without forwards is Equation (45) minus Equation (37), which reduces to

$$\frac{\rho_S^2 \zeta_D \eta}{\alpha\sigma^2(1-\rho_S^2)} > 0,$$

and is also linearly increasing in η. The difference between I_D with versus without forwards is Equation (47) minus Equation (39), which is zero when $\rho = 0$, as assumed for this section.

■

In Appendix C, we instead introduce a green bond that gives investor M positive non-pecuniary utility, and that has a pecuniary payoff positively correlated with firm C. This captures the idea that a subset of the clean firm's activities might be marketed as helping the environment. Analogous to Proposition 5, introducing the green bond doesn't change aggregate clean investment I_C. Purely financial green innovations may trade and yet not affect the externality in equilibrium.

However the forward does alter *who* invests in the polluting firm: environmentalists (type M) invest relatively less in polluters, whereas the remaining investors (type L) invest relatively more. The greater the environmentalist's share of the population, the "less effective" the green alternative is, as $I_{D,M}$ falls only slightly with the introduction of the forward when η is large. This parallels the result that the forward is neutral with respect to aggregate investment I_D: the forward only enables environmentalists to disinvest in equilibrium to the extent that non-environmentalists are available to take the other side of the forward contract, which they in turn hedge by increasing investment in the polluting firm.

Our forward contract is an abstraction, but we have in mind carbon emissions futures contracts as an example. The long-term relationship between carbon emissions allowance (EA) prices and the stock returns of heavy CO_2 emitters is potentially complicated and has not, to our knowledge, been studied extensively.[14] However Aatola et al. (2013) documents a significant and negative relationship between the price of coal and the price of EU carbon EAs, whereas the comparatively clean alternative natural gas price shows a significant and positive relationship with EA prices. Therefore $\rho_S > 0$ is plausible and we set $\rho_S = 0.8$ for our numerical example. Remaining parameter values are in Table 1. Proposition 5 requires only $\rho_S \neq 0$.[15]

Figure 3 illustrates the magnitude of the effects described in Proposition 5 with our example parameters. Relative to the case with disutility only, introducing the forward flips the sign of the relationship between η and $I_{D,M}$. When $\eta \approx 0$, investor M invests almost nothing in firm D, whereas investor L invests far more than M in firm D. Both $I_{D,L}$ and $I_{D,M}$ are sharply increasing in η. The implications for individual investment are particularly different from Figure 1, where M correctly views disinvestment by his type as more impactful as η increases, and so $I_{D,M}$ is decreasing in η, the opposite of Figure 3.

Include Figure 3 about here.

Trade and price changes in the forward market change individual investment, as illustrated in Figure 4. For $\eta \in (0, 1)$, investor L sells forwards, whereas investor M buys them. As η increases, the forward price F rises, and L sells more forwards to M. However L hedges his increased short position by investing more in firm D, which has returns positively correlated with $\tilde{S}$, and hence negatively correlated with a short forward position. While investor

[14]For a detailed characterization of emissions permit prices and related futures contracts separately from emitter stock prices, see Hitzemann and Uhrig-Homburg (2018).

[15]The sign of ρ_S affects which investors take long or short positions in the forward, but does not affect productive investment. The extent to which $I_{D,L}$ and $I_{D,M}$ are impacted by introduction of the forward is sensitive to ρ_S^2, and would be small for $\rho_S \approx 0$.

L sells more forwards on a per-capita basis, his share of the population is declining, so from the perspective of investor M the forward market is drying up as $\eta \to 1$. Eventually the forward price rises to the point where M views it as too costly an alternative to investment in firm D, so $f_M = 0$, and $I_{D,M}$ peaks, as previously illustrated in Figure 3.[16]

Include Figure 4 about here.

6 Related Evidence and Conclusion

Several studies report an exclusion premium for dirty stocks, including Baker et al. (2018); Bolton and Kacperczyk (2021); Engle et al. (2020); Hsu et al. (2022); Larcker and Watts (2020). The hedging channel in our model increases the demand for dirty assets relative to clean assets, reducing the exclusion premium. To capture the exclusion effect, we modify our endowment economy to include the possibility that the M investors with fraction η of the population are constrained from investing in the dirty stock D. Let $\lambda_L \tilde{X}$ be the disutility from pollution for the fraction $1 - \eta$ unconstrained from investing in the dirty stock.

In the resulting equilibrium, we measure the exclusion premium as the difference in prices between the clean and dirty stocks assuming a tax rate of zero:

$$P_C - P_D = (1 - \rho^2)\alpha\sigma^2 \frac{\eta}{1 - \eta} - (1 - \rho)\alpha\sigma^2(1 + \eta\rho)\lambda_L. \tag{48}$$

The first-term in Equation (48) is the exclusion premium with no hedging effect. The second term is strictly decreasing in λ_L and measures the importance of the hedging effect to the exclusion premium.

The exclusion premium is 0 when there is no hedging effect and all investors hold the dirty stock $\eta = 0$. To get a sense of the importance of the hedging effect, assume $\eta = 25\%$ of investors are constrained, and the remaining fraction of investors $1 - \eta$ are unaffected by the

[16]Forward positions f_L and f_M are defined on a per-capita basis. Adjusted for population shares, type L always sells as many forwards as type M buys.

externality i.e. $\lambda_L = 0$. Remaining parameters are $\rho = 0$, $\alpha = 3$ and $\sigma = 15\%$. In this case, the exclusion effect is 2.25%. The coefficient multiplying λ_L is 0.0675: $\lambda_L = 1\%$ results in a 0.07% drop in the exclusion premium. With all other parameters the same and increasing ρ to 0.5, the exclusion effect is 2.25% at $\lambda_L = 0$, and drops by 0.05% at $\lambda_L = 1\%$.

If unconstrained investors have more extreme sensitivity, the hedging channel has a larger effect. For example, if $\lambda_L = 5\%$ then the hedging channel reduces the exclusion premium by about 0.34% with $\rho = 0$ and 0.19% with $\rho = 0.5$. Overall, the hedging channel reduces the exclusion effect, but not by a large amount unless the the externality has a large impact on the utility of the unconstrained investors.

Nevertheless, the existence of the hedging channel highlights that demand for exclusion from environmentally conscious investors is more at odds with conventional theory than previously recognized. Therefore we consider two nonstandard countervailing factors — coordination among environmentalists, and nonpecuniary disutility from dirty investment — to explain demand for SRI. Each explanation has distinct implications for public policy and for the social value of green financial innovation.

Macroscopically, coordination to reduce externalities can be ruled out if the disinvestment initiative is limited to a small minority of the wealth-adjusted population. In our model, the population share of coordinating investors must be sufficiently large for their actions to have substantial aggregate impact, otherwise they will choose little or no disinvestment.

According to the Forum for Sustainable and Responsible Investment, in excess of 25% of US professionally managed assets were under some form of SRI or ESG restriction in 2018, with climate change the most commonly cited issue. A 2018 survey by Pew Research found that 59% of the US population viewed climate change as a major threat. Comparable statistics are larger for some other regions. For example, in Europe around 48% of managed assets were under ESG/SRI restrictions in 2018, and EU country views on climate change as a major threat ranged from 55% for Poland to 90% for Greece. Coordinated action to reduce the externality via disinvestment seems viable.

A distinguishing macroscopic feature of coordination is that the amount of disinvestment should be superlinear in the population share of environmentalists, whereas nonpecuniary motives should yield an approximately linear relationship. Although available data is inadequate to distinguish between these hypotheses, US data shows no significant relationship between SRI investment share and views on the environment. Gallup asks a broad panel of US respondents whether environmental protection should be prioritized, potentially at the cost of economic growth, or whether economic growth should be prioritized, potentially at the expense of the environment.[17]

Figure 5 plots Gallup survey results alongside investment shares of SRI from 1995-2018, when overlapping data is available. Although the SRI share of investment and support for prioritizing the environment have both increased since 2012, support for prioritizing the environment was much higher in the past, and SRI investment did not decline with support for environmental protection leading up to 2010.[18]

Include Figure 5 about here.

At a more granular level, surveys and experiments with individual investors shed some light on the motives for SRI. A series of papers study attitudes and behaviors of retail investors regarding SRI (Bauer and Smeets, 2015; Riedl and Smeets, 2017; Anderson and Robinson, 2019; Bauer et al., 2021). This work most strongly supports nonpecuniary utility from SRI investment.

However there is also survey and experimental evidence consistent with coordination. Riedl and Smeets (2017) find that social preferences predict the decision to invest in SRI funds more strongly than signaling. Their measure of social preferences is more closely aligned with coordination than nonpecuniary utility. Bauer et al. (2021) study an explicit mechanism for enforcing coordination among a pension plan's participants, who vote in support of a more

[17] https://news.gallup.com/poll/1615/environment.aspx Respondents may also voluntarily state equal priority or no opinion, but comparatively few do so.

