

DOI: <https://doi.org/10.38035/dijefa><https://creativecommons.org/licenses/by/4.0/>

Received: 13 September 2026 | Revised: 15 September 2026 | Accepted: 27 September 2026

The Effect of ESG Score on Financial Performance and Financial Distress with ESG Controversies as a Moderating Variable in ASEAN and Indian Emerging Markets

Stefanus Dimas R. Saputro^{1*}, Enrique Griffin Imanuel Marbun², Elfriana Nely Royana Situmorang³

¹Universitas Pelita Harapan, Tangerang, Indonesia, stefanus.saputro@uph.edu

²Universitas Pelita Harapan, Tangerang, Indonesia, enrikoiman9@gmail.com

³Universitas Pelita Harapan, Tangerang, Indonesia, elfriananely@gmail.com

*Corresponding Author: stefanus.saputro@uph.edu¹

Abstract: This study examines the association between ESG Score and two financial outcomes, namely financial performance and financial health, and the moderating role of ESG Controversies in both relationships. Prior studies report mixed findings and rarely examine the two outcomes within a single framework in Asian emerging markets. The sample consists of 1,168 firm-year observations from 146 non-financial firms in Indonesia, Malaysia, the Philippines, Thailand, and India over the period 2017–2024, obtained from LSEG Workspace. Financial performance is proxied by return on equity (ROE) and financial health by the Altman Z"-Score, for which a higher value indicates lower distress risk. The data are analysed using panel regression with firm fixed effects, year fixed effects, and standard errors clustered at the firm level, and all inference is two-tailed. ESG Score shows no significant association with either ROE ($\beta = 0.00051$; $p = 0.984$) or the Z"-Score ($\beta = -0.01100$; $p = 0.457$), so the first and second hypotheses are not supported. The ESG $\times$ Controversies interaction is negative and significant for ROE ($\beta = -0.001316$; $p = 0.009$) and positive and significant for the Z"-Score ($\beta = 0.000899$; $p < 0.001$), supporting the third and fourth hypotheses. Robustness tests that hold the estimator fixed and vary one element at a time show that the two results are not equally secure: the positive interaction on the Z"-Score survives the exclusion of ROA, country $\times$ year fixed effects, a one-year lag, random effects, and the restoration of firms removed by an extreme-value screen, whereas the negative interaction on ROE keeps its direction but loses significance without ROA and reverses when those firms are restored. Marginal-effects analysis locates the sign change at a controversy level of 8.5 for ROE and 20.3 for the Z"-Score. The contribution of this study lies in the evidence that a single moderating construct is associated with opposite interaction signs across two financial outcomes, which suggests that ESG Score and controversy information should be assessed jointly rather than separately. Because the design is observational, the findings should be interpreted as conditional associations rather than as causal effects.

Keywords: ESG Score, ESG Controversies, Financial Performance, Financial Health, Altman Z"-Score.

INTRODUCTION

Evidence on whether environmental, social, and governance (ESG) performance is associated with better financial outcomes remains unsettled, and three unresolved issues motivate this study. First, findings on ESG and profitability are inconsistent across markets and estimation designs. Second, ESG ratings reward the completeness of policies and disclosure, yet a firm may hold a high score while its actual conduct produces negative incidents, so the credibility of the ESG signal depends on the controversy record that accompanies it. Third, and least examined, ESG may operate differently on the profitability channel, which shareholders and product markets evaluate, and on the solvency channel, which creditors evaluate; because most studies model a single outcome, this difference has rarely been tested within one framework. This study therefore distinguishes two complementary financial outcomes. Financial performance reflects the ability of a firm to convert shareholders' capital into profit and is proxied by return on equity (ROE). Financial health reflects distance from default and is proxied by the Altman Z"-Score for non-manufacturing firms in emerging markets; because a higher Z"-Score indicates a healthier financial condition, the score moves inversely to distress risk, and every interpretation in this study follows that direction explicitly. The setting is one in which these questions carry practical weight: firms in Asian emerging markets face interest-rate and exchange-rate volatility, supply-chain disruption, inflationary pressure, and an accelerating sustainability transition, so financial resilience has become a strategic rather than a purely accounting concern.

