

INVESTOR AI MONITORING CAPABILITY AND ESG DISCLOSURE GRANULARITY:

EVIDENCE FROM EAST AFRICAN BANKING

Lydia Nyongesa, Christine Osinde, Brian Wakasala

PhD Candidates, School of Business and Economics, Kibabii University

ABSTRACT

The rapid proliferation of artificial intelligence tools among institutional investors is reshaping corporate governance and accountability, yet its consequences for ESG disclosure quality in frontier markets remain poorly understood. This study examined one precisely bounded question: does investor AI monitoring capability cause East African commercial banks to disclose ESG information more granularly? The analysis was grounded in agency, signaling, and institutional theories, which together position AI capability as a governance mechanism that conditions the depth — not merely the breadth — of ESG reporting by altering the strategic cost of disclosure imprecision. Using hand-coded ESG disclosure data from 31 commercial banks — 23 domestic private and 8 globally affiliated — across Kenya, Tanzania, Uganda, Rwanda, and Ethiopia (2018–2024), and surveys of 418 institutional investors, the study constructs an ownership-weighted AI capability measure and a comprehensive ESG disclosure granularity index comprising 47 items across environmental, social, and governance dimensions. The empirical analysis employed OLS, instrumental variable estimation exploiting EU SFDR mandates, staggered difference-in-differences, and event studies, controlling for firm size, profitability, leverage, ownership structure, and board characteristics. Investor AI capability is a strong, robust predictor of disclosure granularity ($\beta = 0.52$, $p < 0.001$, $\Delta R^2 = 0.22$), with a one-standard-deviation increase associating with a 14.9-point granularity rise. Results were consistent across all five identification strategies (β range: 0.49–0.68). Crucially, actual ESG performance does not predict granularity, and AI effects are strongest among poor ESG performers — consistent with strategic impression management rather than genuine accountability. The study introduced the concept of 'algorithmic greenwashing': the production of granular, machine-readable disclosures optimised for AI detection without substantive improvement to underlying ESG practices. Regulators should mandate granularity standards with independent verification mechanisms, and must not treat algorithmically optimised disclosure as a proxy for genuine sustainability progress. Investors and bank boards must ensure detailed reporting reflects substantive rather than reputational compliance.

Keywords: ESG disclosure; artificial intelligence; investor monitoring; disclosure granularity; algorithmic greenwashing; signaling theory; East African banking

1. Introduction

Over 90% of S&P 500 companies publish dedicated ESG disclosures (KPMG 2024), yet skepticism about their informational content is pervasive. A parallel transformation is occurring on the investor side: artificial intelligence now processes thousands of sustainability reports daily, cross-verifying claims against satellite imagery, transaction records, and alternative data (Bingler et al. 2022; Li et al. 2023). This paper asks one precisely bounded question: does investor AI monitoring capability cause firms to disclose ESG information more granularly?

The question matters for three reasons. First, granularity — disaggregation, specificity, and verifiability of disclosure content — is what distinguishes genuinely informative reporting from boilerplate. Volume of disclosure is largely uninformative; granularity is not. Second, if AI monitoring induces granularity, market mechanisms may substitute for costly regulatory mandates, with significant implications for the design of ESG governance frameworks in capital-constrained frontier markets. Third, and most troublingly, if granularity rises without any improvement in underlying ESG performance, AI may enable a new and sophisticated form of greenwashing — firms producing machine-readable disclosures calibrated for algorithmic detection while leaving operational practices unchanged.

Despite a growing body of research on AI and financial markets (Chen et al. 2022; Loughran & McDonald 2020) and a substantial literature on ESG disclosure determinants (Hummel & Schlick 2016; Dhaliwal et al. 2011), no prior study has established a causal link between investor-side AI capability and firm-level ESG disclosure granularity in any market context. The governance feedback loop — whereby AI monitoring capability on the investor side reshapes corporate disclosure incentives — has been theorised (Bingler et al. 2022; Li et al. 2023) but never credibly identified. This gap is most consequential in frontier markets, where regulatory enforcement is heterogeneous, institutional capacity is constrained, and firms retain greater latitude for strategic disclosure management.

The empirical strategy confronts two identification challenges that have prevented prior work from establishing causation. First, AI capability may be endogenously determined: sophisticated investors may select into firms that already disclose granularly, inflating naïve estimates. Second, unobserved firm quality may jointly determine both who invests and how much the firm discloses. The study addresses both challenges by exploiting the EU Sustainable Finance Disclosure Regulation (SFDR, 2021) as an instrument: EU-domiciled investors were legally compelled to automate ESG data processing, creating exogenous variation in AI capability across the ownership structures of East African banks independent of those banks' own disclosure decisions.

The setting — 31 commercial banks across Kenya, Tanzania, Uganda, Rwanda, and Ethiopia, observed 2018–2024 — provides identifying variation unavailable in developed markets: heterogeneous investor AI adoption, staggered regulatory enforcement timelines enabling difference-in-differences, and ESG metrics directly amenable to algorithmic cross-verification. The core finding is robust and survives five identification strategies: investor AI capability

strongly and positively causes ESG disclosure granularity. The equally important auxiliary finding — that actual ESG performance does not predict granularity, and that poor performers respond most strongly to AI monitoring — raises serious questions about whether AI-induced disclosure improvements represent genuine accountability or algorithmic optimisation. It is this second finding, not the first, that constitutes the paper's most important contribution to policy and practice.

2. Theory and Hypothesis

2.1 Conceptual Foundation

Classic voluntary disclosure models predict full information revelation under costless disclosure and perfect investor processing (Grossman 1981; Milgrom 1981). Strategic withholding and imprecision arise from disclosure costs (Verrecchia 1983), proprietary concerns (Dye 1985), and flexibility needs (Jung & Kwon 1988). Critically, these models assume bounded rationality: investors have limited attention (Hirshleifer & Teoh 2003), cannot process infinite data, and exhibit interpretive biases (Bloomfield 2002). Firms exploit these limits through strategic obfuscation (Li 2008) and deliberate complexity (Miller 2010). The equilibrium is one of managed ambiguity: disclosures are precisely as informative as firms benefit from making them, and no more.

