The Greenwashing Rat Race: Endogenous Information Degradation from ESG Score Manipulation

[Author Name]*

April 4, 2026

Abstract

When firms manipulate ESG scores using an imprecise technology, aggregate manipulation endogenously degrades score informativeness for all firms. Each firm's greenwashing adds idiosyncratic noise to its score; in a symmetric equilibrium, the market strips out the average manipulation but cannot remove the noise. Social welfare is hump-shaped in the fraction of ESG-mandated capital: some ESG demand is beneficial, but excessive demand triggers a wasteful manipulation arms race that destroys the information on which capital allocation depends. In a two-type extension with green and brown firms, the externality falls disproportionately on genuinely green firms, who lose the ability to signal their true quality. A Pigouvian tax on manipulation has a double dividend, and optimal enforcement intensity increases with ESG demand. The fundamental Sharpe ratio of green-minus-brown portfolios declines as ESG demand grows; the total Sharpe ratio declines when the cash-flow cost of pollution is large relative to the taste parameter. The results characterize optimal ESG demand, anti-greenwashing regulation, and the interaction between mandatory disclosure and carbon taxes.

JEL Classification: G11, G14, G18, D82, Q58
Keywords: ESG investing, greenwashing, informational externality, green finance, sustainable investing, manipulation

*[Affiliation]. Email: [email]. For helpful comments, I thank [acknowledgments].

1 Introduction

Over $35 trillion in assets are now managed under ESG mandates, yet the evidence on whether sustainable investing improves environmental outcomes remains mixed. ESG rating agencies disagree substantially on which firms are green (Berg et al., 2022). Firms respond to ESG pressure by divesting pollutive assets rather than reducing pollution, and the buyers simply continue the same operations (Duchin et al., 2025). Sustainable investing may even be counterproductive: brown firms become browner when their cost of capital rises, while green firms do not become much greener when their cost of capital falls (Hartzmark and Shue, 2023). A natural question arises: does ESG investor demand improve capital allocation, or does it trigger a corporate response that undermines the very information on which allocation depends?

ESG demand can be self-defeating through a specific mechanism. When firms can manipulate their ESG scores through cosmetic actions (asset reclassification, selective disclosure, supply chain restructuring), and these actions are inherently imprecise, aggregate manipulation degrades the informativeness of ESG scores for all firms. The degradation is an informational externality: each firm's manipulation adds idiosyncratic noise to its own score, but the resulting loss of signal quality affects the entire market's ability to distinguish genuinely green firms from brown ones. No individual firm internalizes the effect of its manipulation on the aggregate information environment. The result is a rat race in which all firms manipulate, manipulation has zero average effect on capital allocation (the market strips out the mean), but the noise remains and destroys information.[1]

The model features a continuum of firms, each choosing genuine abatement (which reduces pollution) and manipulation (which improves ESG scores without reducing pollution). Manipulation is imprecise: higher manipulation effort generates proportionally more noise in the firm's ESG score. Green capital flows to firms with better scores, with the allocation governed by a Bayesian updating coefficient that depends on aggregate noise. In the unique symmetric equilibrium, the updating coefficient falls as ESG demand rises, because more demand triggers more manipulation, which adds more noise. Social welfare is hump-shaped in ESG demand. At low levels, ESG capital directs investment toward greener firms and corrects an uninternalized pollution externality. At high levels, the manipulation rat race dominates: waste mounts, information degrades, and green capital becomes poorly allocated.

The unique symmetric equilibrium has score informativeness declining in ESG demand

[1]The term "rat race" is used here to describe a zero-sum arms race rather than the dynamic escalation in Akerlof (1976). The mechanism is closer to a prisoner's dilemma with an informational twist: each firm manipulates because unilateral deviation is costly, and the aggregate outcome is socially wasteful. The static model does not feature the dynamic escalation that the original usage implies.

(Proposition 3). Equilibrium manipulation exceeds the social optimum because firms ignore the informational externality (Proposition 4). A two-type extension with green and brown firms and heterogeneous manipulation costs yields cross-sectional predictions: brown firms greenwash more, and the informational externality falls disproportionately on genuinely green firms, who lose the ability to signal their true quality (Proposition 5). Social welfare is hump-shaped in ESG demand, with a closed-form optimal demand level for the benchmark case (Proposition 6). A Pigouvian tax on manipulation has a double dividend: it saves waste and restores information quality, which amplifies genuine abatement (Proposition 8). Mandatory disclosure interacts with carbon taxation in a specific way: combining the two instruments causes over-abatement, requiring a downward adjustment in the carbon tax (Proposition 10). Optimal anti-greenwashing enforcement intensity increases with ESG demand and with manipulation imprecision (Proposition 11). The fundamental Sharpe ratio of green-minus-brown portfolios declines as ESG demand grows; the total Sharpe ratio (including the taste premium) declines when the cash-flow cost of pollution is large relative to the taste parameter (Proposition 13).

In the literature on ESG and asset prices, Pástor et al. (2021) show that green assets earn lower expected returns from taste-based demand, and Berk and van Binsbergen (2025) calibrate the cost-of-capital channel and find it is quantitatively small. Goldstein et al. (2022) show that heterogeneous ESG preferences can reduce price informativeness through a demand-side mechanism: investors with different ESG tastes trade in opposite directions on the same signal. The channel here is different and operates on the supply side: firms strategically add noise to their own scores through imprecise manipulation, and the aggregate noise degrades informativeness even when all investors agree on how to interpret scores. The two channels are likely complements: demand-side disagreement and supply-side noise amplify each other when both are present (Section 6).

In the literature on greenwashing and ESG market design, Leippold et al. (2025) model greenwashing as degrading trust and information quality in a dynamic framework with self-reinforcing equilibrium traps. The complementary static Bayesian model here yields closed-form equilibria, an explicit welfare optimum, and asset pricing predictions. Cayirli (2025) shows that ESG demand can backfire when the cost curvature of genuine action determines whether firms switch to greenwashing; the mechanism there operates through the "washing set" rather than through endogenous noise degradation. Cartellier et al. (2023) characterize individual firms' optimal greenwashing policies in a dynamic setting but do not model the aggregate informational externality. On the rating design side, Azarmsa and Shapiro (2025) study competition between ESG rating agencies and show that strong investor demand can lead raters to generalize rather than specialize, reducing informativeness. Inderst and Opp

(2025) analyze fund-level greenwashing and the welfare case for mandatory taxonomies.

More broadly, the mechanism belongs to the literature on strategic information degradation. Frankel and Kartik (2019) study settings where senders optimally muddle information; Pérez-Richet and Skreta (2022) analyze test design when agents can game the test. The model here differs in that noise is a side effect of manipulation (not a strategic choice) and the externality operates through a market-level sufficient statistic (λ) that aggregates across a continuum of agents. The ESG application matters because the welfare stakes ($35 trillion in AUM) and the policy instruments (the EU Corporate Sustainability Reporting Directive, SEC climate disclosure rules, anti-greenwashing regulation) are first-order. The results hold for any noise function that is increasing in manipulation effort (Appendix C); the proportional specification serves as a tractable special case.

Section 2 presents the model, including the two-type extension. Section 3 derives the equilibrium and the informational externality, and characterizes the two-type equilibrium. Section 4 analyzes welfare and policy instruments, including imperfect enforcement. Section 5 derives asset pricing implications. Section 6 discusses limitations, extensions, and connections to empirical evidence. Section 7 concludes. Proofs appear in the appendix unless short enough for the main text.

2 Model

2.1 Environment

A continuum of firms indexed by $i \in [0, 1]$ operate in a single period. Firm i has true pollution θ_i, drawn independently from $N(\bar{\theta}, \sigma_\theta^2)$. Each firm privately observes its own θ_i.

Each firm simultaneously chooses two actions. **Abatement** $a_i \geq 0$ is genuine pollution reduction: the firm's true pollution becomes $\theta_i - a_i$, at a cost of $c_a a_i^2/2$. **Manipulation** $m_i \geq 0$ is cosmetic score improvement that does not reduce true pollution, at a cost of $c_m m_i^2/2$.

Noisy manipulation technology. Manipulation is imprecise. When firm i exerts manipulation effort m_i, the realized score improvement is $m_i + \nu_i$, where $\nu_i \sim N(0, \psi^2 m_i^2)$ is idiosyncratic manipulation noise, independent across firms and independent of θ_i and ε_i. The parameter $\psi > 0$ measures manipulation imprecision. Manipulation effort generates proportionally more noise: greenwashing activities (reclassifying assets, restructuring supply chains, selective disclosure) produce uncertain score improvements, and the gap between intended and realized improvement grows with scale.

ESG scores. Firm i's publicly observable ESG score is

$$S_i = \theta_i - a_i - m_i - \nu_i + \varepsilon_i, \tag{1}$$

where $\varepsilon_i \sim N(0, \sigma_\varepsilon^2)$ is exogenous measurement noise. Lower S_i corresponds to a better (greener) score. The score cannot distinguish abatement from manipulation, and the manipulation noise ν_i adds variance that degrades informativeness.

2.2 Green Capital Allocation

A mass $\mu > 0$ of green capital is allocated across firms based on ESG scores. Green investors observe only scores and use Bayesian updating. In the normal-normal framework, the optimal linear allocation rule assigns green capital

$$K_i = \mu\big[\kappa - \lambda(S_i - \bar{S})\big], \tag{2}$$

where λ is the Bayesian updating coefficient (derived below), $\bar{S}$ is the expected score, and κ is a normalization constant ensuring $\int_0^1 K_i \, di = \mu\kappa$. Firms with lower (greener) scores receive more green capital. Each unit of green capital provides a marginal benefit $\eta > 0$ to the firm (e.g., cheaper financing, favorable contracts, regulatory goodwill).

True pollution $\theta_i - a_i$ imposes social damage at rate $\delta_S > 0$ per unit. Each firm privately internalizes a fraction $\delta \in [0, \delta_S]$ of the damage cost (e.g., through carbon taxes or regulatory penalties).

2.3 Bayesian Updating and the Informational Externality

In a symmetric equilibrium where all firms choose abatement a^* and manipulation m^*, the ESG score of firm i is $S_i = \theta_i - a^* - m^* - \nu_i + \varepsilon_i$, where $\nu_i \sim N(0, \psi^2(m^*)^2)$. The market knows the equilibrium strategies and strips out the constants. The informative component of the score is

$$S_i + a^* + m^* = \theta_i + \varepsilon_i - \nu_i,$$

which is a noisy signal of θ_i with total idiosyncratic noise variance

$$\sigma_n^2(m^*) \equiv \sigma_\varepsilon^2 + \psi^2(m^*)^2. \tag{3}$$

Standard normal-normal Bayesian updating gives the posterior expectation $\mathbb{E}[\theta_i \mid S_i] =$

$\bar{\theta} + \lambda(m^*) \cdot (\theta_i - \bar{\theta} + \varepsilon_i - \nu_i)$, where the **Bayesian updating coefficient** is

$$\lambda(m) \equiv \frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_\varepsilon^2 + \psi^2 m^2}. \tag{4}$$

The informational externality is visible in equation (4). Each firm's equilibrium manipulation m^* enters the noise variance $\sigma_n^2(m^*)$ and hence $\lambda(m^*)$. When all firms increase manipulation by $\mathrm{d}m$, total noise rises by $2\psi^2 m^* \,\mathrm{d}m$, reducing λ and degrading score informativeness for every firm. No individual firm internalizes the effect: each firm is atomistic and takes λ as given.

2.4 Timing and Equilibrium Concept

The game is static and simultaneous:

1. Nature draws $\{\theta_i\}$, $\{\varepsilon_i\}$, $\{\nu_i\}$.
2. Each firm observes its θ_i and chooses (a_i, m_i).
3. ESG scores realize. Green capital is allocated. Payoffs realize.

Definition 1 (Symmetric equilibrium)**.** *A* ***symmetric equilibrium*** *consists of strategies (a^*, m^*) such that every firm optimally chooses (a^*, m^*) taking $\lambda = \lambda(m^*)$ as given, and the market's beliefs are consistent: $\lambda = \lambda(m^*)$.*

Firm i's problem. Taking the strategies of all other firms (and hence λ) as given:

$$\max_{a_i \geq 0,\, m_i \geq 0} \; \eta \cdot \mathbb{E}[K_i \mid \theta_i, a_i, m_i] - \frac{c_a}{2} a_i^2 - \frac{c_m}{2} m_i^2 + \delta \cdot a_i. \tag{5}$$

The term $\delta \cdot a_i$ captures the private benefit of abatement (avoided damage costs).