[18] We also examined potential relationships across countries using statistics from the GSIA and Pew Research, but overlapping data is extremely limited. There is no correspondence between percentages viewing climate change as a major threat and SRI investment shares in 2018, but some positive correspondence when climate change is ranked relative to other potential threats.

stringent SRI management strategy. Public institutions constituted over 50% of SRI assets under management in 2018 (USSIF, 2018), so such coordination mechanisms may apply broadly if public pension plans are accountable to voters.

Both coordination and nonpecuniary utility mechanisms are plausibly at play in practice. Given the proliferation of green bonds and the growth of SRI and ESG investment funds, more work is required to distinguish between these motives, because they have different implications for the efficacy of such investment. Coordinating investors are concerned with ends, i.e., a reduction in negative externalities. Financial innovation oriented towards such investors will therefore serve societal goals such as the reduction of negative externalities. For example, Oehmke and Opp (2020) show that green bonds with technology covenants can help coordinating investors to achieve socially desirable outcomes. In contrast, our examples show that financial innovations motivated by nonpecuniary utility may find a large market without facilitating reduction of negative externalities. If the needs of coordinating investors can be served by tailored financial instruments, their development should be emphasized.

References

Aatola, P., Ollikainen, M., and Toppinen, A. (2013) Price determination in the eu ets market: Theory and econometric analysis with market fundamentals *Energy Economics* 36, 380–395.

Allen, K. (2018) Strict US market rules limit corporate sellers of green bonds *Financial Times* URL `https://www.ft.com/content/baa217c4-157c-11e8-9376-4a6390addb44`.

Anderson, A. and Robinson, D. T. (2019) Knowledge, fear and beliefs: Understanding household demand for green investments *Swedish House of Finance Research Paper* .

Baker, M., Bergstresser, D., Serafeim, G., and Wurgler, J. (2018) Financing the response to climate change: The pricing and ownership of us green bonds Technical report National Bureau of Economic Research.

Barber, B. M., Morse, A., and Yasuda, A. (2021) Impact investing *Journal of Financial Economics* 139, 162–185.

Barrage, L. (2019) Optimal Dynamic Carbon Taxes in a Climate–Economy Model with Distortionary Fiscal Policy *The Review of Economic Studies* 87, 1–39.

Bauer, R., Ruof, T., and Smeets, P. (2021) Get Real! Individuals Prefer More Sustainable Investments *The Review of Financial Studies* 34, 3976–4043.

Bauer, R. and Smeets, P. (2015) Social identification and investment decisions *Journal of Economic Behavior & Organization* 117, 121–134.

Bolton, P. and Kacperczyk, M. (2021) Do investors care about carbon risk? *Journal of Financial Economics* 142, 517–549.

Brav, A. and Heaton, J. (2021) Brown assets for the prudent investor *Harv. Bus. L. Rev. Online* 12, 1.

Chava, S. (2014) Environmental externalities and cost of capital *Management Science* 60, 2223–2247.

Engle, R. F., Giglio, S., Kelly, B., Lee, H., and Stroebel, J. (2020) Hedging climate change news *The Review of Financial Studies* 33, 1184–1216.

Fama, E. F. and French, K. R. (2007) Disagreement, tastes, and asset prices *Journal of Financial Economics* 83, 667–689.

GSIA (2018) Global sustainable investment review Technical report Global Sustainable Investment Alliance (GSIA) URL `http://www.gsi-alliance.org/`.

Heinkel, R., Kraus, A., and Zechner, J. (2001) The effect of green investment on corporate behavior *Journal of Financial and Quantitative Analysis* 36, 431–449.

Hitzemann, S. and Uhrig-Homburg, M. (2018) Equilibrium price dynamics of emission permits *Journal of Financial and Quantitative Analysis* 53, 1653–1678.

Hong, H. and Kacperczyk, M. (2009) The price of sin: The effects of social norms on markets *Journal of Financial Economics* 93, 15–36.

Hong, H. and Kostovetsky, L. (2012) Red and blue investing: Values and finance *Journal of Financial Economics* 103, 1–19.

Hsu, P.-H., Li, K., and Tsou, C.-Y. (2022) The pollution premium *Journal of Finance, Forthcoming* .

Humphrey, J., Kogan, S., Sagi, J. S., and Starks, L. T. (2020) The asymmetry in responsible investing preferences Technical report SSRN.

Landier, A. and Lovo, S. (2021) Socially responsible finance: How to optimize impact? .

Larcker, D. F. and Watts, E. M. (2020) Where's the greenium? *Journal of Accounting and Economics* 69, 101312.

Michel, P. and Rotillon, G. (1995) Disutility of pollution and endogenous growth *Environmental and Resource Economics* 6, 279–300.

Morewedge, C. K., Tang, S., and Larrick, R. P. (2018) Betting your favorite to win: Costly reluctance to hedge desired outcomes *Management Science* 64, 997–1014.

Nordhaus, W. D. (1991) To slow or not to slow: the economics of the greenhouse effect *The Economic Journal* 101, 920–937.

Nordhaus, W. D. (2015) Climate clubs: Overcoming free-riding in international climate policy *American Economic Review* 105, 1339–70.

Nordhaus, W. D. et al. (1992) The 'dice' model: Background and structure of a dynamic integrated climate-economy model of the economics of global warming Technical report Cowles Foundation for Research in Economics, Yale University.

Oehmke, M. and Opp, M. M. (2020) A theory of socially responsible investment Technical report CEPR.

Pástor, L., Stambaugh, R. F., and Taylor, L. A. (2021) Sustainable investing in equilibrium *Journal of Financial Economics* 142, 550–571.

Pedersen, L. H., Fitzgibbons, S., and Pomorski, L. (2021) Responsible investing: The esg-efficient frontier *Journal of Financial Economics* 142, 572–597.

Pigou, A. C. (1920) The economics of welfare, 1920 *McMillan&Co., London* .

Pindyck, R. S. (2013) Climate change policy: What do the models tell us? *Journal of Economic Literature* 51, 860–72.

Poushter, J. and Huang, C. (2019) Climate change still seen as the top global threat, but cyberattacks a rising concern Technical report Pew Research Center URL https://www.pewresearch.org/global/2019/02/10/climate-change-still-seen-as-the-top-glob

Riedl, A. and Smeets, P. (2017) Why do investors hold socially responsible mutual funds? *The Journal of Finance* 72, 2505–2550.

Roth Tran, B. (2019) Divest, disregard, or double down? philanthropic endowment investments in objectionable firms *American Economic Review: Insights* 1, 241–56.

Stein, C. M. (1981) Estimation of the mean of a multivariate normal distribution *The Annals of Statistics* 9, 1135–1151.

Stokey, N. L. (1998) Are there limits to growth? *International Economic Review* 39, 1–31.

USSIF (2018) Report on socially responsible investing trends in the united states Technical report US Forum for Sustainable and Responsible Investment (USSIF) URL https://www.ussif.org/currentandpast.

Xepapadeas, A. (2005) Economic growth and the environment *Handbook of Environmental Economics* 3, 1219–1271.

Zerbib, O. D. (2019) The effect of pro-environmental preferences on bond prices: Evidence from green bonds *Journal of Banking & Finance* 98, 39–60.

Data Availability Statement

The datasets we use are available in the public domain.

The data underlying this article are available from:

- The US SIF (Sustainable Investment Forum, aka The Forum for Sustainable and Responsible Investment) at https://www.ussif.org/currentandpast, (USSIF (2018))
- biannual reviews of the GSIA (Global Sustainable Investment Alliance) available at http://www.gsi-alliance.org, (GSIA (2018))
- Investment Company Institute and Strategic Insight Simfund, available at https://data.worldbank.org/indicator/CM.MKT.LCAP.CD?locations=US,
- Gallup: https://news.gallup.com/poll/1615/environment.aspx,
- Pew Research, "Climate Change Still Seen as the Top Global Threat, but Cyberattacks a Rising Concern" at https://www.pewresearch.org/global/2019/02/10/climate-change-still-seen-as-the-top-global-threat-but-cyberattacks-a-rising-concern/ (Poushter and Huang (2019))

A Comparative Statics for Portfolios with Generalized Preferences and Environmental Damage Functions

Our comparative static results for portfolios in the basic setting — without coordination or nonpecuniary utility — reflect a hedging motive that is quite general. Here we show similar results for more general preferences and damage functions.

To provide a general analysis, we solve for portfolios as a function of net return characteristics, with net returns defined as

$$\tilde{r}_i = \frac{\tilde{Y}_i - P_i}{P_i},\ i \in \{C, D\},$$

in the endowment economy or

$$\tilde{r}_i = \tilde{\mu}_i - 1,\ i \in \{C, D\},$$

in the production economy.