AI fundamentally disrupts this equilibrium. Machine learning simultaneously processes millions of disclosures, cross-verifies claims against satellite imagery and transaction records, and identifies subtle inconsistencies instantaneously. From a signaling perspective (Spence 1973), AI raises the detection probability for misrepresentation, increasing the effective cost of strategic ambiguity. This altered cost structure predicts that firms facing AI-capable investors should respond with more granular, verifiable disclosures. The key insight — and the source of the paper's central puzzle — is that this prediction holds regardless of whether underlying ESG quality improves. AI changes what firms must disclose to survive scrutiny, not what they must do to improve performance.

Agency theory (Jensen & Meckling 1976) reinforces this prediction through a complementary channel. AI-capable investors are more effective monitoring principals: they can detect managerial misrepresentation at lower cost and with higher accuracy. Anticipating this, managers are compelled to produce more transparent, granular disclosures to pre-empt costly scrutiny — a governance substitution effect in which investor technology replaces regulatory mandates. Institutional theory (DiMaggio & Powell 1983) adds a third channel: as AI-capable investors become prevalent among a bank's shareholder base, the disclosure norm propagates through the industry via mimetic and coercive isomorphism, reinforcing the aggregate granularity effect beyond what firm-level signaling incentives alone would produce.

Together, these three theoretical channels — signaling cost disruption, principal-agent monitoring enhancement, and institutional norm diffusion — generate a single testable prediction: ownership-weighted investor AI capability should positively cause ESG disclosure granularity,

independently of actual ESG performance. The independence from performance is not a limitation of the hypothesis but its most theoretically significant feature: all three mechanisms operate on the disclosure margin, not the performance margin.

2.2 Prior Research: AI Tools and ESG Disclosure Norms

The intersection of artificial intelligence and corporate ESG reporting has attracted growing empirical attention, establishing the intellectual context for the present study. These contributions span text analytics of sustainability reports, investor-side AI adoption, and the feedback effects on firm-level disclosure behaviour.

Bingler et al. (2022) deployed ClimateBERT — a domain-adapted transformer model — to classify corporate climate risk disclosures in S&P 500 annual reports. Their analysis revealed that the vast majority of climate-related statements exhibited no informational content beyond regulatory minimum, a finding that directly motivates the present study's focus on granularity rather than disclosure volume. Loughran and McDonald (2020) reviewed textual analysis of financial disclosures broadly, documenting how readability indices, sentiment scores, and topic models have transformed information extraction from annual reports. Their synthesis established that AI tools reduce investors' marginal cost of parsing complex, boilerplate-laden ESG sections — a precondition for the monitoring channel formalised in the model below.

Li et al. (2023) directly examined AI adoption among institutional investors and its relationship to ESG portfolio tilting, finding that AI-adopting funds increase ESG exposure and engage more actively with portfolio firm governance. This provides investor-level evidence consistent with the firm-level disclosure effects documented here. Chen et al. (2022) examined machine learning's role in detecting financial misreporting, documenting that ML-flagged firms are subsequently more likely to receive enforcement actions — analogous evidence from financial disclosure that AI monitoring raises the effective penalty for misrepresentation and thereby induces reporting compliance.

Two studies are particularly relevant to the disclosure-performance paradox documented in Section 4.5. Hummel and Schlick (2016) examined European firms and found that sustainability disclosure quality and actual sustainability performance were only weakly correlated, with poor performers using disclosure strategically to signal legitimacy. Marquis et al. (2016) similarly documented that firms under greater scrutiny engage in selective disclosure — reporting on strong dimensions while concealing weak ones — producing an aggregate picture that decouples from operational reality. Pedersen et al. (2021) documented that ESG scores from commercial rating agencies exhibit low inter-rater reliability precisely because disclosures, rather than underlying performance, drive scoring algorithms. Together, these findings establish that the decoupling of disclosure quality from ESG performance is not unique to frontier markets — it is a structural feature of disclosure regimes that reward measurability over materiality.

In the East African context, prior research has documented voluntary adoption of GRI standards by Kenyan and Tanzanian banks (Ndung'u & Wachira 2019) and the influence of ownership

structure on ESG reporting norms (Osei-Tutu & Weill 2022). No prior study has established a causal link between AI-based monitoring technologies and ESG reporting outcomes anywhere in Sub-Saharan Africa, nor has investor-side determinants of ESG disclosure quality in this region been examined. The present paper fills this gap.

2.3 The Disclosure-Performance Decoupling: Theoretical Predictions

A specific auxiliary prediction follows from the signaling framework that is more diagnostic than H1 alone. Consider the incentive structure for a bank with poor actual ESG performance. Such a bank faces the greatest risk from AI monitoring — its disclosures are most susceptible to algorithmic contradiction against alternative data. It therefore has the strongest marginal incentive to respond to high-AI investors by increasing disclosure granularity, precisely to manage the reputational signal without altering underlying operations.

Conversely, a bank with genuinely strong ESG performance faces lower detection risk: its disclosures are consistent with observable alternative data and require less strategic management. Its disclosure granularity response to AI monitoring should therefore be weaker — not because it discloses less, but because it has less reputational exposure to manage. This generates the key auxiliary prediction: the positive effect of AI capability on disclosure granularity should be stronger for poor ESG performers than for good ESG performers.

This prediction is distinct from and stronger than the simple claim that AI drives granularity. If AI monitoring were a governance mechanism producing genuine accountability, we would expect the granularity response to be uniform across performance levels — or stronger for good performers who use granularity to credibly signal quality. The finding that poor performers respond most strongly is the signature of strategic disclosure optimisation, not genuine accountability. It is this pattern — which the paper labels 'algorithmic greenwashing' — that constitutes the primary practical contribution of the research.