2.5 Social Welfare

The social planner values environmental outcomes at the full social cost δ_S and accounts for the allocative efficiency of green capital. Social welfare per firm is

$$W(\mu) = \delta_S \cdot a^* + \eta\mu \cdot \mathrm{AE}(\lambda) - \frac{c_a}{2}(a^*)^2 - \frac{c_m}{2}(m^*)^2 - \delta_S \bar{\theta}, \tag{6}$$

where the **allocative efficiency** $\mathrm{AE}(\lambda) \equiv \mathrm{Cov}(K_i, -(\theta_i - a^*))/\mu = \lambda\sigma_\theta^2$ measures how well green capital tracks true environmental quality (see the derivation in Appendix A).

The welfare function counts environmental outcomes only. Green investors' taste parameter (introduced in the asset pricing extension of Section 5) affects equilibrium prices

but does not enter social welfare. This convention follows the standard approach in environmental economics: welfare reflects real pollution outcomes, not investor preferences over portfolio composition.

2.6 Discussion of Modeling Choices

Why proportional noise? The specification $\text{Var}(\nu_i) = \psi^2 m_i^2$ is economically natural: larger-scale greenwashing involves more complex activities with more uncertain outcomes. Formally, the results require only that noise is increasing in manipulation effort. All main results hold for any noise function $g(m)$ with $g(0) = 0$ and $g'(m) > 0$ (Appendix C). The proportional case is a convenient special case that yields closed forms.

Why static? The static setting isolates the informational externality without learning dynamics. In a dynamic model, investors could learn about greenwashing over time, partially restoring informativeness. The static prediction is relevant when the manipulation technology evolves at least as fast as investor learning, or when new firms continuously enter.

Why a continuum? The continuum eliminates strategic interaction and maximizes the externality. With N firms, each would internalize $1/N$ of the informational effect; for concentrated industries, the policy case for intervention is correspondingly weaker.

All parameters are strictly positive: $c_a, c_m, \eta, \mu, \sigma_\theta^2, \sigma_\varepsilon^2, \psi, \delta_S > 0$, and $\delta \geq 0$.

Remark 2 (Sign convention). Lower S_i corresponds to a better (greener) score. This convention simplifies the Bayesian updating expressions but reverses the sign relative to most commercial ESG rating providers, where higher scores indicate better performance. All comparative statics should be read accordingly.

2.7 Two-Type Extension: Green and Brown Firms

The baseline model has symmetric firms. This section introduces two types to generate cross-sectional predictions about which firms greenwash more and how the informational externality distributes costs unevenly.

A fraction $\alpha \in (0, 1)$ of firms are *green* (type G), with pollution $\theta_i^G \sim N(\bar{\theta}^G, \sigma_\theta^2)$ where $\bar{\theta}^G < \bar{\theta}$, and manipulation cost c_m^G. The remaining fraction $1 - \alpha$ are *brown* (type B), with $\theta_i^B \sim N(\bar{\theta}^B, \sigma_\theta^2)$ where $\bar{\theta}^B > \bar{\theta}$, and manipulation cost c_m^B. Green firms have lower baseline pollution. The market knows the population shares but does not observe firm types directly; it observes only the ESG score S_i.

Each type $\tau \in \{G, B\}$ chooses abatement a^τ and manipulation m^τ to maximize the same objective (5), taking the aggregate updating coefficient λ as given. The noise variance is now a weighted average across types:

$$\sigma_n^2 = \sigma_\varepsilon^2 + \psi^2[\alpha(m^G)^2 + (1-\alpha)(m^B)^2]. \tag{7}$$

The updating coefficient $\lambda = \sigma_\theta^2/(\sigma_\theta^2 + \sigma_n^2)$ depends on both types' manipulation levels. Each type's first-order conditions are:

$$m^\tau = \frac{\eta\mu\lambda}{c_m^\tau}, \qquad a^\tau = \frac{\eta\mu\lambda + \delta}{c_a}, \quad \tau \in \{G, B\}. \tag{8}$$

Abatement is identical across types (it depends on λ but not on c_m^τ), while manipulation differs: firms with lower manipulation costs manipulate more. Section 3.3 derives the equilibrium and cross-sectional results.

3 Equilibrium and the Informational Externality

3.1 Equilibrium Characterization

Appendix A derives the firm's first-order conditions from a general model with type-dependent strategies. Both the abatement and manipulation first-order conditions are independent of θ_i, so the equilibrium strategies are type-independent. Lemma 14 shows this conclusion extends to arbitrary (non-linear) strategy spaces: under the Gaussian information structure, no non-linear equilibrium exists. The equilibrium first-order conditions are:

$$a^* = \frac{\eta\mu\lambda(m^*) + \delta}{c_a}, \qquad m^* = \frac{\eta\mu\lambda(m^*)}{c_m}. \tag{9}$$

The second equation defines a fixed-point problem: m^* appears on both sides through $\lambda(m^*)$.

Proposition 3 (Greenwashing Rat Race)**.** *There exists a unique symmetric equilibrium. In this equilibrium:*

(a) The equilibrium manipulation level m^* is the unique positive solution to

$$m^* = \frac{\eta\mu}{c_m} \cdot \frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_\varepsilon^2 + \psi^2(m^*)^2}. \tag{10}$$

Given m^*, abatement is $a^* = (\eta\mu\lambda(m^*) + \delta)/c_a$.

(b) *Score informativeness $\lambda(m^*)$ is strictly decreasing in ESG investor demand: $d\lambda/d\mu < 0$.*

(c) *Manipulation is strictly increasing in ESG demand: $dm^*/d\mu > 0$.*

(d) *The rat race is wasteful. Manipulation has zero average effect on capital allocation (the market strips out the mean m^*), but it imposes two costs: direct manipulation expense $c_m(m^*)^2/2$ per firm, and degraded score informativeness (reduced λ), which worsens capital allocation for all firms.*

Proof sketch. Define $f(m) \equiv \eta\mu\sigma_\theta^2/[c_m(\sigma_\theta^2 + \sigma_\varepsilon^2 + \psi^2 m^2)]$ and $h(m) \equiv m - f(m)$. Since $f(0) > 0$, we have $h(0) < 0$, and since $f(m) \to 0$ as $m \to \infty$, $h(m) \to \infty$. By the intermediate value theorem, h has at least one root. For uniqueness, $f'(m) < 0$ for all $m > 0$, so $h'(m) = 1 - f'(m) > 1 > 0$: h is strictly increasing and crosses zero exactly once. Parts (b) and (c) follow from implicit differentiation. The full proof is in Appendix A. □

The mechanism has two layers. The first is a prisoner's dilemma. Each firm gains green capital by reducing its score relative to others. In equilibrium, all firms manipulate equally, so no firm gains on average. But no firm can unilaterally stop: doing so makes its score appear worse relative to all others, causing it to lose green capital.

The second layer is the informational externality, which distinguishes the model from a standard rent-seeking contest. Because manipulation is imprecise ($\psi > 0$), each firm's manipulation adds idiosyncratic noise ν_i to its score. In aggregate, the noise degrades the signal-to-noise ratio of ESG scores for all firms. The market's ability to distinguish genuinely green firms from dirty ones deteriorates. Green capital becomes less well-allocated. When manipulation is perfectly precise ($\psi = 0$), the model reduces to a standard Tullock contest: manipulation is wasteful but does not degrade information. All of the economic interest flows from $\psi > 0$.

3.2 The Informational Externality

Proposition 4 (Excessive Manipulation)**.** *The socially optimal manipulation level m^{SP} is strictly below the equilibrium level m^*. That is, private manipulation is excessive.*

Proof. The social planner's welfare as a function of m, holding μ fixed and letting abatement respond optimally ($a = (\eta\mu\lambda(m) + \delta)/c_a$), satisfies

$$\frac{dW}{dm} = \eta\mu\lambda'(m)\left[\frac{\delta_S - \delta - \eta\mu\lambda(m)}{c_a} + \sigma_\theta^2\right] - c_m m.$$

Evaluating at $m = m^*$ and using the private first-order condition $c_m m^* = \eta\mu\lambda(m^*)$:

$$\left.\frac{dW}{dm}\right|_{m=m^*} = \underbrace{\eta\mu\lambda'(m^*)}_{<0}\underbrace{\left[\frac{\delta_S - \delta - \eta\mu\lambda(m^*)}{c_a} + \sigma_\theta^2\right]}_{>0} - \underbrace{c_m m^*}_{>0} < 0.$$

Welfare is strictly decreasing in manipulation at the private equilibrium, so the social optimum requires lower manipulation: $m^{SP} < m^*$. □

Each firm ignores two effects of its manipulation on others. First, higher aggregate manipulation reduces allocative efficiency: λ falls, and green capital becomes less correlated with true greenness. Second, lower λ weakens the green capital incentive for abatement in all firms: since $a^* = (\eta\mu\lambda + \delta)/c_a$, lower λ reduces abatement across the board. The wedge between private and social manipulation reflects the magnitude of both channels.

Comparative statics. Table 1 summarizes how the key equilibrium objects respond to parameter changes. The signs follow from implicit differentiation of the fixed-point condition (10) (see Appendix A).

Table 1: Comparative statics

Parameter	m^*	a^*	λ	Intuition
μ (ESG demand)	+	+	−	More demand raises both actions; noise degrades λ
c_m (manipulation cost)	−	+	+	Less manipulation restores informativeness
ψ (noise intensity)	−	ambig.	−	Less effective manipulation, but direct noise effect
η (green capital benefit)	+	+	−	Stronger incentive drives both actions
σ_ε^2 (exogenous noise)	−	−	−	Noisier scores reduce rewards for both actions

The comparative static $\partial a^*/\partial c_m > 0$ when $\psi > 0$ is central: making manipulation costlier increases abatement. The channel is informational. Costlier manipulation reduces m^*, which raises λ, which strengthens the green capital incentive for genuine action. With $\psi = 0$, this channel is shut down (λ is constant) and $\partial a^*/\partial c_m = 0$. The noisy manipulation model breaks the separability of abatement and manipulation decisions.

The effect of noise intensity ψ on abatement is genuinely ambiguous. Higher ψ reduces manipulation (the noisy technology is less rewarding), which tends to raise λ and hence a^*. But higher ψ also directly degrades λ by increasing noise per unit of manipulation. A formal condition for the sign follows from implicit differentiation. Write $a^* = (\eta\mu\lambda(m^*(\psi)) + \delta)/c_a$. Then

$$\frac{\partial a^*}{\partial \psi} = \frac{\eta\mu}{c_a}\frac{d\lambda}{d\psi} = \frac{\eta\mu}{c_a}\left[\lambda'(m^*)\frac{\partial m^*}{\partial \psi} + \left.\frac{\partial \lambda}{\partial \psi}\right|_m\right].$$

The first term is positive (less manipulation raises λ) and the second is negative (direct noise effect lowers λ). Abatement rises in ψ if and only if the manipulation-reduction channel dominates the direct noise channel.

3.3 Two-Type Equilibrium

With green and brown firms as defined in Section 2.7, the equilibrium exhibits differential manipulation and an informational externality that falls disproportionately on green firms.

Proposition 5 (Two-Type Equilibrium)**.** *There exists a unique equilibrium with the following properties:*

(a) If $c_m^B < c_m^G$ (brown firms find it cheaper to manipulate), then $m^B > m^G$: brown firms greenwash more.

(b) Both types choose the same abatement: $a^G = a^B = (\eta\mu\lambda + \delta)/c_a$.

(c) The informational externality falls disproportionately on green firms. Define the green premium loss as the reduction in the expected green capital that a firm of type τ receives due to aggregate manipulation noise (relative to the no-manipulation benchmark). Green firms lose more green capital than brown firms from the degradation of λ:

$$\Delta K^G \equiv \mu(\lambda_0 - \lambda)\,(\bar{\theta} - \bar{\theta}^G) > \mu(\lambda_0 - \lambda)\,(\bar{\theta} - \bar{\theta}^B) = -\Delta K^B, \tag{11}$$

where $\bar{\theta}^G < \bar{\theta} < \bar{\theta}^B$ implies $\Delta K^G > 0 > \Delta K^B$.

(d) Aggregate manipulation noise is $\psi^2[\alpha(m^G)^2 + (1-\alpha)(m^B)^2]$. If $c_m^B < c_m^G$, brown firms contribute disproportionately to the noise: the fraction of aggregate noise contributed by brown firms exceeds their population share $1 - \alpha$.

Proof. Parts (a)–(b) follow directly from the first-order conditions (8): $m^\tau = \eta\mu\lambda/c_m^\tau$, so $c_m^B < c_m^G$ implies $m^B > m^G$. Abatement does not depend on c_m^τ.