ESG performance has developed into an important framework for investors, creditors, regulators, and management in assessing the quality of corporate sustainability management. Conceptually, ESG is rooted in the triple bottom line idea popularised by Elkington (1997), which asserts that corporate sustainability is judged not only by profit but also by people and planet. Stakeholder theory (Freeman, 1984) explains that firms which manage the interests of employees, customers, suppliers, creditors, regulators, and communities effectively obtain the support and resources that improve performance and resilience. Signalling theory (Spence, 1973; Connelly et al., 2011) positions ESG Score as a signal of the quality of non-financial risk management to parties with less information, while legitimacy theory (Suchman, 1995; Deegan, 2019) explains the effort of firms to ensure that their activities are regarded as appropriate within prevailing social norms.

The relationship between ESG and financial outcomes is nevertheless not automatic. Environmental and social investment may generate long-term benefits but can also raise costs and depress earnings in earlier periods. A firm may also obtain a high ESG Score because of the completeness of its policies and disclosures, even when its actual conduct produces negative incidents. This gap between claim and conduct is captured by ESG controversies, which are event-based negative information drawn from news and public sources. Elamer and Boulhaga (2024) show that ESG controversies are negatively associated with corporate performance across 5,360 firm-year observations and that governance quality and ESG practices alter the magnitude of that impact, while Nirino et al. (2021) find that corporate controversies reduce firm value with ESG practices acting as a moderator.

Empirical evidence on ESG and profitability remains inconsistent. Alareeni and Hamdan (2020) report findings that support the benefit of aggregate ESG disclosure, whereas Rao et al. (2023) document heterogeneous pillar effects along the distribution of ROE in India, and Narula et al. (2024) show that ESG effects on firm performance differ across emerging market contexts. Evidence on financial distress is comparatively more consistent in positioning ESG

as a risk-mitigation mechanism: Boubaker et al. (2020) find that stronger corporate social responsibility profiles are associated with lower distress risk, Shi et al. (2023) find that the social and governance pillars reduce distress risk in the air transport industry, and Zhou et al. (2025) and Song et al. (2024) show that ESG information improves distress prediction. Results by pillar, sector, and region nevertheless remain uneven, and Rahmadini and Hartanti (2025) find no significant moderating role of controversies in the Indonesian setting.

This inconsistency points to several gaps. First, a number of studies use the term financial performance interchangeably with firm value, so that Tobin's Q is treated as a performance construct even though it captures market valuation rather than the return generated on equity. Second, most studies examine a single financial outcome, so the possibility that ESG operates differently on the profitability channel and on the solvency channel is rarely tested within one framework. Third, ESG controversies are more often modelled as a direct explanatory variable than as a moderator, even though their theoretical role is to condition the credibility and salience of the ESG signal. Fourth, the raw Refinitiv ESG Controversies Score is inversely scaled, since a higher raw score indicates fewer controversies, which creates a persistent risk of sign misinterpretation when the score is used without transformation. Fifth, regional evidence is fragmented: Southeast Asian studies frequently rely on short windows or on return on assets, while Indian evidence is usually produced separately.

This study addresses those gaps by examining non-financial firms in four ASEAN emerging markets, namely Indonesia, Malaysia, the Philippines, and Thailand, together with India, over the period 2017–2024. The five markets differ in market size, ownership structure, investor protection, disclosure quality, and institutional enforcement, yet share the characteristics of emerging markets that make sustainability information economically relevant. The eight-year window covers the pre-pandemic phase, the COVID-19 shock, and the recovery period, so that variation in ESG, controversies, profitability, and distress risk can be observed across distinct economic conditions.

The contribution of this study is stated relative to the studies closest to it rather than as an absolute novelty claim. Table 1 maps the nearest prior work along four dimensions: setting, financial outcome, the role assigned to ESG controversies, and the resulting difference from the present design. Three patterns emerge. First, studies that model controversies almost always place them as a direct explanatory variable, or reverse the moderation by letting ESG practices moderate the effect of controversies (Nirino et al., 2021; Elamer & Boulhaga, 2024); the specification used here, in which controversies condition the association of the ESG Score, appears in a small number of studies, all confined to Indonesia and to firm value or performance (Rahmadini & Hartanti, 2025; Sihotang & Siregar, 2025). Second, studies of financial distress and ESG rarely model controversies at all (Boubaker et al., 2020; Landi et al., 2022; Shi et al., 2023; Song et al., 2024; Lohmann et al., 2025; Zhou et al., 2025), and where distress appears jointly with controversies it is typically as an antecedent of firm value rather than as an outcome (Yadav & Asongu, 2025). Third, regional coverage is fragmented: Southeast Asian evidence is usually produced without India (Le, 2024; Darsono et al., 2025), Indian evidence usually without ASEAN (Rao et al., 2023; Yadav & Asongu, 2025), and syntheses of the field note this fragmentation as an open issue (Zheng et al., 2025); evidence drawn from banking and other financial-sector samples (Buallay, 2019; Chiaramonte et al., 2022) is not directly comparable, since those firms are excluded here.