In addition to the two risky assets, investors may also trade a risk-free asset with interest rate normalized to zero.

We assume the following regarding risky asset returns.

Assumption 1 (Risky asset returns) *Returns are jointly normally distributed and uncorrelated:*

$$\begin{bmatrix} \tilde{r}_C \\ \tilde{r}_D \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mu_C \\ \mu_D \end{bmatrix}, \begin{bmatrix} \sigma_C^2 & 0 \\ 0 & \sigma_D^2 \end{bmatrix} \right),$$

with strictly positive expectations and standard deviations.

The assumption of uncorrelated returns simplifies the proofs. In the endowment economy, the mean and standard deviation of returns on the stocks are endogenous, depending on equilibrium prices. Our modified formulation allows for this feature of the endowment economy.

However the case of independent and identically distributed (i.i.d.) returns for both firms provides a useful point of reference in some instances, and allows for additional results in others. Therefore while we allow for heterogeneous returns in most of our derivations that follow, we specialize to the following assumption in some propositions.

Assumption 1-alt (i.i.d. risky asset returns) *Returns for each firm are identically and independently normally distributed with mean $\mu_C = \mu_D = \mu > 0$, and standard deviation $\sigma_C = \sigma_D = \sigma$.*

The assumption of competitive behavior and the definition of the externality as in Section 2 allow us reduce environmental damage to a function of firm D returns. In the endowment economy, $\tilde{X} = (1 + \tilde{r}_D)P_D$, whereas in the production economy $\tilde{X} = (1 + \tilde{r}_D)I_D$. Investors take P_D (respectively I_D) as given, hence for purposes of analyzing the portfolio choice of an atomistic investor we can redefine environmental damage as a function of $\tilde{r}_D$. Each investor $j \in \{L, M\}$ has environmental damage function $f_j(\tilde{r}_D)$.

For the moment we place no further restrictions on the damage functions, but introduce the following definition.

Definition 1 (Sensitivity to environmental damage) *Investor j is sensitive to environmental damage if $f'_j(r_D) > 0$, $\forall r_D$.*

Subsequently we consider how L might be less sensitive to environmental damage than M, depending upon the characteristics of their respective damage functions.

We allow for general preferences, although some results require that investors have identical preferences up to heterogeneous damage functions.

Assumption 2 (Preferences) *Investors have identical preferences over terminal wealth net of environmental damage, $U(W_j - f_j(r_D))$, with $U'(.) > 0$ and $U''(.) < 0$.*

Each investor of type $j \in \{L, M\}$ solves

$$\max_{\theta_C,\theta_D} \mathbb{E}[U(\tilde{W}_j - f_j(\tilde{r}_D))], \tag{A1}$$
$$\text{subject to } \tilde{W}_j = \theta_{C,j}\tilde{r}_C + \theta_{D,j}\tilde{r}_D.$$

The formulation above implicitly assumes that the domain of U is the real line, since returns are normally distributed. The definition of wealth can be augmented to include a constant endowment without materially altering our remaining derivations. For sufficiently large endowments and under technical conditions, our results would extend in the usual way to utility with a domain on the positive real line.

The first order condition (FOC) with respect to $\theta_{i,j}$ is

$$\mathbb{E}\left[U'(\tilde{W}_j - f_j(\tilde{r}_D))\tilde{r}_i\right] = 0. \tag{A2}$$

Given Assumption 1, optimal $\theta_{C,j}$ satisfies

$$\begin{aligned} 0 &= \mathbb{E}\left[U'\left(\tilde{W}_j - f_j(\tilde{r}_D)\right)\right]\mu_C + cov\left(U'\left(\tilde{W}_j - f_j(\tilde{r}_D)\right), \tilde{r}_C\right) \\ &= \mathbb{E}\left[U'\left(\tilde{W}_j - f_j(\tilde{r}_D)\right)\right]\mu_C + \mathbb{E}\left[U''\left(\tilde{W}_j - f_j(\tilde{r}_D)\right)\theta_{C,j}\sigma_C^2\right], \\ \Rightarrow \theta_{C,j} &= \frac{-\mathbb{E}\left[U'\left(\tilde{W}_j - f_j(\tilde{r}_D)\right)\right]\mu_C}{\mathbb{E}\left[U''\left(\tilde{W}_j - f_j(\tilde{r}_D)\right)\right]\sigma_C^2}, \end{aligned} \tag{A3}$$

where the second line uses Stein's Lemma.[19]

[19] See Stein (1981), Lemma 2 and discussion regarding application to arbitrary normal random variables, pages 1136-1137.

Similarly, optimal $\theta_{D,j}$ satisfies

$$
\begin{aligned}
0 &= \mathbb{E}\left[U'\left(\tilde{W}_j - f_j(\tilde{r}_D)\right)\right]\mu_D + cov\left(U'\left(\tilde{W}_j - f_j(\tilde{r}_D)\right), \tilde{r}_D\right) \\
&= \mathbb{E}\left[U'\left(\tilde{W}_j - f_j(\tilde{r}_D)\right)\right]\mu_D + \mathbb{E}\left[U''\left(\tilde{W}_j - f_j(\tilde{r}_D)\right)\left(\theta_{D,j}\sigma_D^2 - f_j'(\tilde{r}_D)\sigma_D^2\right)\right], \\
\Rightarrow \theta_{D,j} &= \frac{-\mathbb{E}\left[U'\left(\tilde{W}_j - f_j(\tilde{r}_D)\right)\right]\mu_D}{\mathbb{E}\left[U''\left(\tilde{W}_j - f_j(\tilde{r}_D)\right)\right]\sigma_D^2} + \frac{\mathbb{E}\left[U''\left(\tilde{W}_j - f_j(\tilde{r}_D)\right)f_j'(\tilde{r}_D)\right]}{\mathbb{E}\left[U''\left(\tilde{W}_j - f_j(\tilde{r}_D)\right)\right]}.
\end{aligned}
\tag{A4}
$$

Let $\theta_{i,j}^*$ and $\tilde{W}_j^*$ denote optimal portfolios and wealth distributions, respectively. Our first result follows directly from the above optimality conditions.

Proposition A1 *Assume i.i.d. returns as in Assumption 1-alt and preferences as in Assumption 2. Then any investor who is sensitive to environmental damage invests more in firm D than in firm C.*

Proof.

$$
\theta_{D,j}^* - \theta_{C,j}^* = \frac{\mathbb{E}\left[U''\left(\tilde{W}_j^* - f_j(\tilde{r}_D)\right)f_j'(\tilde{r}_D)\right]}{\mathbb{E}\left[U''\left(\tilde{W}_j^* - f_j(\tilde{r}_D)\right)\right]} > 0, \tag{A5}
$$

since utility is concave and $f_j'(r_D) > 0$. ■

Proposition A1 does not require that investors L and M have identical utility functions U.

An obvious implication of Proposition A1 is that Propositions 2 and 4 are not reliant on the preference or damage function specifications assumed in the main text. The crucial assumptions are that at least some investors in the economy are sensitive to environmental damage whereas no investors have $f_j'(r_D) < 0$, and that the characteristics of firms C and D are symmetrical. Aggregate investment $I_D > I_C$ in the production economy (Proposition 4) with i.i.d. returns, whereas in the endowment economy the cost of capital for firm D must fall to $\mu_D < \mu_C$ in order to clear markets (Proposition 2) if the two firms instead have i.i.d. dividends.

By contrast it is more difficult to show under general conditions that investors who are more sensitive to environmental damage always invest relatively more in firm D than investors who are less sensitive, along the lines of Propositions 1 and 3. We differentiate investor sensitivity to environmental damage as follows.

Assumption 3 (Relative Sensitivity to Environmental Damage) *Investor M is more sensitive to environmental damage than investor L in that the damage functions satisfy condition*

$$\frac{\mathbb{E}\left[U''\left(\tilde{W}_M^* - f_M(\tilde{r}_D)\right) f_M'(\tilde{r}_D)\right]}{\mathbb{E}\left[U''\left(\tilde{W}_M^* - f_M(\tilde{r}_D)\right)\right]} > \frac{\mathbb{E}\left[U''\left(\tilde{W}_L^* - f_L(\tilde{r}_D)\right) f_L'(\tilde{r}_D)\right]}{\mathbb{E}\left[U''\left(\tilde{W}_L^* - f_L(\tilde{r}_D)\right)\right]}. \tag{A6}$$

Assumption 3 can be interpreted as a requirement that, in the neighborhood of their respective optimal portfolios, the standardized change in marginal utility has greater covariance with marginal environmental damage for investor M than it does for investor L. We present two categorical examples of model primitives that satisfy the assumption.