For part (c), under the no-manipulation benchmark ($m^G = m^B = 0$, $\lambda = \lambda_0$), firm i of type G receives expected green capital $\mu[\kappa - \lambda_0(\bar{\theta}^G - \bar{\theta})]$, which exceeds the population mean $\mu\kappa$ because $\bar{\theta}^G < \bar{\theta}$. Under equilibrium manipulation, the expected green capital is $\mu[\kappa - \lambda(\bar{\theta}^G - \bar{\theta})]$. The difference is $\mu(\lambda_0 - \lambda)(\bar{\theta} - \bar{\theta}^G)$, which is positive for green firms and negative for brown firms. Green firms lose the most because they had the most to gain from accurate scoring.

Part (d): brown firms contribute $\psi^2(1-\alpha)(m^B)^2$ to aggregate noise. Since $m^B > m^G$, the brown share of noise is $(1-\alpha)(m^B)^2/[\alpha(m^G)^2 + (1-\alpha)(m^B)^2] > 1 - \alpha$. □

Part (c) captures a distributional consequence absent from the symmetric model. Genuinely green firms benefit from accurate ESG scoring because accurate scores distinguish

them from the brown majority. When brown firms greenwash heavily (adding noise that degrades λ), green firms lose the ability to signal their true quality. The informational externality is regressive: the firms that pollute the most create the most noise, and the firms that pollute the least bear the largest cost.

Part (d) generates a testable cross-sectional prediction: in industries where manipulation costs are lower for dirty firms (e.g., asset-intensive industries where reclassification is easy), ESG score disagreement should be higher and the correlation between scores and true quality should be lower. The prediction distinguishes the model from symmetric-agent frameworks in which all firms contribute equally to noise.

4 Welfare and Policy

4.1 Optimal ESG Demand

Proposition 6 (Unique Welfare Maximum)**.** *Social welfare $W(\mu)$ has a unique interior maximum for all $\psi \geq 0$. Specifically:*

(a) $W'(0) > 0$ whenever $\delta_S > \delta$ or $\sigma_\theta^2 > 0$. Some ESG demand is always beneficial.

(b) $W'(\mu) < 0$ for μ sufficiently large. Excessive ESG demand is always harmful.

(c) The equilibrium manipulation product $P(\mu) \equiv c_m m^*(\mu) = \eta\mu\lambda(m^*(\mu))$ satisfies the cubic relation

$$(\sigma_\theta^2 + \sigma_\varepsilon^2)\, P + \frac{\psi^2}{c_m^2}\, P^3 = \eta\sigma_\theta^2\, \mu. \tag{12}$$

The function $P(\mu)$ is strictly increasing; it is linear for $\psi = 0$ and strictly concave on $(0, \infty)$ for $\psi > 0$.

(d) W is strictly concave as a function of P, with $d^2W/dP^2 = -(1/c_a + 1/c_m) < 0$. The sign of $W'(\mu)$ is determined by

$$\operatorname{sign}\big(W'(\mu)\big) = \operatorname{sign}\left(\sigma_\theta^2 + \frac{\delta_S - \delta}{c_a} - \left(\frac{1}{c_a} + \frac{1}{c_m}\right) P(\mu)\right). \tag{13}$$

Since P is strictly increasing from 0 to ∞, W' changes sign exactly once, giving a unique maximizer μ^*.

(e) For $\psi = 0$, W is globally strictly concave with a unique interior maximum at

$$\mu^*_{\psi=0} = \frac{\dfrac{\delta_S - \delta}{c_a} + \sigma_\theta^2}{\eta\lambda_0\left(\dfrac{1}{c_a} + \dfrac{1}{c_m}\right)}, \tag{14}$$

where $\lambda_0 = \sigma_\theta^2/(\sigma_\theta^2 + \sigma_\varepsilon^2)$.

(f) For $\psi > 0$, W *is quasi-concave (single-peaked) but not globally concave:* $W''(\mu) > 0$ *for* μ *sufficiently large. The inflection point lies strictly past the maximizer, so* W *is strictly concave on* $[0, \mu^*]$ *and on a neighborhood of* μ^*. *The marginal welfare condition at* $\mu = 0$ *is identical for all* $\psi \geq 0$ *(Proposition 22 in Appendix E).*

Proof. See Appendix A. □

Corollary 7 (Exact Welfare Characterization)**.** *Increasing ESG demand from* μ *improves social welfare if and only if* $\mu < \mu^*$, *where*

$$\mu^* = \frac{1}{\eta\sigma_\theta^2}\left[(\sigma_\theta^2 + \sigma_\varepsilon^2)\,\frac{K}{\beta} + \frac{\psi^2}{c_m^2}\left(\frac{K}{\beta}\right)^3\right], \tag{15}$$

with $K = \sigma_\theta^2 + (\delta_S - \delta)/c_a$ *and* $\beta = 1/c_a + 1/c_m$. *The optimal demand* μ^* *is strictly increasing in the uninternalized externality* $\delta_S - \delta$, *in firm heterogeneity* σ_θ^2, *and in manipulation imprecision* ψ *(for* $\psi > 0$*).*

Proof. Immediate from Proposition 6(c)–(d): the sign of $W'(\mu)$ is that of $K - \beta P(\mu)$, and P is strictly monotone, so $W'(\mu) > 0$ iff $P(\mu) < K/\beta \equiv P^*$ iff $\mu < \mu^*$. Evaluating the cubic (12) at P^* gives (15). The comparative statics follow from direct differentiation: $\partial\mu^*/\partial(K/\beta) > 0$ (the bracketed expression is increasing in P^*), and $\partial\mu^*/\partial\psi = 2\psi(K/\beta)^3/(c_m^2\eta\sigma_\theta^2) \geq 0$. □

Three forces determine the optimal ESG demand. The *externality correction* pushes μ^* up: when $\delta_S > \delta$, firms under-abate, and ESG demand provides an additional abatement incentive. The *allocative value* of green capital also pushes μ^* up: directing capital toward genuinely greener firms has social value as long as firms differ in their pollution ($\sigma_\theta^2 > 0$). Even when the externality is fully internalized ($\delta = \delta_S$), some ESG demand is beneficial for its allocative role: $\mu^*_{\psi=0} = \sigma_\theta^2/[\eta\lambda_0(1/c_a + 1/c_m)] > 0$.

The *manipulation rat race* pushes μ^* down. More ESG demand means more manipulation, higher direct costs ($c_m(m^*)^2/2$ per firm), and degraded information quality (lower λ, worse capital allocation). At the optimum, the marginal benefit of ESG demand (externality correction plus allocative value) equals the marginal cost of the rat race.

The closed-form (15) clarifies the trade-offs. The optimal demand is higher when the uninternalized externality $\delta_S - \delta$ is larger, when firm heterogeneity σ_θ^2 is larger (more scope for allocation), and when action costs c_a and c_m are higher (each unit of ESG demand triggers less manipulation waste). The cubic term $\psi^2(K/\beta)^3/c_m^2$ captures the additional margin from the manipulation rat race: noisier manipulation ($\psi > 0$) makes manipulation less effective

per unit of effort, allowing the planner to tolerate more ESG demand before the rat race cost dominates. For $\psi = 0$, the expression simplifies to (14).

4.2 Pigouvian Tax on Manipulation

Proposition 8 (Double Dividend)**.** *A tax t per unit of manipulation changes the firm's manipulation first-order condition to $m_i^*(t) = (\eta\mu\lambda - t)/c_m$ for $\eta\mu\lambda > t$, and $m^* = 0$ otherwise.*

(a) The tax $t^0 = \eta\mu\lambda_0$ eliminates all manipulation ($m^* = 0$) and restores score informativeness to λ_0.

(b) The welfare gain from eliminating manipulation has two components: the direct cost saving $c_m(m^*)^2/2$ per firm, and the allocative efficiency gain $\eta\mu[\lambda_0 - \lambda(m^*)]\sigma_\theta^2$.

(c) Abatement increases when manipulation is taxed away, because the restoration of λ amplifies the green capital incentive for genuine action. The double dividend requires $\psi > 0$; when $\psi = 0$, the tax saves waste but does not affect abatement.

Proof. See Appendix A. □

The double dividend is the central result of the noisy manipulation model. In a standard rent-seeking model ($\psi = 0$), taxing manipulation saves the direct cost but does not change abatement (the first-order conditions are separable). With $\psi > 0$, eliminating manipulation raises λ, which raises $a^* = (\eta\mu\lambda + \delta)/c_a$. The information channel creates a complementarity between anti-manipulation policy and genuine environmental improvement.

Remark 9. Proposition 8 is a theoretical benchmark that requires contractible manipulation effort. In practice, distinguishing manipulation from abatement is precisely what makes the problem hard. The tax should be interpreted as applying to specific observable greenwashing activities (e.g., cosmetic asset reclassifications, misleading disclosures). Mandatory disclosure (Proposition 10) is a more implementable policy instrument.

4.3 Mandatory Disclosure

Proposition 10 (Mandatory Disclosure)**.** *Suppose a regulator mandates verified disclosure of true pollution $\theta_i - a_i$, so that ESG scores become $\tilde{S}_i = (\theta_i - a_i) + \tilde{\varepsilon}_i$, where $\tilde{\varepsilon}_i \sim N(0, \sigma_{\tilde{\varepsilon}}^2)$ is residual measurement noise. Then:*

(a) $m^* = 0$: manipulation has no effect on scores, so firms do not manipulate.

(b) Abatement is $a^* = (\eta\mu\tilde{\lambda} + \delta)/c_a$ where $\tilde{\lambda} = \sigma_\theta^2/(\sigma_\theta^2 + \sigma_{\tilde{\varepsilon}}^2)$.

(c) All manipulation waste and manipulation-induced noise are eliminated.

(d) Mandatory disclosure combined with a full carbon tax ($\delta = \delta_S$) causes over-abatement: $a^* = (\eta\mu\tilde{\lambda} + \delta_S)/c_a > \delta_S/c_a = a^{FB}$. The first-best requires adjusting the carbon tax downward to $\delta = \delta_S - \eta\mu\tilde{\lambda}$ to offset the ESG scoring incentive.

Proof. See Appendix A. □

Part (d) reveals a policy interaction: mandatory disclosure and carbon taxes are partial substitutes for incentivizing abatement. When ESG scores are informative and green capital rewards low-pollution firms, the carbon tax needed to achieve the first-best is lower. Setting $\delta = \delta_S$ without accounting for the ESG channel leads to excessive abatement. This crowding of policy instruments is standard in the environmental economics literature on overlapping climate regulations; the ESG scoring channel provides a specific quantification of the interaction.

The result connects to ongoing policy debates. The EU Corporate Sustainability Reporting Directive (CSRD) mandates standardized pollution disclosure, and the SEC has proposed climate-related disclosure rules. The welfare rationale for mandatory disclosure here is distinct from the usual investor protection argument: mandatory disclosure eliminates the manipulation rat race and its associated informational externality (Proposition 10).

4.4 Imperfect Enforcement

Propositions 8 and 10 provide first-best benchmarks. In practice, regulators detect manipulation imperfectly. This section models imperfect enforcement and derives how optimal policy depends on the detection probability.

Suppose a regulator audits each firm's ESG report with probability $p \in [0, 1]$. If audited, the firm's manipulation m_i is detected and penalized at rate τ per unit. If not audited (probability $1 - p$), the manipulation goes undetected. The firm's expected penalty from manipulation is $p\tau m_i$. The manipulation first-order condition becomes:

$$m_i^*(p, \tau) = \frac{\eta\mu\lambda - p\tau}{c_m}, \quad \text{for} \quad \eta\mu\lambda > p\tau. \tag{16}$$

When $p\tau \geq \eta\mu\lambda_0$, the expected penalty exceeds the marginal benefit and $m^* = 0$.

Proposition 11 (Optimal Policy with Imperfect Enforcement)**.** *(a) The equilibrium manipulation level is strictly decreasing in the enforcement intensity $p\tau$.*

(b) For any detection probability $p > 0$, the penalty $\tau^*(p) = \eta\mu\lambda_0/p$ eliminates all manipulation.

(c) When the social cost of enforcement is $\kappa(p)$ per firm (with $\kappa' > 0$, $\kappa'' > 0$), the optimal enforcement intensity (p^*, τ^*) satisfies

$$\kappa'(p^*) = \left.\frac{dW}{dp}\right|_{p^*}, \tag{17}$$

where $dW/dp > 0$ for $p < p^{elim}$ (the probability that eliminates manipulation). The optimal detection probability balances the marginal enforcement cost against the marginal welfare gain from reduced manipulation and restored informativeness.