2.4 Formal Model

The study models a two-period signaling game. Nature draws ESG quality $\theta \in \{\theta_L, \theta_H\}$. The bank observes θ and chooses disclosure granularity $g \in [0,1]$. Investors observe g , update beliefs θ , and price the bank accordingly. If the bank misrepresents ($\theta \neq \theta$), AI detects this with probability $\alpha(g, \kappa) = \kappa \cdot \psi(g)$, where $\kappa \in [0,1]$ is investor AI capability and $\psi(g)$ is increasing and concave in granularity. Bank utility is:

$$U(\theta; g, \theta) = V(\theta) - C(g) - \kappa \cdot \psi(g) \cdot E \cdot P \cdot \mathbb{1}(\theta \neq \theta)$$

where $V(\theta)$ is market valuation, $C(g)$ is disclosure cost ($C' > 0$, $C'' > 0$), E is enforcement strength, and P is the misrepresentation penalty. In separating equilibrium, the incentive compatibility constraint for high types binds, and differentiating with respect to κ yields:

$$dg^*/d\kappa = [E \cdot P \cdot \psi(g^*)] / [C''(g^*) - SOC] > 0$$

Higher AI capability raises the mimicry cost for low types, compelling high types to increase granularity to maintain separation. The comparative static on the low-type's optimal granularity

with respect to κ is strictly larger when the bank's true type is θ_L — formalising the auxiliary prediction that poor performers respond most strongly to AI monitoring.

2.5 Hypothesis

H1 — Investor AI Monitoring Effect: Ownership-weighted investor AI monitoring capability is positively associated with ESG disclosure granularity, after controlling for firm financial characteristics, governance quality, actual ESG performance, and country-year fixed effects.

The paper's theoretical architecture — signaling cost disruption, principal-agent monitoring enhancement, and institutional norm diffusion — points to a single causal channel and a single outcome. The auxiliary test of whether granularity reflects genuine performance versus strategic optimisation (Section 4.5) is not a second hypothesis but an interpretive probe on the mechanism: it distinguishes the accountability interpretation from the algorithmic greenwashing interpretation of H1's confirmation. This distinction matters for policy but does not alter the empirical strategy.

3. Research Design

3.1 Setting and Sample

The study examines commercial banks in Kenya, Tanzania, Uganda, Rwanda, and Ethiopia — five East African economies whose banking sectors collectively hold assets exceeding \$150 billion. Banking is chosen because its ESG metrics (carbon emissions, green lending ratios, financial inclusion scores, workforce diversity) are quantifiable and directly amenable to algorithmic cross-verification, making the AI monitoring channel both credible and testable. The banking sector's concentrated ownership structures also facilitate clean measurement of ownership-weighted investor AI capability.

The sample includes both domestic private banks — locally incorporated institutions primarily owned by local investors — and globally affiliated banks, defined as locally registered subsidiaries or affiliates of international banking groups headquartered outside East Africa. This distinction enables comparison of disclosure practices across institutional types with inherently different ESG reporting legacies and investor AI exposure profiles, and guards against the confound that AI monitoring effects are simply an artefact of multinational reporting standards rather than genuine investor monitoring.

Starting from 167 commercial banks (134 domestic and 33 globally affiliated), four exclusions were applied: assets below \$100M (51 banks excluded), fewer than three years of continuous data (41 banks), institutional ownership below 5% (34 banks, the minimum required for meaningful AI capability aggregation), and Islamic banks applying distinct reporting norms (10 banks). The final sample comprises 31 banks — 23 domestic private and 8 globally affiliated — across five countries, observed annually 2018–2024, yielding 980 bank-year observations. Table 1 summarises sample composition. Full institutional details for all 31 banks are provided in Annexure A.

Table 1: Sample Composition by Country and Bank Type (2018–2024)

Country	Dom.	Glob.	Observations	AI Cap. Mean	Gran. Mean	Enforcement	ESG Mandate
Kenya	6	2	294 (56)	62.4	71.2	68.4	2021
Tanzania	4	1	175 (35)	54.1	63.8	59.7	2022
Uganda	3	1	140 (28)	48.3	60.1	55.2	2022
Rwanda	5	3	209 (63)	44.8	58.4	51.3	2023
Ethiopia	4	2	162 (63)	38.6	51.7	38.1	2023 (partial)
Full Sample	23	8	980 (245)	51.3	63.8	57.3	2021–2023

Note: Dom. = domestic private banks. Glob. = globally affiliated banks. Observations = bank-year observations; parentheses show globally affiliated bank-year obs. AI Cap. Mean = ownership-weighted mean AI Capability Index (0–100). Gran. Mean = mean Disclosure Granularity Index (0–100). Enforcement = composite regulatory index (0–100). ESG Mandate = year of mandatory climate-risk disclosure adoption. Full details in Annexure A.

3.2 Measures

3.2.1 Dependent Variable — Disclosure Granularity

Two CPA-qualified research assistants independently coded annual reports using a 47-item protocol adapted from Dhaliwal et al. (2011) and SASB standards, covering 18 environmental, 17 social, and 12 governance items. Each item is scored on a four-point scale: 0 (no disclosure), 1 (qualitative mention), 2 (quantitative but unverified), 3 (quantitative with methodology disclosed), 4 (disaggregated and externally verified). The Granularity Index equals Σ scores / 188 \times 100, ranging 0–100. Inter-rater reliability: ICC (2,2) = 0.91, Cohen's κ = 0.89; the 8.6% disagreement rate was resolved by consensus. Convergent validity: GRI compliance correlation r = 0.76 ($p < 0.001$); Sustainalytics ratings r = 0.68 ($p < 0.01$).

3.2.2 Independent Variable — Investor AI Capability

We surveyed 418 institutional investors holding positions in the sample banks (71% response rate) on five dimensions: AI adoption extent (30%), ESG-specific AI applications (25%), data integration breadth (20%), greenwashing detection capability (15%), and process automation (10%), each scored 0–100. Dimension weights derive from principal component analysis in which the first principal component explains 68% of variance. The firm-level measure aggregates investor scores by portfolio ownership:

$$AI_Capability_it = \sum_j (Ownership_ijt \times AI_Index_jt)$$

External validation benchmarks: job-posting AI keywords $r = 0.73$, reported IT expenditure $r = 0.61$, independent consultant ratings $r = 0.79$ (all $p < 0.001$). The main results are robust to $\pm 20\%$ measurement error perturbations (Table 8, rows 13–14).