Against this mapping, the present study differs in a specific and demonstrable way rather than in an absolute one: to the best of our knowledge, and subject to the search procedure described next, it is the first to test the same controversy moderator on a profitability outcome and a solvency outcome within a single panel, and thereby the first to document that the estimated interaction sign differs between the two. The claim is bounded by that procedure: Scopus, Web of Science, and Google Scholar were searched to [month, year] using

combinations of “ESG”, “ESG controversies”, “financial performance”, “return on equity”, “financial distress”, “Altman Z-score”, “emerging market”, “ASEAN”, and “India”, with backward and forward citation tracing on the studies in Table 1. Studies published after that date, in languages other than English and Indonesian, or outside indexed outlets, may not be captured.

Table 1. Positioning Relative to the Closest Prior Studies

Study	Setting	Financial outcome(s)	Role of ESG controversies	Difference from this study
Alareeni & Hamdan (2020)	US, S&P 500	ROA, ROE, Tobin's Q	Not modelled	Developed market; no controversy moderator; no distress outcome
Buallay (2019)	Europe, banking	ROA, ROE, Tobin's Q	Not modelled	Financial sector, excluded here; no controversies
Rao et al. (2023)	India	Financial performance (quantile)	Not modelled	Single country; performance only; no controversies
Narula et al. (2024)	Emerging markets	Firm performance	Not modelled	No controversy moderation; no distress outcome
Le (2024)	Southeast Asia	Financial performance	Not modelled	ASEAN without India; performance only; no moderator
Darsono et al. (2025)	Indonesia	Financial performance, cost of debt	Not modelled (assurance is the moderator)	Single country; different moderator; no distress outcome
Nirino et al. (2021)	International	Firm value	Direct explanatory variable; ESG practices moderate	Moderation runs in the opposite direction; no distress outcome
Elamer & Boulhaga (2024)	International	Corporate performance	Direct explanatory variable; governance and ESG practices moderate	Moderation runs in the opposite direction; no distress outcome
Rahmadini & Hartanti (2025)	Indonesia	Financial performance	Moderator of the ESG–performance link	Single country; performance only; smaller sample
Sihotang & Siregar (2025)	Indonesia	Firm value, financial distress	Moderator	Firm value rather than ROE; single country
Boubaker et al. (2020)	International	Financial distress (Z-score)	Not modelled	CSR rather than ESG Score; no moderation; no profitability outcome
Landi et al. (2022)	Europe	Corporate financial risk	Not modelled	Risk rather than Z”; developed market; no controversies
Shi et al. (2023)	Global, air transport	Financial distress	Not modelled (region is the moderator)	Single industry; no controversies
Song et al. (2024)	China, energy	Distress prediction	Not modelled	Prediction rather than association; single sector

Lohmann et al. (2025)	International	Financial distress	Not modelled	No controversy moderation; no profitability outcome
Zhou et al. (2025)	China	Distress models	Not modelled	Single country; no controversies
Yadav & Asongu (2025)	India	Company value (distress as antecedent)	ESG performance is the moderator	Reversed ordering; value rather than ROE
Broadstock et al. (2021)	China	Crisis-period returns	Not modelled	Event window; market returns
Chiaramonte et al. (2022)	Europe, banking	Bank stability	Not modelled	Financial sector, excluded here
Zheng et al. (2025)	Meta-analysis	Financial stability	Not applicable	Synthesis rather than a primary test
This study	Indonesia, Malaysia, Philippines, Thailand, India	ROE and Altman Z"-Score jointly	Moderator of the ESG–outcome link on both outcomes	—

Source: Compiled by the authors.