(d) The optimal enforcement intensity is higher when ESG demand μ is larger (the externality is more severe) and when manipulation imprecision ψ is larger (the informational cost of manipulation is higher).

Proof. Part (a): the fixed-point equation becomes $c_m m + p\tau = \eta\mu\lambda(m)$, or equivalently $m = [\eta\mu\lambda(m) - p\tau]/c_m$. Higher $p\tau$ shifts the right-hand side down, reducing m^* by the same intermediate value argument as Proposition 3.

Part (b): with $p\tau = \eta\mu\lambda_0$ and $m = 0$, the first-order condition gives $\eta\mu\lambda_0 - p\tau = 0$, so $m^* = 0$ is an equilibrium. Since $\lambda(0) = \lambda_0 \geq \lambda(m)$ for all m, no deviation to $m > 0$ is profitable.

Parts (c)–(d): the welfare function $W(p)$ includes the enforcement cost $\kappa(p)$. Since $dW/dp > 0$ from the manipulation reduction and $\kappa'(p) > 0$, the interior optimum exists. The comparative static $dp^*/d\mu > 0$ follows because higher μ raises the marginal welfare gain from enforcement (the externality is larger). Analogously, $dp^*/d\psi > 0$ because higher ψ increases the informational damage per unit of manipulation. □

Detection probability and penalty size are perfect substitutes for deterrence: $\tau^*(p) = \eta\mu\lambda_0/p$: a low-probability, high-penalty regime achieves the same outcome as a high-probability, low-penalty one. However, risk-neutral firms may respond differently to these regimes under limited liability, suggesting that the analysis extends naturally to a setting with penalty caps (see Appendix A for details).

The comparative static in part (d) has a direct policy implication: as ESG demand grows (increasing μ), optimal enforcement should tighten. Regulators should invest more in detecting greenwashing precisely when ESG investing is most popular, which is when the informational externality is most severe.

Remark 12 (Randomized auditing). Proposition 11 and the perturbation result (Appendix D, Proposition 17(d)) together yield a useful policy insight. Making manipulation noisier (increasing ψ) reduces equilibrium manipulation because the noisy technology is less rewarding per unit of effort. Randomized auditing, which introduces uncertainty about whether specific

manipulative actions will survive regulatory scrutiny, effectively increases ψ from the firm's perspective. Randomized auditing can therefore be welfare-improving even without directly penalizing manipulation: by raising the noise in the manipulation technology, the regulator reduces the incentive to manipulate. The optimal enforcement mix combines direct penalties ($p\tau$) with indirect deterrence through randomized scrutiny.

5 Asset Pricing Implications

The equilibrium analysis in Sections 3–4 characterizes manipulation, abatement, and welfare without specifying how ESG scores affect asset prices. The following analysis embeds the model in a one-period CAPM economy to derive equilibrium prices, expected returns, and the Sharpe ratio of ESG-sorted portfolios.

5.1 Setup

Each firm i generates a random cash flow

$$X_i = \bar{X} + bZ + \sigma_x \epsilon_i - \gamma(\theta_i - a_i), \tag{18}$$

where $Z \sim N(0,1)$ is a common market factor, $\epsilon_i \sim N(0,1)$ is firm-specific risk (independent across firms and of all other random variables), $b > 0$ is the market beta loading, $\sigma_x > 0$ is idiosyncratic cash-flow volatility, and $\gamma > 0$ captures the cash-flow cost of pollution (future regulatory penalties, carbon tax exposure, stranded-asset risk). Two types of investors hold total wealth normalized to one. Traditional investors (mass $1-\mu$) maximize mean-variance utility $\mathbb{E}[W_T] - \frac{\rho}{2}\text{Var}(W_T)$. Green investors (mass μ) maximize the same objective but add a non-pecuniary penalty $\phi \cdot \mathbb{E}[\theta_i - a_i \mid S_i]$ per unit held, where $\phi > 0$ is the ESG taste parameter.

Market clearing with one share outstanding per firm yields the equilibrium price (derived in Appendix B):

$$P_i - \bar{P} = -\frac{(\gamma + \mu\phi)\lambda(m^*)}{R_f}(S_i - \bar{S}), \tag{19}$$

where $\bar{P}$ is the average price and R_f is the gross risk-free rate. Firms with lower (greener) scores command higher prices.

5.2 The Green Premium and Greenwashing Waste

The correlation between the green premium received by a firm and its true greenness captures how much of the ESG-related price differential rewards genuine environmental quality:

$$\text{Corr}(P_i - \bar{P},\ -(\theta_i - a^*)) = \sqrt{\lambda(m^*)}. \tag{20}$$

The fraction of the green premium that rewards true greenness is $\lambda(m^*)$; the fraction $1 - \lambda(m^*)$ flows to noise, including manipulation noise. As ESG demand grows and λ falls, a larger share of the green premium flows to greenwashers. Equation (20) connects the informational externality directly to investor wealth transfers: green investors overpay for firms that appear green due to favorable manipulation noise draws rather than low true pollution.

5.3 Declining Fundamental Sharpe Ratio

Proposition 13 (Declining GMB Sharpe Ratio)**.** *The green-minus-brown (GMB) portfolio is long the greenest decile and short the brownest decile of firms sorted by ESG score S_i. The fundamental Sharpe ratio of the GMB portfolio, isolating the cash-flow channel by setting $\phi = 0$, is*

$$\text{SR}^{\text{CF}}_{\text{GMB}} \propto \frac{\gamma\sqrt{\lambda(m^*)}\,\sigma_\theta}{\sqrt{\gamma^2(1-\lambda(m^*))\sigma_\theta^2 + \sigma_x^2}}. \tag{21}$$

This expression is strictly increasing in λ and hence strictly decreasing in μ (through the information degradation channel). As ESG demand grows, manipulation degrades the informativeness of the ESG sort, and the Sharpe ratio of exploiting genuine pollution differences declines.

Proof. Let $f(\lambda) = \lambda/[\gamma^2(1-\lambda)\sigma_\theta^2 + \sigma_x^2]$, which is proportional to $(\text{SR}^{\text{CF}}_{\text{GMB}})^2$. Then $f'(\lambda) = (\gamma^2\sigma_\theta^2 + \sigma_x^2)/[\gamma^2(1-\lambda)\sigma_\theta^2 + \sigma_x^2]^2 > 0$. Since λ is strictly decreasing in μ (Proposition 3(b)), the result follows. □

The fundamental Sharpe ratio isolates an economically meaningful channel. An investor who sorts firms by ESG scores to exploit pollution-related cash-flow differences faces a signal extraction problem. The quality of the signal (λ) declines as ESG demand grows, because manipulation degrades the information on which the strategy depends. The green factor "eats itself."

The total GMB Sharpe ratio (including the taste premium from $\phi > 0$) has competing forces. The taste component adds a return spread proportional to $\mu\phi$ that does not depend

on whether the sort accurately identifies true pollution. The total Sharpe ratio is

$$\mathrm{SR}_{\mathrm{GMB}}^{\mathrm{total}} \propto \frac{(\gamma+\mu\phi)\sqrt{\lambda(m^*)}\,\sigma_\theta}{\sqrt{\gamma^2(1-\lambda(m^*))\sigma_\theta^2+\sigma_x^2}}.$$

The numerator combines a declining term ($\gamma\sqrt{\lambda}$, from information degradation) with a growing term ($\mu\phi\sqrt{\lambda}$). The total Sharpe ratio declines in μ if and only if the information channel dominates:

$$\frac{d}{d\mu}\left[(\gamma+\mu\phi)\sqrt{\lambda(m^*(\mu))}\right] < 0 \quad \Longleftrightarrow \quad \phi\sqrt{\lambda} + \frac{(\gamma+\mu\phi)\lambda'(m^*)\dot{m}}{2\sqrt{\lambda}} < 0. \tag{22}$$

Rearranging, the total Sharpe ratio declines when

$$\frac{(\gamma+\mu\phi)|\lambda'(m^*)|\dot{m}}{2\lambda} > \phi,$$

which holds when (i) the cash-flow cost of pollution γ is large relative to the taste parameter ϕ (pollution has real cash-flow consequences, not just taste effects), or (ii) the semi-elasticity of λ with respect to μ is large (manipulation is highly responsive to ESG demand). In the limit $\phi \to 0$, the condition holds for all $\mu > 0$: the total and fundamental Sharpe ratios coincide. In the limit $\gamma \to 0$ (pure taste asset pricing with no cash-flow channel), the total Sharpe ratio always rises in μ because the taste premium dominates.

For calibrations with $\gamma > 0$ and moderate ϕ, the condition can hold at intermediate values of μ, producing a region where information degradation temporarily depresses the total Sharpe ratio. However, the condition fails for all $\mu > \gamma/(2\phi)$: the taste premium, growing linearly in μ while $\sqrt{\lambda}$ decays only as $\mu^{-1/3}$, eventually dominates (Proposition 24 in Appendix E). The total Sharpe ratio therefore rises without bound for large μ. Prediction P2 below applies specifically to the fundamental (cash-flow) component of the GMB Sharpe ratio, which declines monotonically to zero as ESG demand grows.

5.4 Testable Predictions

Three testable predictions follow directly from the equilibrium:

P1: Score informativeness declines with ESG AUM. The cross-sectional correlation between ESG scores and true environmental quality proxies (direct emissions, regulatory violations, third-party audits) should decline as the fraction of ESG-mandated capital grows. Regression specification: regress quality proxies on ESG scores, interacting with ESG AUM share; the interaction coefficient should be negative.

P2: GMB portfolio Sharpe ratio declines over time. As ESG investing has grown (2010–2025), the risk-adjusted performance of green-minus-brown portfolios, measured by the cash-flow component of returns, should have deteriorated. Rolling Sharpe ratios of ESG-sorted portfolios should trend downward.

P3: ESG rating disagreement increases with ESG AUM. More ESG demand triggers more manipulation, which adds more noise to scores. The Berg et al. (2022) disagreement measure should be positively correlated with ESG AUM growth across time or across sectors with different ESG pressure intensity.

6 Discussion

6.1 Generality of the Mechanism

The informational externality from noisy metric gaming applies to any setting where agents game a metric and gaming adds noise. The formal requirement is $g'(m) > 0$: noise must be increasing in manipulation effort (Appendix C). Under this condition, existence, uniqueness, comparative statics, welfare hump shape, and the double dividend all hold regardless of the specific noise function. The proportional specification $g(m) = m^2$ provides closed forms; it is not necessary for any qualitative result.

Applications beyond ESG include credit ratings (firms manage their rating through financial engineering), standardized testing (teaching to the test degrades score informativeness for other students), online reviews (fake reviews add noise that makes ratings less useful for all consumers), and hospital quality scores (selective patient referral adds noise to outcome metrics). The ESG setting has three features that make the analysis different from these other applications. First, the capital allocation channel ($\eta\mu\lambda$) connects the informational externality directly to real investment and welfare through portfolio sorting, rather than through a purchasing or admissions decision. Second, the policy instruments are specific to financial markets (mandatory disclosure rules, Pigouvian taxes on financial manipulation, interactions with carbon pricing), and the over-abatement result (Proposition 10(d)) depends on the interaction between ESG scoring and carbon taxation that has no analogue in education or health care. Third, the two-type extension generates predictions about cross-sectional asset pricing (which firms lose the most green capital from information degradation) that connect to the empirical finance literature on ESG rating disagreement.

6.2 Relationship to Existing Results

The model nests two benchmark cases. When $\psi = 0$ (precise manipulation), the model reduces to a standard Tullock rent-seeking contest: manipulation is pure waste with no information degradation, λ is constant, and the first-order conditions for abatement and manipulation are separable. When $c_m \to \infty$ (no manipulation possible), $m^* = 0$ and the model becomes a standard ESG allocation problem with exogenous score quality.

The noisy manipulation model ($\psi > 0$) adds two features absent from both benchmarks. First, the informational externality through endogenous λ creates a feedback loop: more ESG demand triggers more manipulation, which degrades information, which partially offsets the demand's allocative benefit. Second, the double dividend of anti-manipulation policy (Proposition 8) arises because abatement and manipulation interact through λ. Standard rent-seeking models, in which manipulation is a pure mean shift, cannot generate the double dividend.

The informational externality studied here belongs to a broader class of strategic information degradation problems. Frankel and Kartik (2019) study "muddled information" in which senders optimally add noise to their signals, and Pérez-Richet and Skreta (2022) analyze test design when agents can game the test. The model here differs from these frameworks in two respects. First, the noise here is a side effect of manipulation, not a strategic choice: firms want to improve their scores precisely, but the manipulation technology is inherently noisy. Second, the externality operates through a market-level object (λ) that aggregates across a continuum of agents, generating a prisoner's dilemma that neither the sender-receiver nor the test-design frameworks produce.