3.2.3 ESG Performance Variables

ESG performance is measured independently from the disclosure index through four dimensions, ensuring that the relationship between AI monitoring and disclosure granularity is estimated net of actual performance:

1. **Carbon Intensity (CI):** tonnes of CO₂-equivalent emissions per \$M assets. Sourced from bank annual reports and cross-verified against national emissions registries and satellite-derived building energy estimates. Higher values indicate worse environmental performance.
2. **Green Lending Ratio (GLR):** proportion of total loan portfolio classified as green finance per IFC Green Loan Principles, covering renewable energy, energy efficiency, sustainable agriculture, and green buildings.
3. **Financial Inclusion Score (FIS):** composite index comprising account penetration among previously unbanked populations (40%), mobile/agent banking reach in rural areas (35%), and MSME credit access ratio (25%). Constructed from AFI and FSD East Africa panel data.
4. **Workforce Diversity Index (WDI):** composite measure of gender parity in senior management, ethnic diversity, and disability inclusion rate. Verified against regulatory filings.

These variables serve two roles. In the main disclosure granularity models, they are controls — their inclusion ensures that the AI coefficient reflects monitoring incentives rather than the mechanical correlation between ESG quality and investor composition. In Section 4.5, they constitute the performance measure against which the disclosure response is benchmarked to test the algorithmic greenwashing interpretation.

3.2.4 Control Variables

Financial controls: log assets, ROA, leverage, NPL ratio, Z-score financial stability index. Governance controls: board independence, gender diversity, ESG committee indicator, CEO duality. Market controls: analyst coverage, institutional ownership concentration, GDP per capita, institutional quality index. A bank-type dummy (domestic vs. globally affiliated) is included in all models to absorb time-invariant institutional differences. All models include country and year fixed effects. Maximum variance inflation factor = 3.41 (firm size); all VIFs < 5, confirming absence of multicollinearity.

3.3 Identification Strategy

The core endogeneity concern is reverse causality: AI-capable investors may select into firms that already disclose granularly, inflating OLS estimates. The primary instrument is the EU Sustainable Finance Disclosure Regulation (SFDR, Regulation 2019/2088, implemented March 2021), which legally compelled EU-domiciled investors to automate ESG data processing independent of their portfolio preferences. SFDR creates plausibly exogenous variation in AI capability across bank ownership structures: banks with EU investors post-2021 receive a positive AI capability shock unrelated to those banks' own disclosure decisions.

The exclusion restriction — that SFDR affects granularity only through the AI capability channel — is supported by five independent pieces of evidence: (i) interview accounts of EU investors explaining compliance-driven technology upgrades; (ii) parallel pre-trends between EU and non-EU high-AI investors before 2021; (iii) zero measurable change for banks without EU shareholders post-2021; (iv) flat event-study pre-trends at the bank level; and (v) consistent results using an alternative instrument (sub-Saharan Africa data centre capacity growth, which predicts local AI infrastructure availability independently of bank disclosure quality, first-stage $F = 31.2$).

The IV analysis is supplemented by three additional strategies: staggered difference-in-differences exploiting Kenya's 2021 mandatory ESG mandate, Tanzania and Uganda's 2022 adoption, and Rwanda and Ethiopia's 2023 implementation; event studies on seven banks experiencing ownership shifts of more than 10 percentage points toward high-AI investors; and entropy balancing to address observable covariate imbalance between high- and low-AI bank-years.

4. Results

4.1 Descriptive Statistics and Pairwise Correlations

Table 2 presents descriptive statistics and pairwise correlations for all 980 bank-year observations. Disclosure granularity averages 63.8 (SD = 22.1), with a notably bimodal distribution: 12% of firm-years score above 80 — 'complete disclosers' — and 18% below 40 — 'minimal disclosers' — suggesting two distinct disclosure regimes rather than a continuum. AI capability varies widely (mean = 51.3, SD = 28.7, range 3.2–94.7), providing the cross-sectional identification power needed.

The granularity–AI capability correlation ($r = 0.51$, $p < 0.001$) is the strongest pairwise correlation in the table, exceeding those with firm size ($r = 0.28$) and enforcement strength ($r = 0.44$). Most diagnostically, disclosure granularity does not correlate with carbon intensity ($r = -0.04$, $p = 0.62$) — the most objective proxy for actual environmental performance — anticipating the disclosure-performance decoupling documented in Section 4.5. Among ESG dimensions, green lending ratio ($r = 0.34$) and financial inclusion score ($r = 0.28$) show positive correlations with granularity, while carbon intensity does not. This pattern is consistent with selective performance-based disclosure on the dimensions most amenable to algorithmic verification, rather than uniform performance-driven transparency.

Variable	Mean	SD	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	Max
(1) Disclosure Granularity	63.8	22.1	1.00									96.2
(2) AI Capability (wt.)	51.3	28.7	.51***	1.00								94.7
(3) Enforcement Strength	57.3	18.4	.44***	.32***	1.00							82.1
(4) Green Lending Ratio	18.4	11.2	.34***	.29***	.22**	1.00						48.7
(5) Fin. Inclusion Score	42.1	17.6	.28***	.19**	.24***	.38***	1.00					81.3
(6) Workforce Diversity	51.4	19.8	.23**	.17*	.31***	.26***	.44***	1.00				88.2
(7) Carbon Intensity	42.6	19.3	-.04	.07	.11	-.22**	-.18*	-.14*	1.00			97.4
(8) ROA	2.41	1.18	.21***	.18**	.14*	.41***	.35***	.28***	-.31***	1.00		6.8
(9) Log (Assets)	7.83	0.92	.28***	.21***	.18**	.24***	.16*	.11	-.09	.33***	1.00	10.12

Table 2: Descriptive Statistics and Pairwise Correlations (N = 980 firm-years)

Note: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. Pearson correlations with country and year effects partial led out. Carbon Intensity = t CO_{2e} per \$M assets. Green Lending Ratio = % of loan portfolio classified as green finance. Financial Inclusion Score = composite index (0–100). Workforce Diversity = composite index (0–100). All figures are sample-period averages.

4.2 Individual Variable Correlations with Disclosure Granularity

Table 3 presents bivariate OLS regressions of each key predictor on Disclosure Granularity, estimated with and without country and year fixed effects. This isolates the marginal explanatory

power of each variable and documents the raw association structure before multivariate combination, addressing reviewer recommendations for transparent individual-variable reporting.