Example 1 *Suppose investor L is insensitive to environmental damage, $f_L'(r_D) = 0$, $\forall r_D$, whereas investor M is strictly sensitive to environmental damage, $f_M'(r_D) > 0$, $\forall r_D$. Then Assumption 3 is satisfied.*

Proof. Since Assumption 2 requires $U''(.) < 0$ everywhere, Equation (A6) is satisfied due to the conditions on the derivatives of the damage functions. ■

Although the example above is stark in that L is totally insensitive to environmental damage, we place few restrictions on M's damage function. The environmental damage function $f_M(r_D)$ may exhibit a "tipping point" or otherwise reflect disaster risk such that the probability density of environmental damage is fat tailed, even though the externality itself is normally distributed.

Example 1 does not accommodate the framework of the main text, however, wherein we assume linear damage functions but allow both investors to be sensitive to the externality. This is covered by the second example.

Example 2 *Suppose damage functions are linear such that* $f_M'(r_D) = \lambda_M > \lambda_L = f_L'(r_D), \forall r_D$. *Then Assumption 3 is satisfied.*

Proof. Equation (A6) is satisfied because

$$\frac{\mathbb{E}\left[U''\left(\tilde{W}_M^* - f_M(\tilde{r}_D)\right)\lambda_M\right]}{\mathbb{E}\left[U''\left(\tilde{W}_M^* - f_M(\tilde{r}_D)\right)\right]} - \frac{\mathbb{E}\left[U''\left(\tilde{W}_L^* - f_L(\tilde{r}_D)\right)\lambda_L\right]}{\mathbb{E}\left[U''\left(\tilde{W}_L^* - f_L(\tilde{r}_D)\right)\right]} = \lambda_M - \lambda_L > 0. \tag{A7}$$

■

An analogue of Proposition 3 follows.

Proposition A2 *Under Assumptions 1-alt, 2, and 3, investors of type* M*, who are hurt most by pollution, invest relatively more in polluting firm* D *than do investors of type* L*, who are hurt less by pollution, in the sense that*

$$\theta_{D,M}^* - \theta_{C,M}^* > \theta_{D,L}^* - \theta_{C,L}^*. \tag{A8}$$

Proof. Per Equations (A3) and (A4) under Assumption 1-alt,

$$\theta_{D,j}^* - \theta_{C,j}^* = \frac{\mathbb{E}\left[U''\left(\tilde{W}_j^* - f_j(\tilde{r}_D)\right)f_j'(\tilde{r}_D)\right]}{\mathbb{E}\left[U''\left(\tilde{W}_j^* - f_j(\tilde{r}_D)\right)\right]}.$$

Then the result follows directly from Assumption 3. ■

Although similar in spirit, above we prove a difference in differences, rather than the difference in levels shown in Proposition 3. In addition, the above result does not directly extend Proposition 1, because it is shown under i.i.d. returns, which will not hold in equilibrium in an endowment economy with i.i.d. *dividends*.

It is possible to prove a difference in ratios while allowing for heterogeneous returns, showing an analogue of Proposition 1, but with a more restrictive assumption on damage functions as in Example 1.

Proposition A3 *Under Assumptions 1, 2, and damage functions as in Example 1, investors of type M, who are hurt most by pollution, invest relatively more in polluting firm D than do investors of type L, who are hurt less by pollution, in the sense that*

$$\frac{\theta^*_{D,M}}{\theta^*_{C,M}} > \frac{\theta^*_{D,L}}{\theta^*_{C,L}}. \tag{A9}$$

Proof. Per Equations (A3) and (A4),

$$\frac{\theta^*_{D,j}}{\theta^*_{C,j}} = \frac{\mu_D \sigma^2_C}{\mu_C \sigma^2_D} + \frac{\mathbb{E}\left[-U''\left(\tilde{W}^*_j - f_j(\tilde{r}_D)\right) f'_j(\tilde{r}_D)\right]\sigma^2_C}{\mathbb{E}\left[U'\left(\tilde{W}^*_j - f_j(\tilde{r}_D)\right)\right]\mu_C}.$$

Then per Assumption 2 and Example 1, we have

$$\frac{\mathbb{E}\left[-U''\left(\tilde{W}^*_M - f_M(\tilde{r}_D)\right) f'_M(\tilde{r}_D)\right]\sigma^2_C}{\mathbb{E}\left[U'\left(\tilde{W}^*_M - f_M(\tilde{r}_D)\right)\right]\mu_C} > \frac{\mathbb{E}\left[-U''\left(\tilde{W}^*_L - f_L(\tilde{r}_D)\right) f'_L(\tilde{r}_D)\right]\sigma^2_C}{\mathbb{E}\left[U'\left(\tilde{W}^*_L - f_L(\tilde{r}_D)\right)\right]\mu_C} = 0.$$

■

A.1 More General Utility Functions

We now show how demand for the dirty asset would change if the investor had a more general form of the utility function.

Each investor has initial wealth normalized to W_0, and can in invest two risky assets with characteristics as in Assumption 1-alt. The correlation between $\tilde{X}$ and $\tilde{r}_C$ is zero and the correlation between $\tilde{X}$ and $\tilde{r}_D$ is ρ_X.

Each investor's optimization problem is

$$\max_{\theta^j_C, \theta^j_D} E\left[U^j\left(\tilde{W}, \tilde{X}\right)\right] \quad \text{s.t. } \tilde{W} = W_0 + \sum_{i=C,D} \theta^j_i \tilde{r}_i. \tag{A10}$$

The first-order conditions for portfolio optimization are

$$E\left[U_W^j\left(\tilde{W},\tilde{X}\right)\tilde{r}_i\right]=0,\ i=C,D. \tag{A11}$$

Using the multivariate version of Stein's Lemma, the first order conditions are

$$\begin{aligned}E\left[U_W^j\left(\tilde{W},\tilde{X}\right)\tilde{r}_C\right]&=EU_W^j\left(\tilde{W},\tilde{X}\right)E\left[\tilde{r}_C\right]+cov\left(U_W^j\left(\tilde{W},\tilde{X}\right),\tilde{r}_i\right)\\&=EU_W^j\left(\tilde{W},\tilde{X}\right)\mu+cov(\tilde{W},\tilde{r}_c)EU_{WW}^j(\tilde{W},\tilde{X})+cov(\tilde{X},\tilde{r}_C)EU_{WX}^j(\tilde{W},\tilde{X})\\&=EU_W^j\left(\tilde{W},\tilde{X}\right)\mu+\theta_C^j\sigma^2EU_{WW}^j\left(\tilde{W},\tilde{X}\right)\\&=0,\end{aligned} \tag{A12}$$

and

$$\begin{aligned}E\left[U_W^j\left(\tilde{W},\tilde{X}\right)\tilde{r}_D\right]&=EU_W^j\left(\tilde{W},\tilde{X}\right)\mu+\theta_D^j\sigma^2EU_{WW}^j\left(\tilde{W},\tilde{X}\right)+\rho_X\sigma_X\sigma E\left[U_{WX}^j\left(\tilde{W},\tilde{X}\right)\right]\\&=0,\end{aligned} \tag{A13}$$

where σ_X is the standard deviation of the externality.

Subtracting the (A12) from (A13) and rearranging,

$$\theta_D^j-\theta_C^j=-\rho_X\frac{\sigma_X\sigma E\left[U_{WX}^j\left(\tilde{W},\tilde{X})\right)\right]}{\sigma^2E\left[U_{WW}^j\left(\tilde{W},\tilde{X}\right)\right]}. \tag{A14}$$

Since $U_{WW}^j\left(\tilde{W},\tilde{X}\right)<0$

$$sign\left(\theta_D^j-\theta_C^j\right)=sign\left(\rho_X\left(E\left[U_{WX}^j\left(\tilde{W},\tilde{X}\right)\right]\right)\right). \tag{A15}$$

In our baseline model in the body of the manuscript, $\rho_X=1$ meaning that the effect of the externality on the holdings of the risky asset depends on the effect of the externality on the investor's marginal utility of wealth. Following Michel and Rotillon (1995) and Xepapadeas (2005), we consider three cases.

Case 1: Separability

$$U_{WX}(W, X) = 0. \tag{A16}$$

Here investors would hold the same quantity of clean or dirty assets: $\theta^j_C = \theta^j_D$. The separability assumption holds for separable utility functions, with the utility function

$$U(W, X) = u(W) - v(X), \tag{A17}$$

with $u_W > 0,\ u_{WW} < 0,\ and\ v_X(X) \geq 0.$

A special case of separability is where the investor is indifferent to the externality, or $v(X) = 0$. See Stokey (1998) or Barrage (2019) for examples of environmental models using separability.

Case 2: The Distaste Effect

$$U_{WX}(W, X) < 0. \tag{A18}$$

The marginal utility of wealth decreases in the externality to cause a relative distaste for wealth. A useful framework is the multiplicative form,

$$U(W, X) = u(W)\phi(X), \tag{A19}$$

with the externality reducing environmental quality:

$$U_X(W, X) = u(W)\phi_X(X) < 0. \tag{A20}$$

We also assume $u_W(W) > 0,\ u_{WW}(W) < 0$, and $\phi(X) > 0$.