Interaction with demand-side channels. Goldstein et al. (2022) identify a demand-side channel: heterogeneous ESG preferences among investors reduce price informativeness because investors with different ESG tastes trade in opposite directions on the same signal. The supply-side channel here operates through firms adding noise to signals via imprecise manipulation. The two channels are likely complements rather than substitutes. When price informativeness falls (the GKSX channel), firms face a noisier inference environment and the marginal return to manipulation may change. Conversely, when score informativeness falls (the supply-side channel), investors with heterogeneous ESG preferences face noisier signals, which amplifies the GKSX disagreement mechanism. A combined model would feature two endogenous noise sources (demand-side from investor heterogeneity and supply-side from firm manipulation), and the total information loss would exceed the sum of the parts. The predictions here, particularly the declining fundamental Sharpe ratio, should therefore be interpreted as a lower bound on the total information degradation when both channels

operate simultaneously.

6.3 Connection to Empirical Evidence

Several recent empirical findings are consistent with the model's qualitative predictions, though the specific mechanisms in these papers differ from the model's formalization. Duchin et al. (2025) document that firms targeted by green investors divest pollutive assets without reducing aggregate pollution: the buyers simply continue the same operations. The model treats manipulation as a generic action that improves scores without reducing pollution; the Duchin, Gao, and Xu finding that asset divestiture is a specific channel through which cosmetic ESG improvement occurs is consistent with the model's broader prediction that ESG pressure induces socially wasteful score improvement, though the model does not speak to why firms choose asset sales over other forms of manipulation. Hartzmark and Shue (2023) find that sustainable investing can be counterproductive, with brown firms becoming browner under ESG pressure. One mechanism is the hump-shaped welfare result (Proposition 6): excessive ESG demand generates manipulation waste and information degradation that can outweigh the allocative benefits. Berk and van Binsbergen (2025) show that the direct cost-of-capital channel of ESG divestment is quantitatively negligible. A complementary channel operates here: ESG demand affects firm behavior through the green capital allocation mechanism ($\eta\mu\lambda$), which operates through scoring and information rather than cost of capital.

6.4 Limitations

Static model. The model has no dynamics and no learning. In a dynamic setting, investors could learn about greenwashing and adjust λ downward, partially restoring informativeness. The static prediction is most relevant when the manipulation technology evolves faster than investor learning (e.g., new forms of greenwashing emerge as old ones are detected) or when new firms continuously enter, resetting the learning problem. If investor learning dominates, the steady-state externality is smaller than the static prediction, and the optimal μ^* is correspondingly higher.

Two types versus full heterogeneity. The two-type extension (Section 3.3) introduces green and brown firms and generates the distributional prediction that the externality falls disproportionately on green firms. A model with a continuous distribution of types (heterogeneous c_m and θ_i) would generate richer sorting patterns and allow for a full distributional analysis of who gains and who loses from the rat race. In particular, it could address whether

firms sort into manipulation versus abatement as a function of their type, and whether the equilibrium exhibits endogenous clustering of green and brown firms at different points of the cost distribution.

Continuum of firms. The externality is maximal in the continuum limit. With N firms, each internalizes $1/N$ of the informational effect. For a concentrated industry ($N = 10$), the wedge between private and social manipulation is smaller, and the case for anti-greenwashing regulation is correspondingly weaker.

Welfare with and without taste utility. The baseline welfare function (6) counts environmental outcomes only. An alternative welfare function that includes green investors' taste utility adds a term $\mu\phi\lambda(m^*)\sigma_\theta^2$ to $W(\mu)$, capturing the non-pecuniary benefit green investors derive from holding firms they perceive as clean. Under this extended welfare function, the optimal ESG demand μ^* is higher because the marginal benefit of ESG demand includes the taste satisfaction channel. Formally, the first-order condition gains the additional term $\phi\lambda(m^*)\sigma_\theta^2 + \mu\phi\lambda'(m^*)\dot{m}\sigma_\theta^2$, which is positive for moderate μ. The hump shape and the informational externality remain: both welfare functions generate excessive manipulation (the taste utility does not internalize the noise externality), and both imply an interior optimum. The quantitative difference is that including taste utility raises μ^* by approximately $\phi\sigma_\theta^2/[\eta\lambda_0(1/c_a + 1/c_m)]$ in the $\psi = 0$ case.

The capital allocation rule (2) in the welfare analysis is a reduced-form linear allocation based on Bayesian updating, while the asset pricing section (Section 5) derives equilibrium prices from CAPM investor preferences. The two are consistent: the green capital allocation (2) is the mapping from scores to investment that emerges from the portfolio optimality conditions of Section 5, with λ playing the role of the posterior precision and $\eta\mu$ capturing the effective demand for ESG-linked capital. The welfare analysis takes the allocation rule as given to isolate the externality; the asset pricing section microfounds the rule from investor optimization.

Concavity of welfare for general ψ. Proposition 6 resolves the welfare shape for all $\psi \geq 0$. The key step is reducing the fixed-point equation to the cubic relation (12), which reveals that $P(\mu) = c_m m^*(\mu)$ is the inverse of a strictly convex function and hence globally concave. Because W is strictly concave in P (with $d^2W/dP^2 = -(1/c_a + 1/c_m)$), the sign of $W'(\mu)$ depends only on the linear condition $K - \beta P(\mu)$, where $K = \sigma_\theta^2 + (\delta_S - \delta)/c_a$ and $\beta = 1/c_a + 1/c_m$. Since P is monotone increasing, W' changes sign exactly once, giving uniqueness of the maximizer for all $\psi \geq 0$. For $\psi > 0$, W is not globally concave: $W''(\mu)$

changes sign at an inflection point that lies strictly past the maximizer. The non-concavity occurs only in the decreasing tail (where W is already below its maximum) and has no policy consequence: the optimal level of ESG demand is unique.

7 Conclusion

ESG score manipulation creates an informational externality that degrades score quality for all firms. In the unique symmetric equilibrium, all firms manipulate, the market strips out the average manipulation, but the noise remains and destroys the information on which green capital allocation depends. The externality falls disproportionately on genuinely green firms, who lose the ability to distinguish themselves when brown firms greenwash heavily. Social welfare is hump-shaped in ESG demand: some demand is beneficial because it corrects an uninternalized pollution externality and directs capital toward genuinely greener firms, but excessive demand triggers a wasteful arms race whose costs (direct manipulation expense and information degradation) eventually dominate. Anti-manipulation policy has a double dividend, and optimal enforcement intensity should increase as ESG demand grows. The fundamental Sharpe ratio of green-minus-brown portfolios declines as ESG demand grows; the total Sharpe ratio declines when the cash-flow cost of pollution is large relative to the taste parameter. These predictions distinguish the information degradation channel from taste-based explanations of declining green returns.

References

Akerlof, G. A. (1976). The economics of caste and of the rat race and other woeful tales. *Quarterly Journal of Economics 90*(4), 599–617.

Azarmsa, E. and J. Shapiro (2025). The market for ESG ratings. *Journal of Finance*.

Berg, F., J. F. Kölbel, and R. Rigobon (2022). Aggregate confusion: The divergence of ESG ratings. *Review of Finance 26*(6), 1315–1344.

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

Cartellier, F., P. Tankov, and O. D. Zerbib (2023). Can investors curb greenwashing? SSRN Working Paper 4644741.

Cayirli, E. (2025). When ESG demand backfires: Greenwashing, cost convexity, and policy design. SSRN Working Paper 5581853.

Duchin, R., J. Gao, and Q. Xu (2025). Sustainability or greenwashing: Evidence from the asset market for industrial pollution. *Journal of Finance 80*(2), 699–754.

Frankel, A. and N. Kartik (2019). Muddled information. *Journal of Political Economy 127*(4), 1739–1776.

Goldstein, I., A. Kopytov, L. Shen, and H. Xiang (2022). On ESG investing: Heterogeneous preferences, information, and asset prices. NBER Working Paper 29839.

Hartzmark, S. M. and K. Shue (2023). Counterproductive sustainable investing: The impact elasticity of brown and green firms. *Journal of Finance*. Conditionally accepted.

Inderst, R. and M. M. Opp (2025). Sustainable finance versus environmental policy. *Journal of Financial Economics*.

Leippold, M., C. Colesanti Senni, and S. A. Vaghefi (2025). The social cost of greenwashing. SSRN Working Paper 5483808.

Pástor, v., R. F. Stambaugh, and L. A. Taylor (2021). Sustainable investing in equilibrium. *Journal of Financial Economics 142*(2), 550–571.

Pérez-Richet, E. and V. Skreta (2022). Test design under falsification. *Econometrica 90*(3), 1109–1142.

A Proofs of Main Results

Derivation of Equilibrium from Type-Dependent Strategies

To derive the equilibrium without circularity, suppose the market conjectures linear strategies $a_i = \alpha_0 + \alpha_1\theta_i$ and $m_i = \beta_0 + \beta_1\theta_i$. The ESG score becomes $S_i = (1-\alpha_1-\beta_1)\theta_i - \alpha_0 - \beta_0 - \nu_i + \varepsilon_i$. Define $\phi_c \equiv 1 - \alpha_1 - \beta_1$ as the score sensitivity to type.

A deviating firm i choosing (a_i, m_i) produces score $S_i = \theta_i - a_i - m_i - \nu_i + \varepsilon_i$. The market, believing the equilibrium conjecture, forms the posterior $\mathbb{E}[\theta_i \mid S_i] = \bar{\theta} + \tilde{\lambda}(S_i - \bar{S})$, where $\tilde{\lambda} = \phi_c\sigma_\theta^2/(\phi_c^2\sigma_\theta^2 + \sigma_n^2)$.

The firm's expected green capital benefit is $\mathbb{E}[\eta K_i \mid \theta_i, a_i, m_i] = \eta\mu[\text{const} + (1-\alpha_1)\tilde{\lambda}(a_i + m_i - \text{const}'(\theta_i))]$, so the marginal effect of both a_i and m_i on expected green capital is the same: $\partial\mathbb{E}[\eta K_i]/\partial a_i = \partial\mathbb{E}[\eta K_i]/\partial m_i = \eta\mu(1-\alpha_1)\tilde{\lambda}$.

The first-order conditions are

$$a_i = \frac{\eta\mu(1-\alpha_1)\tilde{\lambda} + \delta}{c_a},$$
$$m_i = \frac{\eta\mu(1-\alpha_1)\tilde{\lambda}}{c_m}.$$

Both are independent of θ_i, so $\alpha_1 = 0$ and $\beta_1 = 0$ in equilibrium, giving $\phi_c = 1$ and $(1-\alpha_1)\tilde{\lambda} = \lambda(m^*)$. The Hessian of the firm's payoff is $\text{diag}(-c_a, -c_m)$, negative definite, confirming the interior solution is the unique optimum.

Lemma 14 (No Non-Linear Equilibria)**.** *Any symmetric equilibrium strategy must be type-independent: $a_i = a^*$ and $m_i = m^*$ for constants (a^*, m^*) independent of θ_i. In particular, no non-linear equilibrium exists.*

Proof. Fix an arbitrary symmetric conjecture: all other firms play measurable strategies $\hat{a}(\theta), \hat{m}(\theta)$ (possibly non-linear). Let λ denote the market's updating coefficient under this conjecture (however determined). A deviating firm i choosing (a_i, m_i) produces score $S_i = \theta_i - a_i - m_i - \nu_i + \varepsilon_i$. The allocation rule (2) gives $K_i = \mu[\kappa - \lambda(S_i - \bar{S})]$, which is linear in S_i.