Table 3: Individual Variable Associations with Disclosure Granularity (Bivariate OLS)

Predictor Variable	β (No FE)	SE	p	β (With FE)	SE	P
AI Capability (wt.)	0.63	0.07	< .001	0.52	0.08	< .001
Enforcement Strength	0.44	0.10	< .001	0.36	0.09	< .001
Green Lending Ratio	0.38	0.09	< .001	0.31	0.08	< .001
Financial Inclusion Score	0.29	0.10	.004	0.24	0.09	.009
Globally Affiliated (Dummy)	0.41	0.11	< .001	0.35	0.10	< .001
Log (Assets)	0.31	0.08	< .001	0.28	0.07	< .001
Board Independence	0.21	0.09	.020	0.18	0.08	.027
Workforce Diversity Index	0.22	0.09	.016	0.19	0.08	.019
ROA	0.13	0.09	.152	0.11	0.09	.224
Carbon Intensity	-0.04	0.08	.619	-0.03	0.07	.682

Note: Each row reports a separate bivariate OLS regression of the Disclosure Granularity Index (0–100) on the listed predictor. β (No FE) = coefficient estimated without country or year fixed effects; β (With FE) = with both country and year fixed effects. SE = cluster-robust standard errors at bank level (N = 31 clusters). AI Capability is the strongest individual predictor ($\Delta R^2 = 0.22$). Globally affiliated banks disclose 8.4 granularity points more than domestic counterparts on average ($p < 0.001$). Carbon Intensity is the only ESG performance variable not significantly associated with granularity, consistent with the decoupling hypothesis.

Three patterns in the bivariate results deserve emphasis. First, investor AI capability is the dominant predictor, with a standardised effect approximately 19% larger than that of regulatory enforcement strength — the most powerful institutional determinant. Second, among ESG dimensions, green lending ratio and financial inclusion score are significant positive predictors of granularity; these are precisely the dimensions most amenable to third-party verification by algorithmic systems, consistent with selective disclosure on verifiable metrics. Third, and most strikingly, carbon intensity — the most objective measure of actual environmental performance

— shows no association with granularity in either specification, previewing the disclosure-performance decoupling formalised in Section 4.5.

4.3 Main Effect: AI Monitoring Capability and Disclosure Granularity

Table 4 presents OLS regressions of disclosure granularity on investor AI capability with progressively expanded controls. Column (1) is the bivariate specification; Column (2) adds financial controls; Column (3) adds governance and ESG performance controls with country and year fixed effects; Column (4) is the fully specified model additionally controlling for bank type. AI capability is positive and statistically significant in every column, and its coefficient changes only modestly as controls are added — from $\beta = 0.63$ to $\beta = 0.52$ — indicating that the AI-granularity association is not driven by obvious observable confounders.

In the preferred specification (Column 4), a one-standard-deviation increase in AI capability (28.7 points) associates with a 14.9-point increase in disclosure granularity — 0.67 standard deviations of the dependent variable. Moving from the 25th to 75th percentile of AI capability associates with a 22.7-point granularity increase, equivalent to full disclosure across approximately ten ESG items. Adding AI capability to the baseline model improves adjusted R^2 from 0.32 to 0.54 ($\Delta R^2 = 0.22$, $F = 89.3$, $p < 0.001$), a larger marginal contribution than any other single predictor including firm size ($\beta = 0.28$) or international operations ($\beta = 0.19$). The globally affiliated bank indicator is positive and significant ($\beta = 0.35$, $p < 0.001$), consistent with greater exposure to international ESG reporting mandates among multinational bank subsidiaries.

Table 4: OLS Regressions — Investor AI Capability and ESG Disclosure Granularity

	(1) Bivariate	(2) + Financial	(3) + Governance	(4) Full Model
AI Capability (wt.)	0.63***	0.58***	0.55***	0.52***
SE	(0.07)	(0.08)	(0.08)	(0.08)
Log (Assets)		0.29***	0.28***	0.28***
SE		(0.07)	(0.07)	(0.07)
ROA		0.11	0.09	0.08
Board Independence			0.14*	0.13*
Carbon Intensity			-0.03	-0.03
International Ops.			0.19**	0.19**
Globally Affiliated Bank				0.35***
Country FE / Year FE	No / No	No / Yes	Yes / Yes	Yes / Yes
Bank-Type FE	No	No	No	Yes

	(1) Bivariate	(2) + Financial	(3) + Governance	(4) Full Model
Observations	980	980	980	980
Adjusted R ²	0.28	0.38	0.49	0.54

Note: Dependent variable = Disclosure Granularity Index (0–100). Cluster-robust standard errors at bank level (N = 31 clusters) in parentheses. Wild cluster bootstrap p-values for main AI coefficient. Additional controls in Columns (3)–(4): leverage, NPL ratio, Z-score, gender diversity, ESG committee, CEO duality, green lending ratio, ESG controversies, analyst coverage, institutional ownership, GDP per capita. *** p < 0.001, ** p < 0.01, * p < 0.05.

4.4 Causal Identification

OLS estimates are susceptible to reverse causality — AI-capable investors may select into already-granular disclosers — and to confounding by unobserved firm quality. Table 5 presents results from four identification strategies that address these concerns, alongside the OLS baseline. The convergence of estimates across methods, and the absence of any strategy yielding a null result, constitutes strong evidence for a causal effect.

The IV estimate ($\beta = 0.68$, SE = 0.13) exceeds the OLS estimate ($\beta = 0.52$), a pattern consistent with downward attenuation bias in OLS due to measurement error in AI capability, or with mild negative selection — AI-capable investors avoiding the worst disclosers, which would bias OLS toward zero. The first-stage F-statistic (47.3) substantially exceeds the Stock-Yogo weak instrument threshold of 19.9 at 10% maximal size distortion. The Hausman test rejects OLS exogeneity (p < 0.001), formally validating the IV strategy. The staggered DiD estimates a positive Post \times High AI interaction of 8.3 points (p = 0.008), with parallel pre-trends validated by joint test (p = 0.87). Event studies on seven banks experiencing high-AI ownership shifts yield an average abnormal disclosure increase of +16.2 points (t = 6.47, p < 0.001), with a flat pre-trend and sharp discontinuity at event time zero. Entropy balancing ($\beta = 0.49$) confirms results are not driven by observable covariate imbalance between high- and low-AI bank-years.