Utility exhibits the distaste effect only if, in addition

$$U_{WX}(W,X) = u_W(W)\phi_X(X) < 0. \tag{A21}$$

In the asset pricing literature, it is common to assume risk-averse utility over a Cobb-Douglas composite good. This can be seen as an example in the multiplicative form illustrating when the distaste effect prevails, and when it does not.[20]

Adapting this idea to our setting and assuming CRRA utility over the composite good, let

$$U(W,X) = \frac{(WX^{-\beta})^{1-\gamma}}{1-\gamma}, \tag{A22}$$

with $WX^{-\beta}$ for $\beta > 0$ interpreted as a composite good, strictly increasing in W and strictly decreasing in X.

This example fits the multiplicative form with $u(W) = \frac{W^{1-\gamma}}{1-\gamma}$ and $\phi(X) = X^{\beta(\gamma-1)}$, since

$$\begin{aligned} &U_W(W,X) = W^{-\gamma}X^{\beta(\gamma-1)} > 0,\ U_{WW}(W,X) = -\gamma W^{-(\gamma+1)}X^{\beta(\gamma-1)} < 0, \\ &\text{and } U_X(W,X) = -\beta W^{1-\gamma}X^{\beta(\gamma-1)-1} < 0. \end{aligned} \tag{A23}$$

The cross derivative is

$$U_{WX}(W,X) = \beta(\gamma-1)W^{-\gamma}X^{\beta(\gamma-1)-1}. \tag{A24}$$

Since $sign(U_{WX}(W,X)) = sign(\gamma - 1)$, the utility function exhibits the distaste effect only if the coefficient of relative risk-aversion is less than one.

Michel and Rotillon (1995) show in an endogenous growth model with pollution that positive growth cannot be sustained in the long term in an economy with preferences exhibiting the distaste effect.

[20] Since our prevailing assumption is X and W normally distributed, we take some leeway here in assuming that the probability of negative realizations is sufficiently small that this utility specification is valid.

When the investor's preferences satisfy the distaste effect, the investor will invest less in the dirty firm than in the clean firm:

$$\theta_D^j < \theta_C^j. \tag{A25}$$

Case 3: The Compensation Effect

$$U_{WX}(W, X) > 0. \tag{A26}$$

The marginal utility of wealth increases in the externality. A prominent example of these preferences is the damage function widely used in the global warming literatures and that we use in the main body of the paper and discuss in the first part of the Appendix. The externality acts to destroy an investor's wealth

$$U(W, X) = u(W - f(X)),\ u'(W - f(X)) > 0, u''(W - f(X)) < 0, f'(X) > 0, \tag{A27}$$

implying that

$$U_W(W, X) = u'(W - f(X)) > 0,\ U_{WX}(W, X) = -u''(W - f(X))f'(X) > 0. \tag{A28}$$

A second example of preferences satisfying the compensation effect is the composite good utility function in (A22) with a coefficient of relative risk aversion greater than one.

In both examples,

$$\theta_D^j > \theta_C^j. \tag{A29}$$

Michel and Rotillon (1995) show in a endogenous growth model with pollution that such preferences are consistent with long run growth in an endogenous growth economy.

B Model Solution when Both Types of Investor Internalize Pollution

Section 4.3 solves a model in which investors of type M internalize their collective contribution to pollution, but investors of type L still treat pollution as an externality. This section summarizes the model solution under the alternative assumption that both L and M internalize.

Investment in the polluter is

$$I_{D,L} = \frac{(1+\eta(\lambda_L - \lambda_M))(\mu - \rho(\mu-1)) - \frac{T_M}{T_L} - \frac{\eta\lambda_L}{T_M}}{(1-\tau_L-\tau_M)\alpha\sigma^2(1-\rho^2)} \tag{B1}$$

$$I_{D,M} = \frac{(1+(1-\eta)(\lambda_M - \lambda_L))(\mu - \rho(\mu-1)) - \frac{T_L}{T_M} - \frac{(1-\eta)\lambda_M}{T_L}}{(1-\tau_L-\tau_M)\alpha\sigma^2(1-\rho^2)}. \tag{B2}$$

Relative to the case where only investor M internalizes, both investors invest less in firm D if investor L also internalizes. $I_{D,L}$ is reduced by

$$\frac{T_M\left(\frac{1}{T_L}-1\right)}{(1-\tau_L-\tau_M)\alpha\sigma^2(1-\rho^2)},$$

while $I_{D,M}$ is reduced by

$$\frac{\lambda_M(1-\eta)\left(\frac{1}{T_L}-1\right)}{(1-\tau_L-\tau_M)\alpha\sigma^2(1-\rho^2)}.$$

Investor M's investment decision reflects less demand to hedge the externality, because the externality is less severe once L internalizes. Provided $\eta < 1$ and $\lambda_L > 0$, the reductions are strictly positive.

Investment in the clean firm is

$$I_{C,L} = \frac{\mu - 1 - \rho(\mu - \frac{1}{T_L})}{\alpha\sigma^2(1-\rho^2)}, \tag{B3}$$

$$I_{C,M} = \frac{\mu - 1 - \rho(\mu - \frac{1}{T_M})}{\alpha\sigma^2(1-\rho^2)}, \tag{B4}$$

for investors L and M, respectively. In general, internalizing pollution only affects investment in the clean firm when $\rho \neq 0$. For example, if $\rho > 0$, then internalizing increases investment in firm C, because the investors invest less in positively correlated firm D.

C Green Financial Innovation: Bonds

Green bonds are intended to finance or refinance projects with small or positive environmental impact, including energy efficient buildings, renewable energy, etc. Legal requirements for green bonds are still evolving, and vary by jurisdiction (Allen, 2018). A subset of investors prefer environmentally friendly projects, and a green certification allows easy identification of related investments. We model this preference in investor M as positive nonpecuniary utility from green bond holdings.

Define risky green bond payoff $\tilde{V}$, which has correlation ρ_V with $\tilde{\mu}_C$, the clean firm's productivity. We have in mind $\rho_V > 0$, since the clean firm also uses technology without environmental impact. However we show that ρ_V is irrelevant for aggregate productive investment, counter to intuition.

P_V is the price of the bond, which is set to clear the market with

$$(1-\eta)v_L + \eta v_M = 0, \tag{C1}$$

where v_L and v_M are positions of investors L and M, respectively. Green bonds are in zero net supply: they are a pure financial innovation, not a new productive technology.

Equivalent to the section introducing the forward contract, we assume in this section that $\rho = 0$ (the clean and dirty firms' productivity are uncorrelated with each other), $\tilde{V}$ is independent of $\tilde{\mu}_D$, and $\tilde{V} \sim \mathcal{N}(\mu, \sigma)$.

Investor M solves:

$$\max_{I_{C,M}, I_{D,M}, v_M} \mathbb{E}[-\exp\{-\alpha(I_{D,M}(\tilde{\mu}_D - 1) + I_{C,M}(\tilde{\mu}_C - 1) + v_M(\tilde{V} + \zeta_C - P_V) - \lambda_M \tilde{X} - \zeta_D I_{D,M})\}]. \tag{C2}$$

In Section 5, investing in firm D caused M nonpecuniary disutility $-\zeta_D < 0$, whereas trading forwards didn't, even though forwards were a statistical substitute for such investment. In this section, investing in firm C produces no nonpecuniary utility for investor M, whereas investing in green bonds produces positive nonpecuniary utility $\zeta_C > 0$.