The key step does not require the posterior $\mathbb{E}[\theta_i \mid S_i]$ to be linear in S_i. It requires only that the allocation rule is linear in S_i and the noise terms have mean zero. Taking expectations conditional on (θ_i, a_i, m_i) and using $\mathbb{E}[\nu_i] = \mathbb{E}[\varepsilon_i] = 0$:

$$\mathbb{E}[K_i \mid \theta_i, a_i, m_i] = \mu\big[\kappa - \lambda\big((\theta_i - \bar{\theta}) - (a_i - \mathbb{E}[\hat{a}]) - (m_i - \mathbb{E}[\hat{m}])\big)\big].$$

The marginal effects $\partial\mathbb{E}[\eta K_i]/\partial a_i = \partial\mathbb{E}[\eta K_i]/\partial m_i = \eta\mu\lambda$ are independent of θ_i. The θ_i

term affects the level of expected capital but not the marginal benefit of either action. The objective (5) is strictly concave in (a_i, m_i) with first-order conditions

$$a_i = \frac{\eta\mu\lambda + \delta}{c_a}, \qquad m_i = \frac{\eta\mu\lambda}{c_m},$$

which are independent of θ_i for every conjecture $(\hat{a}, \hat{m})$ and every $\lambda > 0$. (The non-negativity constraints $a_i \geq 0$ and $m_i \geq 0$ are slack because $\eta\mu\lambda + \delta > 0$ and $\eta\mu\lambda > 0$ under the maintained parameter assumptions; even at a corner, the best response remains θ_i-independent.) Since a single deviator has measure zero in the continuum and cannot affect $\bar{S}$ or λ, the unique best response is type-independent regardless of the conjecture. No equilibrium can have strategies depending on θ_i. □

Derivation of Allocative Efficiency

Green capital is $K_i = \mu[\kappa - \lambda(S_i - \bar{S})]$ and true greenness is $a^* - \theta_i$. The informative part of the score is $S_i - \bar{S} = (\theta_i - \bar{\theta}) + \varepsilon_i - \nu_i$. Then

$$\mathrm{Cov}(K_i, a^* - \theta_i) = \mu\lambda \cdot \mathrm{Cov}(-(\theta_i - \bar{\theta} + \varepsilon_i - \nu_i), -(\theta_i - \bar{\theta})) = \mu\lambda\sigma_\theta^2.$$

Dividing by μ gives $\mathrm{AE} = \lambda\sigma_\theta^2$.

Proof of Proposition 3 (Greenwashing Rat Race)

Existence. Define $f(m) \equiv \eta\mu\sigma_\theta^2/[c_m(\sigma_\theta^2 + \sigma_\varepsilon^2 + \psi^2 m^2)]$ and $h(m) \equiv m - f(m)$. We have $f(0) = \eta\mu\sigma_\theta^2/[c_m(\sigma_\theta^2 + \sigma_\varepsilon^2)] > 0$, so $h(0) < 0$. As $m \to \infty$, $f(m) \to 0$, so $h(m) \to \infty$. Since h is continuous, the intermediate value theorem gives at least one root.

Uniqueness. For $m > 0$, $f'(m) = -2\eta\mu\psi^2 m\sigma_\theta^2/[c_m(\sigma_\theta^2 + \sigma_\varepsilon^2 + \psi^2 m^2)^2] < 0$. Therefore $h'(m) = 1 - f'(m) > 1 > 0$ for all $m > 0$: h is strictly increasing on $(0, \infty)$ and crosses zero exactly once. □

Part (b): Information degradation. $\lambda(m) = \sigma_\theta^2/(\sigma_\theta^2 + \sigma_\varepsilon^2 + \psi^2 m^2)$ is strictly decreasing in m for $m > 0$: $d\lambda/dm = -2\psi^2 m\sigma_\theta^2/(\sigma_\theta^2 + \sigma_\varepsilon^2 + \psi^2 m^2)^2 < 0$.

Part (c): Manipulation increases in μ. Implicit differentiation of $c_m m^* = \eta\mu\lambda(m^*)$:

$$(c_m - \eta\mu\lambda'(m^*))\, dm^* = \eta\lambda(m^*)\, d\mu.$$

Since $\lambda'(m^*) < 0$, the coefficient $c_m - \eta\mu\lambda'(m^*) > c_m > 0$ and $\eta\lambda(m^*) > 0$, giving $dm^*/d\mu > 0$. The chain rule gives $d\lambda/d\mu = \lambda'(m^*) \cdot dm^*/d\mu < 0$. □

Part (d): Wasteful rat race. In symmetric equilibrium, the market correctly anticipates m^* and strips it out. The expected allocation $\mathbb{E}[K_i]$ is the same as it would be with $m^* = 0$ (holding the score structure fixed). But λ is lower: higher m^* reduces λ, shrinking the dispersion of K_i and reducing the correlation between green capital and true greenness. The direct cost is $c_m(m^*)^2/2$ per firm. □

Proof of Proposition 6 (Unique Welfare Maximum)

Step 1: Cubic reduction. From the fixed-point equation $c_m m^* = \eta\mu\lambda(m^*)$ with $\lambda(m) = \sigma_\theta^2/(\sigma_\theta^2 + \sigma_\varepsilon^2 + \psi^2 m^2)$, define $P = c_m m^*$. Substituting $m^* = P/c_m$:

$$P\left(\sigma_\theta^2 + \sigma_\varepsilon^2 + \frac{\psi^2}{c_m^2}P^2\right) = \eta\sigma_\theta^2\mu.$$

Write $B = \sigma_\theta^2 + \sigma_\varepsilon^2$, $\alpha = \psi^2/c_m^2$, $s = \eta\sigma_\theta^2$. The fixed point reduces to $H(P) \equiv BP + \alpha P^3 = s\mu$. Since $H'(P) = B + 3\alpha P^2 > 0$, the function H is strictly increasing, and $P(\mu) = H^{-1}(s\mu)$ is well-defined and unique.

Step 2: P is increasing and concave. Differentiating $H(P(\mu)) = s\mu$:

$$P'(\mu) = \frac{s}{B + 3\alpha P(\mu)^2} > 0.$$

Differentiating again:

$$P''(\mu) = -\frac{6\alpha P(\mu)\, s^2}{(B + 3\alpha P(\mu)^2)^3} \leq 0 \quad (\mu > 0),$$

For $\psi > 0$ ($\alpha > 0$), H is strictly convex ($H'' = 6\alpha P > 0$ for $P > 0$), so its inverse is strictly concave. For $\psi = 0$ ($\alpha = 0$), $P(\mu) = s\mu/B$ is linear. □

Step 3: Welfare as a function of P. Write $W(\mu) = \widetilde{W}(P(\mu))$ where

$$\widetilde{W}(P) = \frac{\delta_S(P + \delta)}{c_a} + \sigma_\theta^2 P - \frac{(P + \delta)^2}{2c_a} - \frac{P^2}{2c_m} - \delta_S\bar{\theta}.$$

Define $K = \sigma_\theta^2 + (\delta_S - \delta)/c_a > 0$ and $\beta = 1/c_a + 1/c_m > 0$. Then:

$$\widetilde{W}'(P) = K - \beta P, \qquad \widetilde{W}''(P) = -\beta < 0.$$

So W is strictly concave as a function of P. □

Step 4: Unique maximizer. By the chain rule:

$$W'(\mu) = \widetilde{W}'(P(\mu)) \cdot P'(\mu) = \frac{s(K - \beta P(\mu))}{B + 3\alpha P(\mu)^2}.$$

The denominator is positive, so $\text{sign}(W'(\mu)) = \text{sign}(K - \beta P(\mu))$. Since $P(\mu)$ is strictly increasing with $P(0) = 0$ and $P(\mu) \to \infty$, the expression $K - \beta P(\mu)$ crosses zero exactly once, at $P^* = K/\beta$. The unique maximizer is

$$\mu^* = \frac{BP^* + \alpha(P^*)^3}{s}, \qquad P^* = \frac{\sigma_\theta^2 + (\delta_S - \delta)/c_a}{1/c_a + 1/c_m}.$$

Part (a): $W'(0) > 0$. At $\mu = 0$: $P(0) = 0$, so $K - \beta \cdot 0 = K > 0$. □

Part (b): $W'(\mu) < 0$ for large μ. For $\mu > \mu^*$: $P(\mu) > P^* = K/\beta$, so $K - \beta P < 0$. □

Part (e): Clean case. When $\psi = 0$: $\alpha = 0$, $P(\mu) = s\mu/B = \eta\lambda_0\mu$, and $W(\mu)$ is quadratic in μ with $W''(\mu) = -\beta(\eta\lambda_0)^2 < 0$. Setting $W'(\mu) = 0$ yields (14). □

Part (f): Inflection point for $\psi > 0$. The second derivative is

$$W''(\mu) = -\frac{s^2\left[\beta B + 6\alpha K P(\mu) - 3\beta\alpha P(\mu)^2\right]}{(B + 3\alpha P(\mu)^2)^3}.$$

Define $q(P) = \beta B + 6\alpha K P - 3\beta\alpha P^2$. This downward parabola has $q(0) = \beta B > 0$ and $q \to -\infty$, so it has a unique positive root:

$$P_{\text{inf}} = \frac{K}{\beta} + \sqrt{\left(\frac{K}{\beta}\right)^2 + \frac{B}{3\alpha}}.$$

Since $P_{\text{inf}} > K/\beta = P^*$, the inflection point lies strictly past the maximizer. Thus $W'' < 0$ on $[0, \mu^*]$ (and beyond, up to μ_{inf}), and $W'' > 0$ only in the tail where W is already below its maximum. □

Proof of Proposition 8 (Double Dividend)

Part (a). With tax t and $m = 0$, the manipulation FOC is $\eta\mu\lambda(0) - t \leq 0$. Setting $t = \eta\mu\lambda_0$ gives $m^* = 0$. With $m^* = 0$, $\lambda = \lambda_0$, and the abatement FOC gives $a^* = (\eta\mu\lambda_0 + \delta)/c_a$. □

Part (b). The welfare gain from moving from $(m^*, \lambda(m^*))$ to $(0, \lambda_0)$ is

$$\Delta W = \frac{c_m}{2}(m^*)^2 + \eta\mu[\lambda_0 - \lambda(m^*)]\sigma_\theta^2 + \text{abatement gain},$$

where the abatement gain captures the increase from $(\eta\mu\lambda(m^*) + \delta)/c_a$ to $(\eta\mu\lambda_0 + \delta)/c_a$, valued at social rate δ_S. □

Part (c). Abatement $a^* = (\eta\mu\lambda + \delta)/c_a$ is increasing in λ. Since $\lambda_0 > \lambda(m^*)$ for $m^* > 0$, eliminating manipulation raises λ and hence a^*. With $\psi = 0$, λ is constant so eliminating manipulation does not affect a^*. □

Proof of Proposition 10 (Mandatory Disclosure)

Part (a). Under mandatory disclosure, $\partial S_i/\partial m_i = 0$. The manipulation FOC becomes $0 - c_m m_i \leq 0$, giving $m^* = 0$. □

Part (d). First-best abatement: $a^{FB} = \delta_S/c_a$. Under mandatory disclosure with $\delta = \delta_S$: $a^* = (\eta\mu\tilde{\lambda} + \delta_S)/c_a > \delta_S/c_a = a^{FB}$ whenever $\mu > 0$ and $\tilde{\lambda} > 0$. Setting $\delta = \delta_S - \eta\mu\tilde{\lambda}$ achieves $a^* = a^{FB}$. □

Proof of Proposition 5 (Two-Type Equilibrium)

Existence and uniqueness. In the two-type model, the fixed-point equation for λ is

$$\lambda = \frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_\varepsilon^2 + \psi^2[\alpha(m^G)^2 + (1-\alpha)(m^B)^2]},$$

where $m^\tau = \eta\mu\lambda/c_m^\tau$. Substituting:

$$\lambda = \frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_\varepsilon^2 + \psi^2(\eta\mu\lambda)^2[\alpha/(c_m^G)^2 + (1-\alpha)/(c_m^B)^2]}.$$

Define $\bar{c}^{-2} \equiv \alpha/(c_m^G)^2 + (1-\alpha)/(c_m^B)^2$ and $\bar{m}(\lambda) \equiv \eta\mu\lambda\sqrt{\bar{c}^{-2}}$. The fixed-point equation has the same structure as the baseline (equation (10) with an effective manipulation level $\bar{m}$), so the existence and uniqueness proof of Proposition 3 applies. □

Part (c): Green premium loss. Under no manipulation, the expected green capital for a type-G firm is $\mathbb{E}[K_i^G] = \mu[\kappa - \lambda_0(\bar{\theta}^G - \bar{\theta})] = \mu\kappa + \mu\lambda_0(\bar{\theta} - \bar{\theta}^G)$. Under equilibrium manipulation, $\mathbb{E}[K_i^G] = \mu\kappa + \mu\lambda(\bar{\theta} - \bar{\theta}^G)$. The loss is $\mu(\lambda_0 - \lambda)(\bar{\theta} - \bar{\theta}^G) > 0$ because $\lambda_0 > \lambda$ and $\bar{\theta} > \bar{\theta}^G$. For brown firms, $\bar{\theta}^B > \bar{\theta}$, so the loss is $\mu(\lambda_0 - \lambda)(\bar{\theta} - \bar{\theta}^B) < 0$: brown firms actually gain from reduced informativeness. □

Proof of Proposition 11 (Imperfect Enforcement)