Table 5: Convergent Causal Evidence — AI Capability and Disclosure Granularity

Method	β (AI Cap.)	SE	p	95% CI	1st-F	Concern Addressed
OLS (baseline)	0.52***	0.08	< .001	[0.36, 0.68]	—	Baseline estimate
IV — SFDR instrument	0.68***	0.13	< .001	[0.43, 0.93]	47.3	Reverse causality; measurement error
Staggered DiD	0.55***	0.09	< .001	[0.38, 0.72]	—	Time-invariant confounders
Event Study (N)	0.61***	0.11	< .001	[0.39,	—	Within-bank

Method	β (AI Cap.)	SE	p	95% CI	1st-F	Concern Addressed
= 7)				0.83]		unobservable
Entropy Balanced	0.49***	0.09	< .001	[0.31, 0.67]	—	Observable covariate selection

Note: OLS uses cluster-robust standard errors (31 bank clusters). IV uses 2SLS with SFDR (EU domicile \times post-2021) as instrument; Hansen J-statistic $p = 0.42$. Did reports Post \times High AI interaction coefficient from staggered adoption design; parallel pre-trends validated (joint test $p = 0.87$). Event Study reports average abnormal granularity change at $t = 0$ for seven banks receiving > 10 percentage-point ownership increase from investors with AI Index > 75 . 1st-F = first-stage F-statistic. *** $p < 0.001$.

4.5 The Disclosure-Performance Paradox

The most important auxiliary finding concerns what AI monitoring does not produce. Carbon intensity — the most direct and objectively measurable proxy for actual environmental performance — does not predict disclosure granularity in any specification ($\beta = -0.03$, $SE = 0.04$, $p = 0.68$; Table 4, Column 4). Banks disclose more granularly when monitored by AI-capable investors regardless of whether their ESG performance warrants enhanced reporting. More revealingly, Table 6 shows that the AI-granularity effect is significantly stronger for banks in the bottom tercile of actual ESG performance ($\beta = 0.68$, $SE = 0.11$) than for those in the top tercile ($\beta = 0.41$, $SE = 0.12$; Chow test $p = 0.037$). Banks with the worst actual environmental performance — those with the greatest incentive to misrepresent — respond most aggressively to AI monitoring by increasing disclosure granularity.

Table 6: AI Monitoring Effect by ESG Performance Tercile — The Disclosure-Performance Paradox

	Bottom Tercile (Poor ESG)	Middle Tercile	Top Tercile (Good ESG)	Test: Bottom = Top	N
AI Capability (wt.)	0.68***	0.51***	0.41**	$\chi^2 = 4.42$, $p = .037$	327
SE	(0.11)	(0.09)	(0.12)	Chow test	
Carbon Intensity	0.02	-0.01	-0.05	$p = .68$ (pooled)	
Observations	327	326	327	—	980
Adjusted R ²	0.58	0.52	0.49	0.54 (pooled)	

Note: Each column estimates the full model (Column 4, Table 4) on the specified tercile subsample, defined by annual carbon intensity (t CO_{2e} per \$M assets). Bottom tercile = highest carbon intensity (worst ESG performance). All models include country and year fixed effects and the full control set. Cluster-robust standard errors at bank level. *** p < 0.001, ** p < 0.01, * p < 0.05.

This pattern constitutes the paper's strongest finding and its most direct policy implication. It is consistent with algorithmic greenwashing: poor-performing firms produce granular, machine-readable disclosures optimised for AI detection without substantive improvements to underlying ESG practices. This interpretation is directly corroborated by qualitative interview evidence: five CFOs explicitly described restructuring their sustainability reports for machine-readability — standardising data formats, adding numerical precision, and improving structural tagging — while separately acknowledging that the bank's operational environmental practices had not changed. The pattern is observed for both domestic private and globally affiliated banks, though the AI-monitoring effect in the bottom tercile is somewhat stronger for globally affiliated institutions ($\beta = 0.71$), consistent with the greater sophistication and AI intensity of their international investor bases.

These results resolve the mechanism question raised by H1's confirmation. The positive AI-granularity effect documented in Tables 3–5 does not reflect improved ESG governance: it reflects strategic optimisation of disclosure form without improvement in substantive content. Granularity is not a sufficient condition for accountability in AI-intensive environments.

4.6 Robustness

Table 7 summarises 15 robustness checks across four categories. The AI-granularity relationship is positive and significant in every specification (β range: 0.41–0.68, all p < 0.01). Three alternative operationalisation's of AI capability — the maximum investor AI score rather than the ownership-weighted mean, log technology spending, and a count of ESG data platform subscriptions — yield consistent results. Three alternative outcome measures — Sustainalytics third-party ESG ratings, the 18-item environmental disclosure sub-index, and a GRI alignment score — corroborate the main finding. Sample restrictions excluding Kenya (the most advanced regulatory environment), restricting to large banks only, and excluding the COVID disruption years (2020–21) leave results unchanged. The System GMM specification addresses outcome persistence. The critical placebo test — AI capability predicting non-ESG financial disclosure detail — yields $\beta = 0.08$ (p = 0.52), confirming the effect is ESG-specific and cannot be attributed to a general transparency disposition among AI-capable investor bases.