Investor L solves the analogous problem without nonpecuniary utility of any sort from investment holdings,

$$\max_{I_{C,L}, I_{D,L}, v_L} \mathbb{E}[-\exp\{-\alpha(I_{D,L}(\tilde{\mu}_D - 1) + I_{C,L}(\tilde{\mu}_C - 1) + v_L(\tilde{V} - P_V) - \lambda_L \tilde{X})\}]. \tag{C3}$$

Because we assume $\rho = 0$, the solutions for $I_{D,M}$ and $I_{D,L}$ are unaffected by introduction of the green bond. The interesting question is how firm C investment will be affected. Investment in the clean firm, and positions in the green bond, are

$$I_{C,L} = \frac{\mu - 1 - \rho_V(\mu - P_V)}{\alpha\sigma^2(1 - \rho_V^2)}, \tag{C4}$$
$$I_{C,M} = \frac{\mu - 1 - \rho_V(\mu + \zeta_C - P_V)}{\alpha\sigma^2(1 - \rho_V^2)}, \tag{C5}$$
$$v_L = \frac{\mu - P_V - \rho_V(\mu - 1)}{\alpha\sigma^2(1 - \rho_V^2)}, \tag{C6}$$
$$v_M = \frac{\mu + \zeta_C - P_V - \rho_V(\mu - 1)}{\alpha\sigma^2(1 - \rho_V^2)}. \tag{C7}$$

Imposing the market clearing condition pins down the bond's price,

$$P_V = \mu + \eta\zeta_C - \rho_V(\mu - 1). \tag{C8}$$

Substituting the solution for the equilibrium price into the investment and bond positions gives reduced expressions in terms of the model parameters,

$$I_{C,L} = \frac{(1-\rho_V^2)(\mu-1)+\rho_V\eta\zeta_C}{\alpha\sigma^2(1-\rho_V^2)}, \tag{C9}$$

$$I_{C,M} = \frac{(1-\rho_V^2)(\mu-1)-\rho_V(1-\eta)\zeta_C}{\alpha\sigma^2(1-\rho_V^2)}, \tag{C10}$$

$$v_L = \frac{-\eta\zeta_C}{\alpha\sigma^2(1-\rho_V^2)}, \tag{C11}$$

$$v_M = \frac{(1-\eta)\zeta_C}{\alpha\sigma^2(1-\rho_V^2)}. \tag{C12}$$

The above results highlight several features of the green bond. First, trade in the green bond is motivated purely by investor M's nonpecuniary utility ζ_C, as $v_L = v_M = 0$ for $\zeta_C = 0$. Provided $\zeta_C > 0$, i.e., the environmentalist likes to hold green bonds, then M takes a long position in the bond, and L takes a short position. When the bonds are highly correlated with firm C productivity ($|\rho_V| \approx 1$), more bonds trade, because the investors are able to hedge their bond exposure by trading firm C's equity.

In the most relevant case where $\rho_V > 0$, investor L's trade can be thought of as refinancing investment in firm C by issuing green bonds: he invests ρ_V in firm C per unit of green bond sold. To compensate investor L for residual unhedged risk, a portfolio that shorts the green bond and goes long firm C pays a positive spread relative to investment in firm C,

$$E[(\tilde{\mu}_C - 1) - (\tilde{V} - P_V)] = (1-\rho_V)(\mu-1) + \eta\zeta_C, \tag{C13}$$

that is decreasing in the amount of hedged risk. In the limiting case where $\rho_V \uparrow 1$, the spread is $\eta\zeta_C$: the joint surplus created when investor L intermediates investment in firm C, by issuing green bonds that generate nonpecuniary utility for investor M, is split according to the population shares of each investor.

It is also evident in the limiting case that investor L increases firm C investment one-for-one with his green bond issues, whereas investor M decreases firm C investment one-for-one with green bond purchases. In fact green bonds have no effect on aggregate green investment for any ρ_V, since

$$I_C = (1-\eta)I_{C,L} + \eta I_{C,M} = \frac{\mu - 1}{\alpha\sigma^2}, \tag{C14}$$

which is unchanged from the case without the green bond in Equation (19) with $\rho = 0$. Therefore introduction of green bonds may improve welfare, by labeling investment in a way that makes certain investors feel good, but equilibrium investment is unchanged from the case without green bonds.

A small but growing literature on green bonds studies ownership and yields of such bonds, which are labeled or certified to fund projects with relatively positive environmental impact. Zerbib (2019) studies the yield differential between green bonds and synthetically constructed conventional equivalents, and finds that green bonds trade at a small but significant premium on average, equivalent to a 2 basis point (bps) reduction in green bond yields. Zerbib (2019) emphasizes the use of a broad variety bonds compliant with a set of Green Bond Principles developed by the International Capital Market Association, and also provides a useful review of related literature.[21] Baker et al. (2018) focus primarily on US green municipal bonds, finding a somewhat larger yield differential of 6 bps for green bonds. They also find that ownership of green bonds is more concentrated than ownership of conventional equivalents, which they suggest as evidence that a subset of investors has a preference for green bonds.

Both Zerbib (2019) and Baker et al. (2018) attribute the green bond premium to nonpecuniary utility among green investors, the theoretical mechanism for which Baker et al. (2018) sketch in a simple model along the lines of Fama and French (2007).

Our study has implications for the interpretation of their findings. We show that investors with nonpecuniary utility for the green bond will be the exclusive (long) holders of those assets, and that equilibrium expected returns of the green bond are lower than direct investment in the clean technology, which can be viewed as the equivalent conventional as-

[21] Not all studies find that green bonds trade at a premium. Zerbib (2019) and Baker et al. (2018) note the importance of carefully controlling for liquidity and tax status when estimating yield differentials. However Larcker and Watts (2020) argue there is no significant green bond premium after appropriate controls.

set. That is, our results are consistent with empirical findings. However introducing green bonds has no impact on productive investment or the severity of the externality. Therefore findings such as a green bond premium or concentrated green bond ownership do not directly support the conclusion that green bonds have any positive environmental impact. Furthermore, although we associate nonpecuniary utility with sensitivity to externalities, there is no theoretical linkage between the two: the mechanisms are separate.

In contrast, if investors did not have nonpecuniary utility but instead coordinated to reduce the externality, then there would be no demand for green bonds as a purely financial innovation. Clean investments would be relatively concentrated among environmentalists, but the labeling of financial instruments as green or otherwise would have no separate bearing on the concentration of holdings and no impact on relative returns. Hence the absence of return spreads or concentrated holdings of bonds *labeled* as green does not preclude highly effective coordination by investors to reduce environmental externalities.

That green bonds would serve no role for coordinating investors in our simple example does not preclude a positive role in reality. However, new channels would be required. Given that coordinating investors are motivated by ends, i.e., meaningful reductions in negative externalities, innovations such as green bonds would be of social value if they helped coordinating investors to accomplish their objectives, whereas their social benefits are less clear if they serve only to generate nonpecuniary utility. Future work should develop such theoretical mechanisms and test their empirical significance.

	Parameter	Value
Mean gross productivity of investment	μ	1.1
Standard deviation of productivity	σ	0.15
Correlation, productivity of firms C and D	ρ	0
Coefficient of absolute risk aversion	α	3
Sensitivity to externality, agent L	λ_L	0.01
Sensitivity to externality, agent M	λ_M	0.03

Table 1: Parameter values. The table reports the baseline parameter values used in our numerical examples. Our parameter values are broadly consistent with the aggregate stock market, and with a reasonable utility loss from the externality.

Section	Frictionless capital adjustment	Pigouvian taxation	Coordination	Nonpecuniary utility	Derivatives
2		✓			
3	✓	✓			
4.1	✓	✓			
4.2	✓	✓			
4.3	✓		✓		
4.4	✓			✓	
5	✓			✓	✓
B	✓		✓		
C	✓			✓	✓

Table 2: Summary of model variants. High level characteristics of the model variants considered in the paper are summarized by section number in the table above.