Four theoretical lenses are used in this study, but they are not treated as interchangeable justifications; each addresses a different link in the argument and is assigned a specific role. The triple bottom line (Elkington, 1997) supplies the scope condition rather than a testable mechanism: it establishes that corporate performance is evaluated on environmental and social dimensions alongside profit, which is why an ESG Score is economically meaningful information at all, and it generates no hypothesis on its own. Stakeholder theory (Freeman, 1984; Donaldson & Preston, 1995) supplies the resource mechanism and generates H1: firms that manage the claims of employees, customers, suppliers, creditors, regulators, and communities effectively secure cooperation, lower conflict and transaction costs, and preferential access to resources, which should raise the return generated on equity. Jensen (2001) qualifies this by noting that stakeholder investment improves firm outcomes only where it is disciplined by long-run value creation, which is one reason the predicted association is expected to be modest and possibly lagged. Signalling theory (Spence, 1973; Ross, 1977; Connelly et al., 2011) supplies the information mechanism and generates H2, H3, and H4 through two of its properties: signal credibility, which contradictory evidence can destroy, so that verifiable event-based controversies act as a counter-signal that discounts policy-based disclosure; and signal salience, since a signal is weighted only when the receiver is attending to the risk it describes. Legitimacy theory (Suchman, 1995; Deegan, 2002, 2019) supplies the buffering mechanism and reinforces H4: firms accumulate legitimacy reserves through prior environmental policies, safety procedures, grievance channels, compliance functions, and board oversight, and when an incident opens a legitimacy gap those reserves are what a firm draws on to contain escalation. Stakeholder salience theory (Mitchell et al., 1997) connects the two: the stakeholder group whose claims become urgent at a given moment is the group whose evaluation drives the observable financial outcome.

The pivot of the argument is that ROE and the Z"-Score are evaluated by different stakeholder groups with different risk orientations, and that controversies do different things to the ESG signal for each group. Proposition 1 (credibility discount, profitability channel): ROE outcomes are shaped by shareholders and product markets, which are oriented toward upside and evaluate the firm continuously. For these receivers the ESG Score competes with, and is dominated by, event-based negative information, because negative events are more diagnostic of conduct than policy disclosure is. As the controversy level rises, an increment in ESG Score is therefore less readily interpreted as an improvement in economic quality; the moderation is

subtractive, and the interaction term on ROE is predicted to be negative. Proposition 2 (signal salience and legitimacy repair, solvency channel): Z"-Score outcomes are shaped by creditors, which are oriented toward downside and evaluate the firm discretely, when default risk becomes salient. Where no controversy is present, non-financial risk is not salient and the ESG Score is treated as routine disclosure that changes little in a solvency assessment. Where a controversy materialises, information that separates firms able to contain the incident from those that are not becomes necessary, and the ESG Score becomes diagnostic; at the same time, the legitimacy reserves it proxies are what permit containment, an argument analogous to Purnanandam (2008), who shows that the value of risk-management capacity is concentrated in states where distress is near. The moderation is activating, and the interaction term on the Z"-Score is predicted to be positive.

The two propositions are not contradictory. They describe one construct evaluated by two audiences under two attention regimes: for continuously attentive, upside-oriented receivers, a counter-signal subtracts credibility; for intermittently attentive, downside-oriented receivers, an adverse event is what makes the signal worth reading at all.

Because the Altman Z"-Score is an inverse measure of distress, the hypotheses are stated directly in terms of the variable that is estimated, so that no translation between "distress" and "financial health" is required when reading the coefficients. H1: ESG Score has a positive association with return on equity (stakeholder theory; predicted sign on β_1 in Model 1 positive). H2: ESG Score has a positive association with the Altman Z"-Score, that is, it is associated with lower financial distress risk (signalling theory; predicted sign on β_1 in Model 3 positive). H3: ESG Controversies weaken the association between ESG Score and return on equity, that is, the $ESG \times CONT$ interaction has a negative association with ROE (Proposition 1; predicted sign on β_3 in Model 2 negative). H4: ESG Controversies strengthen the association between ESG Score and the Altman Z"-Score, that is, the $ESG \times CONT$ interaction has a positive association with the Z"-Score (Proposition 2; predicted sign on β_3 in Model 4 positive). The opposite predicted signs of H3 and H4 are deliberate and follow directly from Propositions 1 and 2 rather than from the sign convention of the dependent variables: CONT is transformed so that higher values always denote greater controversy intensity, and the Z"-Score is stated so that higher values always denote better financial health.