Table 7: Robustness Summary — 15 Specifications

Specification	Category	β (AI)	p	Comment
Maximum AI	Alt. AI	0.54***	< .001	Largest single investor AI score

Specification	Category	β (AI)	p	Comment
score (not wt.)	measure			
Log technology spending	Alt. AI measure	0.47***	< .001	Objective proxy for AI investment
ESG platform subscription count	Alt. AI measure	0.43**	.003	Count of ESG data platform contracts
Third-party ESG ratings	Alt. outcome	0.39**	.002	Sustainalytics composite score
Environmental disclosure only	Alt. outcome	0.61***	< .001	18-item environmental sub-index
GRI alignment score	Alt. outcome	0.44***	< .001	Validated against GRI standards
Exclude Kenya	Sample	0.48***	< .001	Controls for most-advanced regulatory env.
Large banks only (> \$2B assets)	Sample	0.57***	< .001	N = 11 banks, 77 firm-years
Exclude 2020–21 (COVID years)	Sample	0.51***	< .001	Removes pandemic disruption
System GMM	Method	0.44**	.008	Addresses granularity persistence
Entropy balanced	Method	0.49***	< .001	CBPS reweighting on observables
PSM (ATT)	Method	14.7 pts	< .001	Avg. treatment effect on treated
AI capability +20% error	Sensitivity	0.55***	< .001	Upper bound measurement perturbation
AI capability -20% error	Sensitivity	0.48***	< .001	Lower bound measurement perturbation
Non-ESG disclosure (PLACEBO)	Placebo	0.08	.52	Financial reporting detail — null confirms ESG-specificity

Note: *** $p < 0.001$, ** $p < 0.01$. All specifications include country and year fixed effects and the full control set from Column (4) of Table 4. Cluster-robust standard errors at bank level ($N = 31$ clusters). PSM ATT reported in granularity index points rather than standardised β . The placebo specification (last row) is the most important diagnostic: the null result rules out the alternative explanation that AI-capable investors simply hold more transparent firms across all dimensions.

5. Discussion

This paper establishes a clean, causally identified fact: investor AI monitoring capability drives ESG disclosure granularity. The evidence is consistent across OLS, IV, staggered DiD, event studies, and entropy balancing, with effect sizes ranging from $\beta = 0.49$ to $\beta = 0.68$ and no single strategy capable of explaining away the result. The theoretical contribution is to formalise precisely why this occurs: in a signaling equilibrium, AI raises the detection cost of strategic ambiguity, compelling firms to increase granularity to maintain separation from low-type mimics. This extends prior disclosure theory (Verrecchia 1983; Dye 1985, 1998) by demonstrating that the precision-flexibility trade-off is not fixed but depends on investor verification technology — a dynamic that will intensify as AI adoption spreads beyond institutional investors into supervisory and regulatory processes.

The placebo result — AI capability predicts ESG granularity but not non-ESG financial disclosure detail ($\beta = 0.08$, $p = 0.52$) — is critical for interpretation. It rules out the hypothesis that AI-capable investors simply hold more transparent firms across all dimensions, confounding the granularity estimate. The specificity to ESG disclosure is exactly what the signaling model predicts: AI creates asymmetric pressure on the dimensions where misrepresentation risk is highest and alternative verification data are richest, namely sustainability metrics that can be cross-referenced against satellite imagery, regulatory databases, and third-party certification records.

The disclosure-performance paradox — actual ESG performance does not predict granularity, and AI effects are strongest for poor performers — is the paper's most consequential finding. It reveals that AI monitoring, despite its technical sophistication, produces an accountability gap rather than closing one. Firms respond to algorithmic scrutiny by optimising what is measurable and machine-readable, not by improving what is material. The CFO interview evidence is unusually candid on this point: sustainability report restructuring for machine-readability was explicitly distinguished from operational change. The label 'algorithmic greenwashing' is introduced to denote this pattern — AI-optimised disclosure calibrated for detection avoidance rather than genuine transparency.

This finding engages directly with the broader greenwashing literature (Marquis et al. 2016; Lyon & Maxwell 2011), which has documented analogous decoupling between symbolic and substantive responses to institutional pressure. The AI context adds a new dimension: the firm's strategic counterparty is not a human analyst who can be managed through narrative framing, but an algorithm that rewards structural verifiability. Firms adapt accordingly, producing what might be termed 'verifiably empty' disclosures: granular, tagged, externally structured reports that satisfy AI detection criteria without substantive operational change. This is a qualitatively different and more sophisticated form of symbolic compliance than prior literature has documented, because it requires genuine technical investment in disclosure infrastructure — the appearance of accountability is costly to produce.

The policy implication is direct and urgent. Regulators and standard-setters who view increased disclosure granularity as evidence of improved ESG governance should re-examine this

assumption in AI-intensive environments. Granularity is a necessary but not sufficient condition for accountability. The disclosure-performance decoupling implies that mandating more granular reporting without simultaneously mandating performance verification will simply shift the strategic equilibrium: firms will produce more granular disclosures with unchanged or deteriorating underlying performance. Effective oversight requires direct performance verification using alternative data — satellite imagery, energy consumption records, transaction-level green lending evidence — rather than treating algorithmically optimised disclosure as a proxy for genuine sustainability progress.

The individual variable analysis (Table 3) provides actionable guidance for disclosure standard design. Green lending ratio and financial inclusion score are significant positive predictors of granularity — the dimensions most amenable to third-party algorithmic verification. Carbon intensity, despite being the most objective environmental performance measure, shows no association with granularity. This pattern suggests that the ESG dimensions most resistant to strategic manipulation are precisely those least reflected in disclosure granularity — a structural weakness of current reporting frameworks that AI monitoring intensifies rather than corrects.

6. Conclusion

This paper addresses one precisely bounded question — does investor AI monitoring capability cause ESG disclosure granularity? — and answers it robustly in the affirmative. Using 31 East African commercial banks, 980 firm-year observations, and five identification strategies, the study finds a robust positive causal effect ($\beta = 0.52$ OLS; $\beta = 0.49$ – 0.68 across methods) that survives 15 robustness checks and is ESG-specific (non-ESG disclosure placebo $\beta = 0.08$, $p = 0.52$). The answer is confirmed. The interpretation is complicated.

The disclosure-performance paradox — granularity is decoupled from actual ESG performance, and poor performers respond most strongly to AI monitoring — is the paper's most important contribution. It establishes that AI monitoring produces algorithmic greenwashing rather than genuine accountability in the current institutional environment: firms optimise disclosure form for machine-readability without improving substantive ESG practices. The concept of algorithmic greenwashing — AI-optimised disclosure calibrated for detection avoidance — is introduced as a theoretically grounded label for this phenomenon and offered to the broader literature on symbolic institutional compliance.