Part (a). The fixed-point equation with imperfect enforcement is $c_m m + p\tau = \eta\mu\lambda(m)$. Define $h(m) = c_m m + p\tau - \eta\mu\lambda(m)$. We have $h'(m) = c_m + \eta\mu|\lambda'(m)| > 0$ for all $m \geq 0$. At $m = 0$: $h(0) = p\tau - \eta\mu\lambda_0$. If $p\tau < \eta\mu\lambda_0$, then $h(0) < 0$ and $h(m) \to \infty$, so there is a unique root $m^*(p, \tau) > 0$. If $p\tau \geq \eta\mu\lambda_0$, then $h(0) \geq 0$ and $m^* = 0$. Increasing $p\tau$ raises h pointwise, shifting the root to the left. □

Part (b). At $p\tau = \eta\mu\lambda_0$ and $m = 0$: the FOC gives $\eta\mu\lambda_0 - p\tau = 0$. For any deviation $m > 0$: $\eta\mu\lambda(m) < \eta\mu\lambda_0 = p\tau$, so $\eta\mu\lambda(m) - p\tau - c_m m < 0$. No deviation is profitable. □

Parts (c)–(d). The welfare function including enforcement costs is $\tilde{W}(p) = W(m^*(p, \tau), \mu) - \kappa(p)$. Since W is decreasing in m^* at the private equilibrium (Proposition 4) and m^* is decreasing in $p\tau$, $dW/dp > 0$. With κ convex, the interior optimum $\kappa'(p^*) = dW/dp$ exists. For the comparative static: $d^2\tilde{W}/(dp\, d\mu) > 0$ because higher μ increases both the manipulation level and the informational externality, raising the marginal benefit of enforcement. The argument for ψ is analogous. □

B Asset Pricing Derivations

Investor Portfolio Problems

The traditional investor's problem is

$$\max_{\{d_i^T\}} \int_0^1 d_i^T \left(\mathbb{E}[X_i \mid S_i] - R_f P_i\right) di - \frac{\rho}{2} \mathrm{Var}\left(\int_0^1 d_i^T X_i \, di\right),$$

and the green investor's problem is

$$\max_{\{d_i^G\}} \int_0^1 d_i^G (\mathbb{E}[X_i \mid S_i] - R_f P_i) \, di - \frac{\rho}{2} \mathrm{Var}\left(\int_0^1 d_i^G X_i \, di\right) - \phi \int_0^1 d_i^G \cdot \mathbb{E}[\theta_i - a_i \mid S_i] \, di.$$

Equilibrium Price Derivation

Conditional on S_i, firm i's expected cash flow is $\mathbb{E}[X_i \mid S_i] = \bar{X} - \gamma(\bar{\theta} - a^*) - \gamma\lambda(m^*)(S_i - \bar{S})$. The traditional investor's FOC for firm i (focusing on the cross-sectional component): $\mathbb{E}[X_i \mid S_i] - R_f P_i - \rho\sigma_x^2 d_i^T = 0$. The green investor's FOC: $\mathbb{E}[X_i \mid S_i] - R_f P_i - \rho\sigma_x^2 d_i^G - \phi\mathbb{E}[\theta_i - a_i \mid S_i] = 0$.

Market clearing requires $(1-\mu)d_i^T + \mu d_i^G = 1$. Substituting the FOCs:

$$\frac{\mathbb{E}[X_i \mid S_i] - R_f P_i - \mu\phi\mathbb{E}[\theta_i - a_i \mid S_i]}{\rho\sigma_x^2} = 1.$$

Solving for P_i and using $\mathbb{E}[\theta_i - a_i \mid S_i] = (\bar{\theta} - a^*) + \lambda(m^*)(S_i - \bar{S})$:

$$R_f P_i = \bar{X} - (\gamma + \mu\phi)(\bar{\theta} - a^*) - \rho\sigma_x^2 - (\gamma + \mu\phi)\lambda(m^*)(S_i - \bar{S}).$$

The cross-sectional price is $P_i - \bar{P} = -(\gamma + \mu\phi)\lambda(m^*)(S_i - \bar{S})/R_f$.

Green Premium Waste (Equation 20)

From $P_i - \bar{P} = -(\gamma + \mu\phi)\lambda(m^*)(S_i - \bar{S})/R_f$ and $S_i - \bar{S} = (\theta_i - \bar{\theta}) + \varepsilon_i - \nu_i$:

$$\mathrm{Cov}(P_i - \bar{P}, -(\theta_i - \bar{\theta})) = \frac{(\gamma + \mu\phi)\lambda}{R_f}\sigma_\theta^2,$$

$$\mathrm{Var}(P_i - \bar{P}) = \left(\frac{(\gamma + \mu\phi)\lambda}{R_f}\right)^2 (\sigma_\theta^2 + \sigma_n^2) = \left(\frac{\gamma + \mu\phi}{R_f}\right)^2 \lambda\sigma_\theta^2,$$

where the last step uses $\lambda^2(\sigma_\theta^2 + \sigma_n^2) = \lambda\sigma_\theta^2$. The correlation is

$$\mathrm{Corr} = \frac{(\gamma + \mu\phi)\lambda\sigma_\theta^2/R_f}{[(\gamma + \mu\phi)/R_f]\sqrt{\lambda}\,\sigma_\theta \cdot \sigma_\theta} = \sqrt{\lambda(m^*)}. \quad \square$$

Expected Returns

From the pricing equation, $\mathbb{E}[R_i] - R_f = [\mu\phi\mathbb{E}[\theta_i - a_i \mid S_i] + \rho\sigma_x^2]/P_i$. In the cross-section, the market risk premium $\rho\sigma_x^2/P_i$ varies across firms through P_i. When the cross-sectional price dispersion is small relative to the price level (formally, when $(\gamma + \mu\phi)\lambda\sigma_\theta/\bar{P} \ll 1$), the approximation $\rho\sigma_x^2/P_i \approx \rho\sigma_x^2/\bar{P}$ holds to first order, and the green tilt generates cross-sectional variation $(\gamma + \mu\phi)\lambda(m^*)(S_i - \bar{S})/(R_f\bar{P})$. Firms perceived as dirtier (higher S_i) earn higher expected returns.

The GMB portfolio (long green decile, short brown decile) has zero market beta because

all firms share the same market loading b. Its CAPM alpha equals its expected return:

$$\alpha_{\text{GMB}} = -\frac{(\gamma + \mu\phi)\lambda(m^*)}{R_f \bar{P}} \cdot \Delta S,$$

where ΔS is the expected score gap between the top and bottom deciles. The magnitude depends on μ through two competing channels: the direct taste effect (increasing in μ) and the information degradation effect (decreasing λ).

C General Noise Specification

Replace the proportional noise with a general noise function: $\text{Var}(\nu_i) = \psi^2 g(m_i)$, where $g : [0, \infty) \to [0, \infty)$ satisfies:

- (G1) $g(0) = 0$: no manipulation implies no manipulation noise.
- (G2) $g'(m) > 0$ for $m > 0$: more manipulation generates more noise.

The baseline model corresponds to $g(m) = m^2$. Other natural specifications include $g(m) = m^\alpha$ for $\alpha \geq 1$, $g(m) = m$ (linear), and $g(m) = e^m - 1$ (exponential). The Bayesian updating coefficient becomes $\lambda_g(m) = \sigma_\theta^2/(\sigma_\theta^2 + \sigma_\varepsilon^2 + \psi^2 g(m))$, and the fixed-point equation is

$$c_m m = \frac{\eta\mu\sigma_\theta^2}{\sigma_\theta^2 + \sigma_\varepsilon^2 + \psi^2 g(m)}. \tag{23}$$

Proposition 15 (Existence and Uniqueness under General Noise)**.** *Under G1 and G2, equation (23) has exactly one solution $m^* > 0$.*

Proof. Define $f_g(m) = \eta\mu\sigma_\theta^2/[c_m(\sigma_\theta^2 + \sigma_\varepsilon^2 + \psi^2 g(m))]$ and $h_g(m) = m - f_g(m)$. At $m = 0$: $f_g(0) > 0$, so $h_g(0) < 0$. As $m \to \infty$: $g(m) \to \infty$ (by G2 and $g(0) = 0$), so $f_g(m) \to 0$ and $h_g(m) \to \infty$. By the IVT, at least one root exists.

For uniqueness: $f_g'(m) = -\eta\mu\psi^2 g'(m)\sigma_\theta^2/[c_m(\sigma_\theta^2 + \sigma_\varepsilon^2 + \psi^2 g(m))^2] < 0$ under G2. Therefore $h_g'(m) = 1 - f_g'(m) > 1 > 0$: h_g is strictly increasing and crosses zero exactly once. □

Proposition 16 (Comparative Statics under General Noise)**.** *Under G1–G2:*

(a) $dm^*/d\mu > 0$.

(b) $d\lambda_g/d\mu < 0$.

(c) $W'(0) > 0$ *whenever* $\delta_S > \delta$ *or* $\sigma_\theta^2 > 0$.

(d) $W(\mu) \to -\infty$ *as* $\mu \to \infty$.

(e) There exists at least one interior welfare maximum $\mu^* > 0$.

Proof. Part (a): implicit differentiation of $c_m m = \eta\mu\lambda_g(m)$ gives $dm^*/d\mu = \eta\lambda_g(m^*)/(c_m -$

$\eta\mu\lambda_g'(m^*)) > 0$, since $\lambda_g' < 0$ under G2. Part (b): $d\lambda_g/d\mu = \lambda_g' \cdot dm^*/d\mu < 0$. Parts (c)–(e): identical to the baseline proof. □

Necessary conditions. The following table summarizes which conditions on g are needed for each result:

Result	Required condition	Fails when
Existence and uniqueness of m^*	G1 + G2	Never (given G1–G2)
$dm^*/d\mu > 0$	G2	$g' = 0$ (additive noise)
Welfare hump shape	G1 + G2	$g' = 0$
Double dividend	G2	$g' = 0$ (separable FOCs)

The additive noise case ($g(m) = c$, constant, violating G2) eliminates the informational externality: λ_g becomes constant, and the model reduces to the $\psi = 0$ benchmark. Condition G2 ($g' > 0$) is necessary and sufficient for the information degradation mechanism.

D Perturbation Analysis

Write all equilibrium objects as functions of (μ, ψ). At $\psi = 0$, the equilibrium is $m_0^*(\mu) = \eta\mu\lambda_0/c_m$, $a_0^*(\mu) = (\eta\mu\lambda_0 + \delta)/c_a$, $\lambda_0 = \sigma_\theta^2/(\sigma_\theta^2 + \sigma_\varepsilon^2)$.

Proposition 17 (Perturbation Expansion)**.** *For small ψ, the equilibrium objects admit expansions in ψ^2:*

(a) Manipulation: $m^*(\mu, \psi) = m_0^*(\mu) - \psi^2 m_2^*(\mu) + O(\psi^4)$, where $m_2^* > 0$. Manipulation decreases relative to $\psi = 0$ because the noisy technology reduces the marginal benefit of manipulation.

(b) Informativeness: $\lambda(\mu, \psi) = \lambda_0 - \psi^2\lambda_2(\mu) + O(\psi^4)$, where $\lambda_2 = \sigma_\theta^2(m_0^*)^2/(\sigma_\theta^2 + \sigma_\varepsilon^2)^2 > 0$.

(c) Welfare: $W(\mu, \psi) = W_0(\mu) - \psi^2 W_2(\mu) + O(\psi^4)$, where W_0 is the $\psi = 0$ welfare.

(d) Optimal demand: $\mu^*(\psi) = \mu_0^* + \psi^2\tilde{\mu}_2 + O(\psi^4)$, where $\tilde{\mu}_2 > 0$: the welfare-maximizing demand is higher when manipulation is noisy.