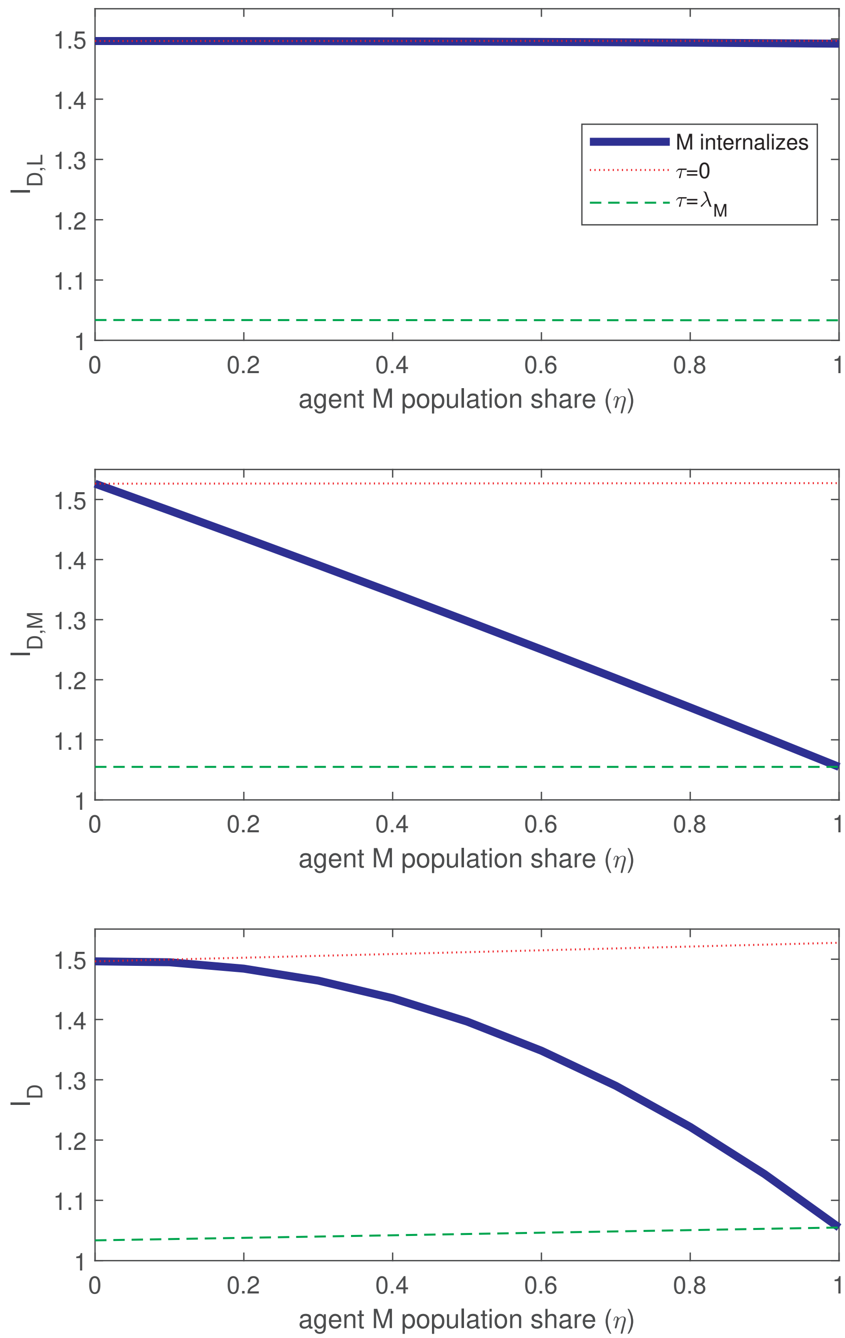

Figure 1: Investment: effect when M internalizes. The figure shows investment in dirty firm D by type L (top), type M (middle), and as a per-capita average (bottom), as a function of type M population share η. We compare three cases: without Pigouvian taxation, with Pigouvian taxation at fixed rate $\tau = \lambda_M$, and when type M internalizes, taking account of the externality generated by the collective investment of type M.

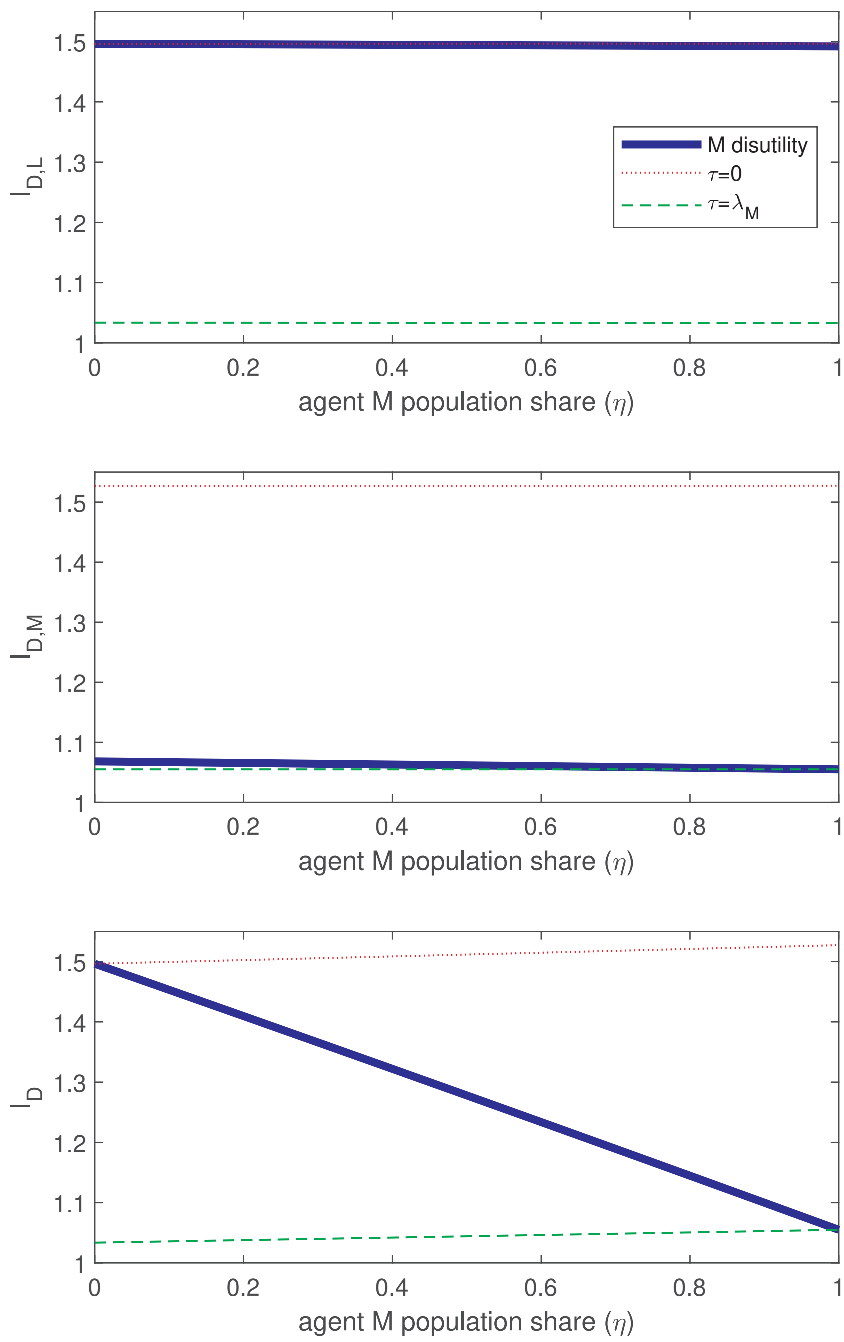

Figure 2: Investment: effect when M has disutility. The figure shows investment in dirty firm D by type L (top), type M (middle), and as a per-capita average (bottom), as a function of type M population share η. We compare three cases: without Pigouvian taxation, with Pigouvian taxation at fixed rate $\tau = \lambda_M$, and when type M has nonpecuniary disutility over investment in firm D with parameter $\zeta_D = \frac{\lambda_M}{1-\lambda_M}$.