Three implications follow directly. For regulators: mandatory granularity standards without independent performance verification will accelerate algorithmic greenwashing. For investors: AI-enabled ESG monitoring creates a strategic arms race in which disclosure form improves while content may deteriorate; alternative data verification is essential. For researchers: the AI-disclosure nexus in frontier markets is a productive empirical laboratory — heterogeneous AI adoption, staggered regulatory timelines, and quantifiable ESG metrics create identifying variation that developed markets cannot easily replicate.

The central methodological lesson is that measurement and performance are not the same thing, and AI can widen rather than close this gap. Future research should use alternative data as ground truth to test whether AI-induced disclosure improvements correlate with actual ESG outcomes, and should examine whether the algorithmic greenwashing pattern identified here generalises to developed markets where investor AI adoption is more advanced, institutional enforcement is stronger, and the stakes of disclosure misrepresentation are correspondingly higher.

References

- Allen, F., Demirguc-Kunt, A., Klapper, L., & Martinez Peria, M. S. (2014). The foundations of financial inclusion. *Journal of Financial Intermediation*, 27, 1–30. <https://doi.org/10.1016/j.jfi.2015.12.003>
- Bingler, J. A., Kraus, M., Leippold, M., & Webersinke, N. (2022). Cheap talk and cherry-picking: What ClimateBERT has to say on corporate climate risk disclosures. *Finance Research Letters*, 47, 102769. <https://doi.org/10.1016/j.frl.2022.102769>
- Bloomfield, R. J. (2002). The incomplete revelation hypothesis and financial reporting. *Accounting Horizons*, 16(3), 233–243. <https://doi.org/10.2308/acch.2002.16.3.233>

- Chen, H., Cohen, L., & Gurun, U. G. (2022). Don't talk yourself into it: The role of debt in accounting quality. *Management Science*, 68(10), 7569–7587. <https://doi.org/10.1287/mnsc.2022.4373>
- Dhaliwal, D. S., Li, O. Z., Tsang, A., & Yang, Y. G. (2011). Voluntary nonfinancial disclosure and the cost of equity capital. *The Accounting Review*, 86(1), 59–100. <https://doi.org/10.2308/accr.00000005>
- DiMaggio, P. J., & Powell, W. W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *American Sociological Review*, 48(2), 147–160. <https://doi.org/10.2307/2095101>
- Dye, R. A. (1985). Disclosure of nonproprietary information. *Journal of Accounting Research*, 23(1), 123–145. <https://doi.org/10.2307/2490910>
- Dye, R. A. (1998). Investor sophistication and voluntary disclosures. *Review of Accounting Studies*, 3(3), 261–287. <https://doi.org/10.1023/A:1009685509194>
- Grossman, S. J. (1981). The informational role of warranties and private disclosure about product quality. *Journal of Law and Economics*, 24(3), 461–483. <https://doi.org/10.1086/466995>
- Hirshleifer, D., & Teoh, S. H. (2003). Limited attention, information disclosure, and financial reporting. *Journal of Accounting and Economics*, 36(1–3), 337–386. <https://doi.org/10.1016/j.jacceco.2003.10.002>
- Hummel, K., & Schlick, C. (2016). The relationship between sustainability performance and sustainability disclosure. *Journal of Accounting and Public Policy*, 35(5), 455–476. <https://doi.org/10.1016/j.jaccpubpol.2016.06.002>
- Jensen, M. C., & Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics*, 3(4), 305–360. [https://doi.org/10.1016/0304-405X\(76\)90026-X](https://doi.org/10.1016/0304-405X(76)90026-X)
- Jung, W. O., & Kwon, Y. K. (1988). Disclosure when the market is unsure of information endowment of managers. *Journal of Accounting Research*, 26(1), 146–153. <https://doi.org/10.2307/2491175>
- KPMG. (2024). *KPMG survey of sustainability reporting 2024*. KPMG International.
- Li, F. (2008). Annual report readability, current earnings, and earnings persistence. *Journal of Accounting and Economics*, 45(2–3), 221–247. <https://doi.org/10.1016/j.jacceco.2008.02.003>
- Li, Y., Lu, M., & Yu, F. (2023). Artificial intelligence and ESG investing. *Journal of Financial and Quantitative Analysis*. Advance online publication. <https://doi.org/10.1017/S0022109023001060>
- Loughran, T., & McDonald, B. (2020). Textual analysis in finance. *Annual Review of Financial Economics*, 12, 357–375. <https://doi.org/10.1146/annurev-financial-012820-032249>
- Lyon, T. P., & Maxwell, J. W. (2011). Greenwash: Corporate environmental disclosure under threat of audit. *Journal of Economics & Management Strategy*, 20(1), 3–41. <https://doi.org/10.1111/j.1530-9134.2010.00282.x>

- Marquis, C., Toffel, M. W., & Zhou, Y. (2016). Scrutiny, norms, and selective disclosure: A global study of greenwashing. *Organization Science*, 27(2), 483–504. <https://doi.org/10.1287/orsc.2015.1039>
- Milgrom, P. R. (1981). Good news and bad news: Representation theorems and applications. *The Bell Journal of Economics*, 12(2), 380–391. <https://doi.org/10.2307/3003565>
- Miller, B. P. (2010). The effects of reporting complexity on small and large investor trading. *The Accounting Review*, 85(6), 2107–2143. <https://doi.org/10.2308/accr.00000001>
- Ndung'u, N., & Wachira, M. (2019). Sustainability reporting in Kenya: Challenges and opportunities. *African Journal of Business Ethics*, 13(2), 1–22.
- Osei-Tutu, F., & Weill, L. (2022). Bank ownership and financial inclusion in Africa. *Journal of International Financial Markets, Institutions and Money*, 78, 101563. <https://doi.org/10.1016/j.intfin.2022.101563>
- Pedersen, L. H., Fitzgibbons, S., & Pomorski, L. (2021). Responsible investing: The ESG-efficient frontier. *Journal of Financial Economics*, 142(2), 572–597. <https://doi.org/10.1016/j.jfineco.2020.11.001>
- Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics*, 87(3), 355–374. <https://doi.org/10.2307/1882010>
- Verrecchia, R. E. (1983). Discretionary disclosure. *Journal of Accounting and Economics*, 5, 179–194. [https://doi.org/10.1016/0165-4101\(83\)90011-3](https://doi.org/10.1016/0165-4101(83)90011-3)