Accordingly, the research questions of this study are formulated as follows: 1. Is ESG Score associated with financial performance? 2. Is ESG Score associated with financial health? 3. Do ESG Controversies moderate the association between ESG Score and financial performance? 4. Do ESG Controversies moderate the association between ESG Score and financial health? These questions are answered in the results and discussion section and summarised in the conclusion.

METHODS

This study applies a quantitative approach with an explanatory research design using secondary panel data. The target population comprises all non-financial firms listed on the Indonesia Stock Exchange, Bursa Malaysia, the Philippine Stock Exchange, the Stock Exchange of Thailand, and the National Stock Exchange and/or Bombay Stock Exchange of India during 2017–2024. Financial sector firms, including banks, insurers, financing companies, securities firms, and other financial institutions, are excluded because their asset structure, leverage, and capital requirements differ fundamentally, so that ROE and the Altman Z"-Score are not comparable with those of non-financial firms.

The sample was selected using purposive sampling (Sugiyono, 2019) with the following criteria: (a) the firm is listed on one of the designated exchanges during 2017–2024 and is classified as a non-financial firm; (b) the firm-year has both an LSEG ESG Score and an LSEG ESG Controversies Score; (c) the firm-year has the net income and average equity required to

compute ROE, together with the working capital, retained earnings, EBIT, total assets, book value of equity, and total liabilities required to compute the Altman Z'' -Score; (d) leverage, firm size, return on assets, country, and year are available; and (e) average equity is positive and non-zero. Table 2 reports the number of firms and firm-years removed at each stage, so that the path from the initial universe to the estimation sample is fully traceable.

Table 2. Sample Selection Procedure

Step	Criterion	Firms removed	Firms remaining	Firm-years
1	Non-financial firms listed on IDX, Bursa Malaysia, PSE, SET, and NSE/BSE with LSEG coverage, 2017–2024 (initial universe). Financial-sector firms are absent by construction: the extract was restricted to non-financial GICS sectors at the data-pull stage.	—	5,390	—
2	Less: firms without an LSEG ESG Score for the observation window	5,186	204	—
3	Less: firms without an LSEG ESG Controversies Score (ESG Score and Controversies Score coverage coincide exactly in this universe)	0	204	—
4	Less: firms missing any Altman Z'' component (X_1 – X_4)	10	194	—
5	Less: firms missing total assets, leverage, or return on assets	12	182	—
6	Less: firms missing return on equity	20	162	1,296
7	Less: firms failing the extreme-value screen, namely $ ROE > 100\%$, $Z'' > 40$, or leverage > 1 (negative book equity) in any year	16	146	1,168
8	Balanced-panel requirement: every remaining firm is observed in all eight years (146×8)	0	146	1,168
9	Winsorisation of all continuous variables at the 1st and 99th percentiles (no observations removed)	0	146	1,168

Source: Compiled by the authors.

ESG Score, ESG Controversies Score, and firm financial data were obtained from LSEG Workspace/Refinitiv, while annual real gross domestic product growth was obtained from the World Bank World Development Indicators. Data were collected through electronic documentation of secondary data and were matched using permanent identifiers such as ISIN or the Refinitiv Instrument Code. The operational definition of each variable is presented in Table 3.

Table 3. Operational Definition of Variables

Variable (Symbol)	Measurement	Source	Interpretation
Financial Performance (ROE)	Return on average common equity (%), i.e. net income after tax divided by average total equity $\times 100\%$.	LSEG Workspace	A higher ROE indicates stronger equity-based profitability.
Financial Distress (Z'')	$Z'' = 6.56X_1 + 3.26X_2 + 6.72X_3 + 1.05X_4$, where X_1 = working capital/total assets, X_2 = retained earnings/total assets, X_3 = EBIT/total assets, X_4 = book value of equity/total liabilities.	LSEG Workspace	A higher Z'' indicates healthier financial condition and lower distress risk.
ESG Score (ESG)	Overall ESG Score on a 0–100 scale.	LSEG Workspace	A higher score indicates better ESG performance.
ESG Controversies (CONT)	$CONT = 100 - \text{raw Refinitiv ESG Controversies Score}$.	LSEG Workspace	A higher CONT indicates a greater level of controversy.
Firm Size (SIZE)	Natural logarithm of total assets.	LSEG Workspace	Firm scale.
Leverage (LEV)	Total liabilities / total assets.	LSEG Workspace	Level of firm obligations.
Return on Assets (ROA)	Net income / average total assets $\times 100\%$.	LSEG Workspace	Asset-based profitability.
GDP Growth (GDPG)	Annual real GDP growth of the country.	World Bank (WDI)	National macroeconomic condition.