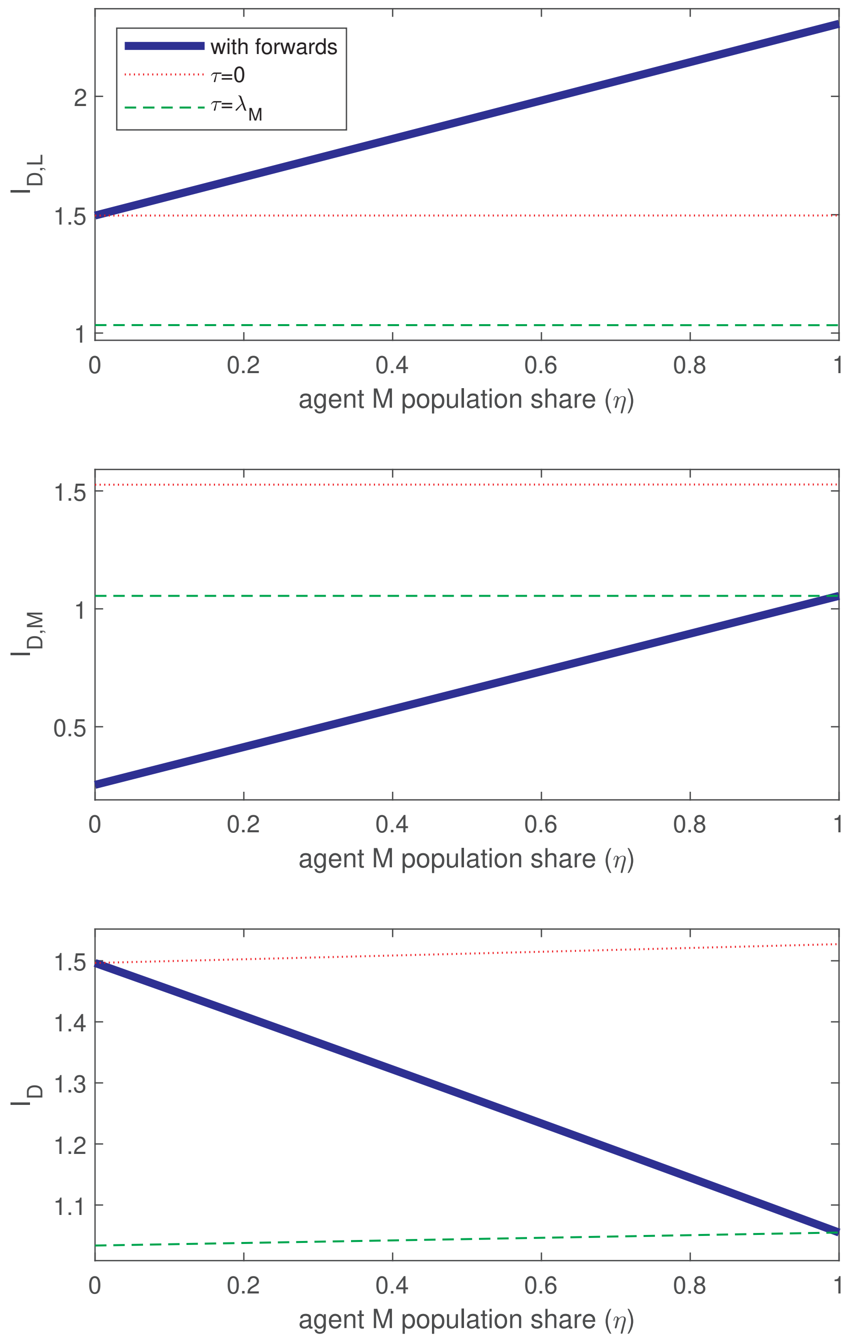

Figure 3: Investment: effect with forward trading. The figure shows investment in dirty firm D by type L (top), type M (middle), and as a per-capita average (bottom), as a function of type M population share η. We compare three cases: without Pigouvian taxation, with Pigouvian taxation at fixed rate $\tau = \lambda_M$, and when type M has nonpecuniary disutility over investment in firm D with parameter $\zeta_D = \frac{\lambda_M}{1-\lambda_M}$, but may trade an alternative forward contract without suffering disutility. Correlation of the forward's payoff with the externality is $\rho_S = 0.8$.

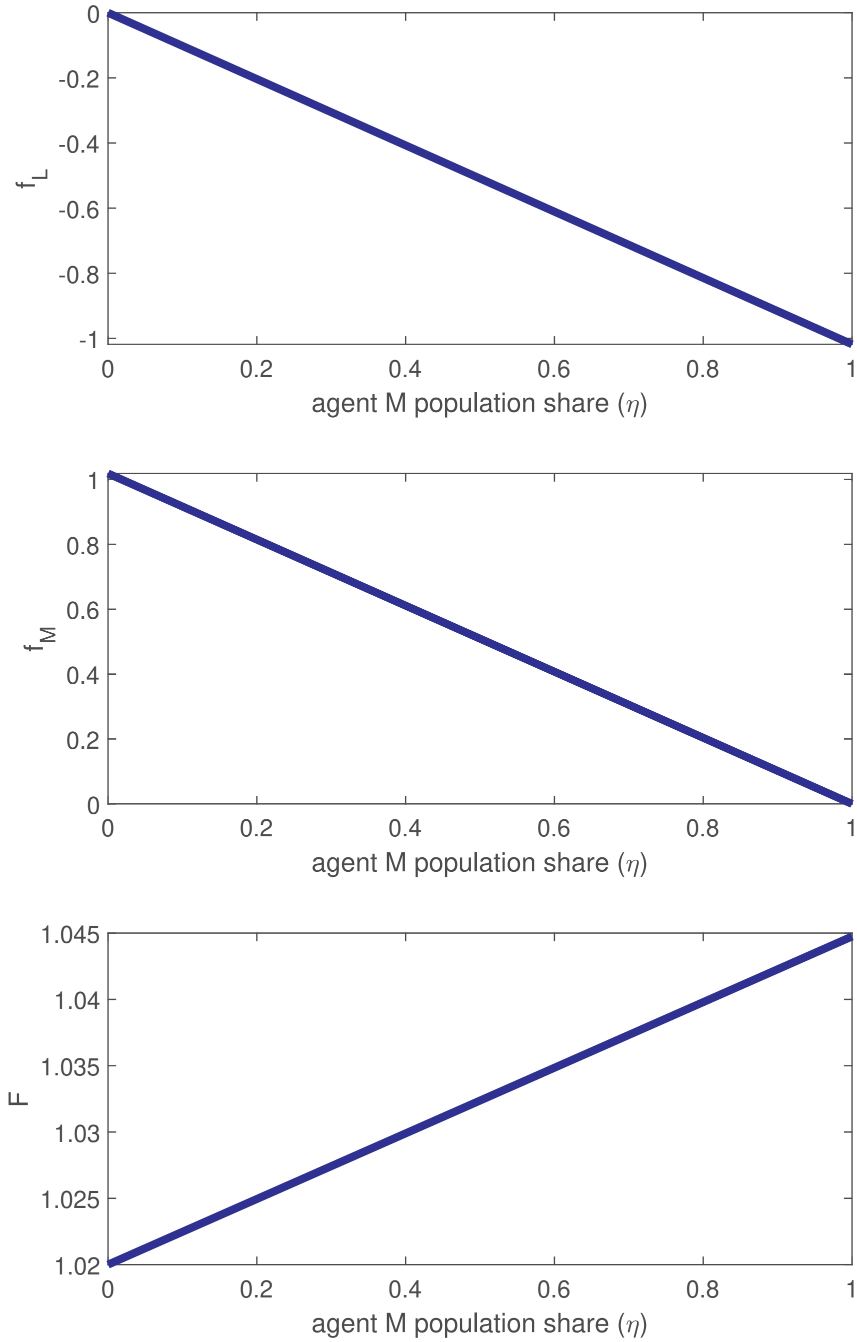

Figure 4: Forwards. The figure shows forward positions of type L (top) and type M (middle), and the forward price (bottom), as a function of type M population share η. Type M has nonpecuniary disutility over investment in firm D with parameter $\zeta_D = \frac{\lambda_M}{1-\lambda_M}$, but may trade an alternative forward contract without suffering disutility. Correlation of the forward's payoff with the externality is $\rho_S = 0.8$.

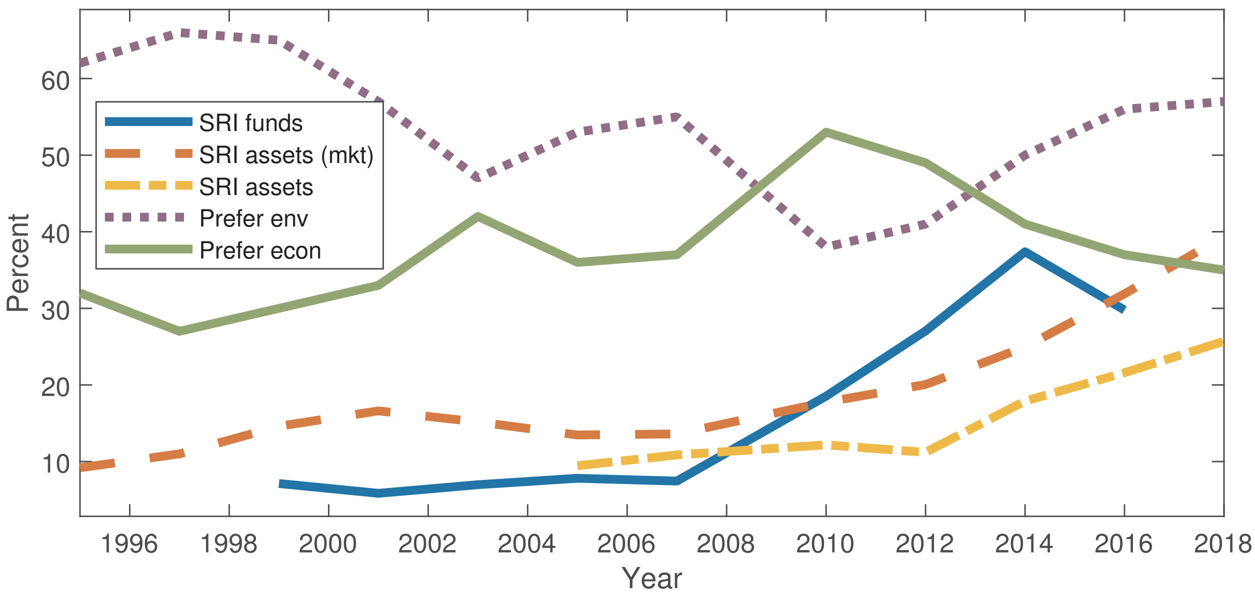

Figure 5: SRI versus environmental opinions, US. SRI funds is SRI fund assets under management (AUM) reported by the Forum for Sustainable and Responsible Investment (USSIF) as a percentage of total AUM for US investment funds, as measured by the Investment Company Institute. SRI assets is the percentage of all professionally managed assets under SRI restrictions, from the USSIF. SRI assets (mkt) is an alternative measure of total SRI share based on AUM reported by the USSIF relative to total US equity market capitalization from the World Bank. Prefer env is the share of respondents to Gallup surveys answering that environmental protection should be prioritized over economic growth, whereas prefer econ is the share of respondents with the opposite preference.