Proof. Expand the fixed-point equation $c_m m = \eta\mu\sigma_\theta^2/(\sigma_\theta^2 + \sigma_\varepsilon^2 + \psi^2 m^2)$ in powers of ψ^2 by writing $m = m_0 + \psi^2\tilde{m}_2 + O(\psi^4)$. Matching at order ψ^0 recovers m_0. Matching at order ψ^2:

$$c_m\tilde{m}_2 = -\frac{\eta\mu\sigma_\theta^2 m_0^2}{(\sigma_\theta^2 + \sigma_\varepsilon^2)^2}, \quad \text{so} \quad \tilde{m}_2 < 0.$$

Parts (b) and (c) follow by substituting the manipulation expansion into λ and W. For part (d), the perturbed optimum satisfies $W'(\mu^*(\psi), \psi) = 0$. Expanding around μ_0^*:

$$\tilde{\mu}_2 = \frac{W_2'(\mu_0^*)}{W_0''(\mu_0^*)}.$$

Since $W_0''(\mu_0^*) < 0$ and $W_2'(\mu_0^*) = \eta\mu_0^*\lambda_2(\mu_0^*) \cdot (-\eta\lambda_0(1/c_a + 1/c_m)) < 0$, we have $\tilde{\mu}_2 > 0$. □

Proposition 18 (Strict Concavity for Small ψ)**.** *For ψ sufficiently small, $W(\mu, \psi)$ is strictly concave in μ on any bounded interval $[0, M]$, and $\mu^*(\psi)$ is the unique welfare maximum.*

Proof. $W''(\mu, \psi) = W_0''(\mu) - \psi^2 W_2''(\mu) + O(\psi^4)$. Since $W_0''(\mu) = -(\eta\lambda_0)^2(1/c_a + 1/c_m) < 0$ and W_2'' is continuous and bounded on $[0, M]$, for ψ small enough $W''(\mu, \psi) < 0$ for all $\mu \in [0, M]$. □

Remark 19. Proposition 6(f) provides the exact global characterization: for any $\psi > 0$, W is strictly concave up to the inflection point $\mu_{\text{inf}} > \mu^*$. Proposition 18 is subsumed by the stronger result but is retained because the perturbation expansion gives explicit ψ-dependence of the concavity bound.

Proposition 20 (Perturbed If-and-Only-If Characterization)**.** *For ψ sufficiently small, increasing ESG demand from μ improves welfare if and only if*

$$\mu < \mu^*(\psi) = \mu_0^*\left(1 + \psi^2 \cdot \frac{(\eta\mu_0^*)^2\sigma_\theta^4}{c_m^2(\sigma_\theta^2 + \sigma_\varepsilon^2)^3}\right) + O(\psi^4).$$

Proof. By Proposition 18, W is strictly concave for small ψ, so $W'(\mu) > 0$ if and only if $\mu < \mu^*(\psi)$. The formula for $\mu^*(\psi)$ follows from Proposition 17(d). □

Remark 21. Corollary 7 provides the exact closed-form for μ^* at all $\psi \geq 0$, subsuming this perturbation approximation. The perturbation expansion is retained because it shows the first-order ψ-dependence explicitly and confirms the direction of the correction.

The direction of the correction ($\tilde{\mu}_2 > 0$) has a clear economic logic: noisier manipulation technology reduces the equilibrium manipulation level (firms get less per unit of effort), which reduces manipulation waste, allowing the planner to tolerate more ESG demand before hitting the welfare peak. The result is specific to the proportional noise specification ($g(m) = m^2$). Under $g(m) = m^\alpha$, the correction term is proportional to $\alpha - 1$ and reverses for sublinear noise ($\alpha < 1$).

E Supplementary Results

Sorting Sharpe Ratio

The sorting Sharpe ratio (in pollution units, not return units) measures how effectively the ESG sort separates genuinely green from brown firms:

$$\mathrm{SR}_{\mathrm{GMB}}^{\mathrm{sort}} \propto \sqrt{\frac{\lambda(m^*)}{1-\lambda(m^*)}}.$$

Since $\lambda(m^*(\mu))$ is decreasing in μ and $\mathrm{SR}_{\mathrm{GMB}}^{\mathrm{sort}}$ is increasing in λ, the sorting effectiveness declines as ESG demand grows.

Manipulation-to-Abatement Ratio

Define $r(\mu) \equiv m^*/a^*$. Using $c_m m^* = \eta\mu\lambda(m^*)$:

$$r = \frac{c_a}{c_m} \cdot \frac{\eta\mu\lambda(m^*)}{\eta\mu\lambda(m^*) + \delta}.$$

The ratio r is increasing in μ (the share of manipulation in total ESG effort rises with demand), with $\lim_{\mu\to 0} r = 0$ and $\lim_{\mu\to\infty} r = c_a/c_m$. Manipulation exceeds abatement ($r > 1$) if and only if $c_a > c_m$ and μ exceeds a threshold.

If-and-Only-If Characterization ($\psi = 0$)

When $\psi = 0$, increasing ESG demand from μ improves welfare if and only if

$$\mu < \mu^*_{\psi=0} = \frac{\dfrac{\delta_S - \delta}{c_a} + \sigma_\theta^2}{\eta\lambda_0\left(\dfrac{1}{c_a} + \dfrac{1}{c_m}\right)}.$$

The condition is equivalent to: the uninternalized externality plus allocative value exceeds the marginal manipulation and abatement cost waste. ESG demand is welfare-destroying from the start ($\mu^* = 0$) if and only if $\delta_S = \delta$ and $\sigma_\theta^2 = 0$ (externality fully internalized and firms are identical).

Marginal Welfare Gain Is Independent of Manipulation Noise

Proposition 22 (Noise-Independence of the Marginal Welfare Condition)**.** *For all $\psi \geq 0$,*

$$\left.\frac{dW}{d\mu}\right|_{\mu=0} = \eta\lambda_0\left[\frac{\delta_S - \delta}{c_a} + \sigma_\theta^2\right],$$

where $\lambda_0 = \sigma_\theta^2/(\sigma_\theta^2+\sigma_\varepsilon^2)$. In particular, some ESG demand is welfare-improving ($dW/d\mu|_{\mu=0} > 0$) if and only if $\delta_S > \delta$ or $\sigma_\theta^2 > 0$, and this condition is independent of ψ.

Proof. The proof proceeds in three steps.

Step 1: Equilibrium at $\mu = 0$. The fixed-point equation is $c_m m^* = \eta\mu\lambda(m^*)$. At $\mu = 0$, $m^* = 0$ is the unique solution (the right-hand side vanishes). Therefore $\lambda(0) = \sigma_\theta^2/(\sigma_\theta^2 + \sigma_\varepsilon^2 + \psi^2 \cdot 0^2) = \lambda_0$ regardless of ψ.

Step 2: $dm^*/d\mu|_{\mu=0}$ is ψ-independent. Differentiating the fixed-point equation with respect to μ:

$$c_m\frac{dm^*}{d\mu} = \eta\lambda(m^*) + \eta\mu\lambda'(m^*)\frac{dm^*}{d\mu}.$$

Evaluate at $\mu = 0$ (where $m^* = 0$). The derivative $\lambda'(m) = -2\psi^2 m\sigma_\theta^2/(\sigma_\theta^2 + \sigma_\varepsilon^2 + \psi^2 m^2)^2$ satisfies $\lambda'(0) = 0$ for every $\psi \geq 0$. Substituting:

$$c_m\left.\frac{dm^*}{d\mu}\right|_{\mu=0} = \eta\lambda_0 + 0 \implies \left.\frac{dm^*}{d\mu}\right|_{\mu=0} = \frac{\eta\lambda_0}{c_m}.$$

This is identical to the $\psi = 0$ result.

Step 3: Welfare derivative at $\mu = 0$. The welfare derivative is:

$$\frac{dW}{d\mu} = \left[\frac{\eta(\delta_S - \delta - \eta\mu\lambda)}{c_a} + \eta\sigma_\theta^2\right](\lambda + \mu\lambda'\dot{m}) - \eta\mu\lambda\dot{m},$$

where $\dot{m} \equiv dm^*/d\mu$. At $\mu = 0$: $m^* = 0$, $\lambda = \lambda_0$, $\lambda'(0) = 0$, and $\eta\mu\lambda = 0$. The term $\mu\lambda'\dot{m}$ vanishes because $\mu = 0$, and the term $\eta\mu\lambda\dot{m}$ vanishes for the same reason. Therefore:

$$\left.\frac{dW}{d\mu}\right|_{\mu=0} = \left[\frac{\eta(\delta_S - \delta)}{c_a} + \eta\sigma_\theta^2\right]\lambda_0 = \eta\lambda_0\left[\frac{\delta_S - \delta}{c_a} + \sigma_\theta^2\right].$$

The expression contains λ_0 but not ψ. The sign is positive if and only if $\delta_S > \delta$ or $\sigma_\theta^2 > 0$. Since neither condition involves ψ, the threshold for beneficial ESG demand is ψ-independent.

□

Remark 23. The intuition is that at $\mu = 0$ there is no ESG demand and hence no incentive to manipulate, so $m^* = 0$ regardless of ψ. The manipulation noise technology is irrelevant when no manipulation occurs. Introducing a small amount of ESG demand generates first-order benefits (correcting the externality and improving information) but only second-order manipulation costs (because m^* starts at zero with zero slope sensitivity to ψ). The ψ parameter affects the curvature of the welfare function away from $\mu = 0$ — determining how fast welfare deteriorates and where the optimal μ^* lies (Proposition 17) — but not whether ESG demand is beneficial at the margin.

Asymptotic Behavior of the Total Sharpe Ratio

Proposition 24 (Asymptotic Behavior of the Total Sharpe Ratio)**.** *Fix $\psi > 0$, $\phi > 0$, and all other parameters. As $\mu \to \infty$:*

(a) The fundamental Sharpe ratio satisfies $\mathrm{SR}_{\mathrm{GMB}}^{\mathrm{CF}} \to 0$.

(b) The total Sharpe ratio satisfies $\mathrm{SR}_{\mathrm{GMB}}^{total} \to \infty$. The taste premium eventually dominates the information degradation channel.

(c) There exists a finite threshold $\bar{\mu}$ such that $\mathrm{SR}_{\mathrm{GMB}}^{total}$ is strictly increasing in μ for all $\mu > \bar{\mu}$. In the leading-order asymptotic approximation, $\bar{\mu} = \gamma/(2\phi)$.

Proof. Part (a). Since m^* is strictly increasing in μ (Proposition 3(c)), $\psi^2(m^*)^2 \to \infty$ and $\lambda(m^*(\mu)) \to 0$. The numerator $\gamma\sqrt{\lambda}\,\sigma_\theta \to 0$ while the denominator converges to $\sqrt{\gamma^2\sigma_\theta^2 + \sigma_x^2} > 0$.

Asymptotic scaling. For large μ, $\psi^2(m^*)^2$ dominates the denominator of the fixed-point (10), giving $c_m m^* \approx \eta\mu\sigma_\theta^2/(\psi^2(m^*)^2)$ and hence

$$m^* \sim \left(\frac{\eta\sigma_\theta^2}{c_m\psi^2}\right)^{1/3} \mu^{1/3}, \qquad \lambda(m^*(\mu)) \sim C_\lambda\, \mu^{-2/3},$$

where $C_\lambda = \sigma_\theta^2(c_m/(\eta\sigma_\theta^2))^{2/3}\psi^{-2/3} > 0$.

Asymptotic derivative of λ. Implicit differentiation of $c_m m = \eta\mu\lambda(m)$ and the asymptotic $\eta\mu|\lambda'(m^*)| \to 2c_m$ (from $\lambda'(m) \approx -2\sigma_\theta^2/(\psi^2(m^*)^3)$ and $m^{*3} \approx \eta\mu\sigma_\theta^2/(c_m\psi^2)$) give $dm^*/d\mu \approx m^*/(3\mu)$ and

$$\frac{d\lambda}{d\mu} \approx -\frac{2\lambda}{3\mu}.$$

Parts (b) and (c). The total Sharpe ratio numerator is $N(\mu) = (\gamma + \mu\phi)\sqrt{\lambda(\mu)}$. Then

$$\frac{dN}{d\mu} = \phi\sqrt{\lambda} + (\gamma + \mu\phi)\frac{d\lambda/d\mu}{2\sqrt{\lambda}} \approx \sqrt{\lambda}\left[\frac{2\phi}{3} - \frac{\gamma}{3\mu}\right].$$

This is positive for $\mu > \gamma/(2\phi)$. Since $N(\mu) \sim \phi C_\lambda^{1/2} \mu^{2/3} \to \infty$ and the denominator converges to a positive constant, $\mathrm{SR}_{\mathrm{GMB}}^{\mathrm{total}} \to \infty$ and is eventually strictly increasing. □

Remark 25 (Correction to main-text characterization). The main text (Section 5) states that at high ESG demand, the information degradation channel overtakes the taste channel and the total Sharpe ratio declines. Proposition 24 shows this is incorrect asymptotically: the taste premium grows linearly in μ while $\sqrt{\lambda}$ decays only as $\mu^{-1/3}$, producing net numerator growth of $\mu^{2/3}$. The condition for decline (22) fails for all $\mu > \gamma/(2\phi)$. The correct picture: (i) for $\phi = 0$, the SR declines monotonically to zero; (ii) for $\phi > 0$ and large μ, the taste premium dominates and the SR rises without bound. The testable prediction P2 (declining GMB Sharpe ratio) applies to the fundamental cash-flow component, not the total SR.