Source: Compiled by the authors from Altman (1968, 2000), Refinitiv ESG/LSEG Workspace methodology, and World Development Indicators.

Measurement of ESG Controversies

Because ESG Controversies is the central moderating construct, its underlying definition and its transformation are documented in full. The LSEG (formerly Refinitiv) ESG Controversies Score is built from 23 ESG controversy topics covering, among others, anti-competitive behaviour, business ethics, human rights, environmental damage, workforce practices, and product responsibility. It is expressed on a 0–100 percentile scale and is computed as a percentile rank within the firm’s TRBC industry group, so it measures controversy exposure relative to industry peers rather than an absolute count of incidents. To prevent large firms from being penalised for greater media coverage, incident severity is weighted by market-capitalisation class (0.33 for large cap of at least USD 10 billion, 0.67 for mid cap of at least USD 2 billion, and 1.00 for small cap below USD 2 billion). Critically for this study, companies with no recorded controversies receive a raw score of 100 (LSEG, 2024).

The raw score is therefore inversely scaled with respect to the construct of interest: a raw score of 100 denotes the absence of controversy and lower raw scores denote greater controversy exposure. Used untransformed, a positive coefficient would have to be read as “fewer controversies”, which is a recurrent source of sign misinterpretation in this literature. The variable is therefore defined as $CONT = 100 - \text{raw ESG Controversies Score}$. This is a strictly monotone, order-preserving linear reflection of the raw percentile: it changes no information content and no rank ordering, and only reverses the direction so that a higher CONT denotes greater controversy intensity and the sign of every coefficient and interaction can be read directly in the direction in which the hypotheses are stated.

Because a controversy-free firm-year receives a raw score of 100, $CONT = 0$ means “no recorded controversy in that firm-year” rather than “the maximum of the raw scale” in any substantive sense; the two coincide by construction. This is consistent with the observed distribution: the median and the 75th percentile of CONT are both zero, so more than three-quarters of firm-years record no controversy, while the 90th percentile rises to approximately

43.02, and the 95th to 63.64. Two consequences follow and are carried through the interpretation of the results. First, CONT is a censored-at-zero, right-skewed variable rather than an approximately continuous one. Second, identification of the interaction rests on the minority of firm-years in which controversies are recorded, which is why marginal effects are reported at specified points of the CONT distribution rather than only at its mean.

The treatment of missing controversy observations does not arise as an issue in this sample. Coverage of the ESG Score and coverage of the ESG Controversies Score coincide exactly in the universe from which the sample is drawn: every one of the 204 firms with an ESG Score also has a Controversies Score, and no firm was lost at that step (Table 2, step 3). No observation was therefore excluded for a missing controversy score, and none was recoded. The distinction would otherwise be consequential, because an absent score would mean the firm-year is not covered by the controversy data collection, whereas a raw score of 100 is an affirmative statement that no controversy was recorded; recoding the former to $CONT = 0$ would conflate “no coverage” with “no controversy” and would mechanically inflate the controversy-free group. In the estimation sample 78.4 per cent of firm-years record $CONT = 0$, and every one of them is an affirmative zero.

Four models are estimated separately. Model 1 and Model 3 test the direct effects, while Model 2 and Model 4 add the moderating term. The subscript i denotes the firm, t denotes the year, μ_i represents firm fixed effects, λ_t represents year fixed effects, and ε_{it} is the error term. In the moderation models, ESG Score and CONT are mean-centred and the interaction term is formed as the product of the two centred variables in order to limit non-essential multicollinearity and to allow the coefficient of the constituent variables to be read at the sample average.

Model 1 – Financial Performance (H1)

$$ROE_{it} = \alpha + \beta_1 ESG_{it} + \beta_2 LEV_{it} + \beta_3 SIZE_{it} + \beta_4 ROA_{it} + \beta_5 GDPGC_t + \mu_i + \lambda_t + \varepsilon_{it}$$

Model 2 – Financial Performance with ESG Controversies as Moderator (H3)

$$ROE_{it} = \alpha + \beta_1 ESG_{it} + \beta_2 CONT_{it} + \beta_3 (ESG \times CONT)_{it} + \beta_4 LEV_{it} + \beta_5 SIZE_{it} + \beta_6 ROA_{it} + \beta_7 GDPGC_t + \mu_i + \lambda_t + \varepsilon_{it}$$

Model 3 – Financial Distress (H2)

$$Z''_{it} = \alpha + \beta_1 ESG_{it} + \beta_2 LEV_{it} + \beta_3 SIZE_{it} + \beta_4 ROA_{it} + \beta_5 GDPGC_t + \mu_i + \lambda_t + \varepsilon_{it}$$

Model 4 – Financial Distress with ESG Controversies as Moderator (H4)

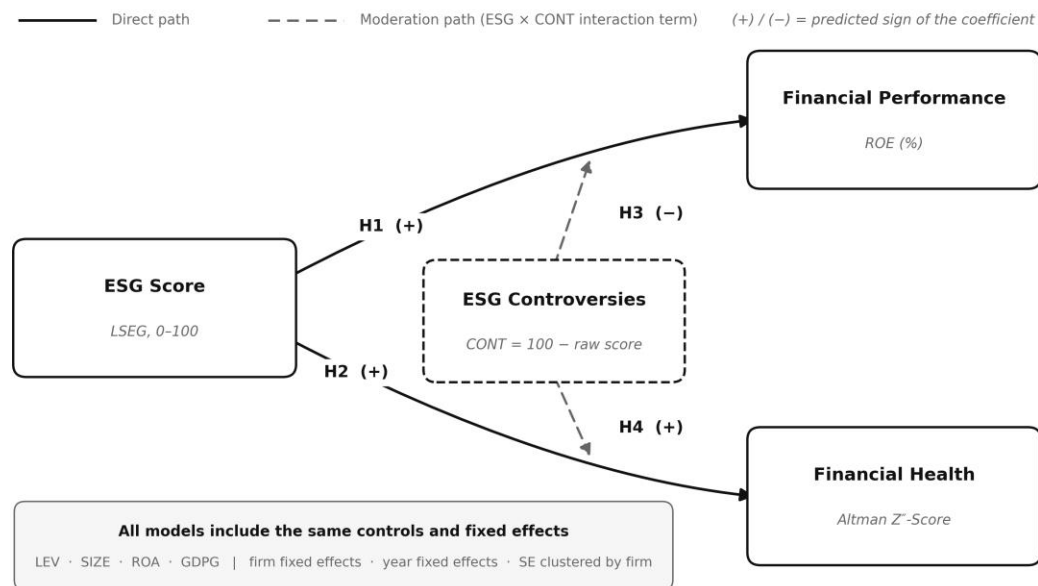
$$Z''_{it} = \alpha + \beta_1 ESG_{it} + \beta_2 CONT_{it} + \beta_3 (ESG \times CONT)_{it} + \beta_4 LEV_{it} + \beta_5 SIZE_{it} + \beta_6 ROA_{it} + \beta_7 GDPGC_t + \mu_i + \lambda_t + \varepsilon_{it}$$

Control Variables and Overlap with the Dependent Variables

The choice of ROA and LEV as controls requires justification, because both are mechanically related to the dependent variables. By the DuPont identity, $ROE = ROA \times (\text{total assets} / \text{average equity})$, so ROA is not an ordinary covariate of ROE but one of its two multiplicative components; conditioning ROE on ROA therefore changes the estimand, since β_1 ceases to be the association of ESG Score with equity profitability and becomes its association with the equity-multiplier component of profitability, holding asset profitability fixed. The sample correlation of 0.746 reflects this construction rather than an incidental collinearity. The overlap is more direct for the distress model: in $Z'' = 6.56X_1 + 3.26X_2 + 6.72X_3 + 1.05X_4$, the term $X_3 = \text{EBIT} / \text{total assets}$ is a pre-tax, pre-interest analogue of ROA and carries the largest weight in the index, while $X_4 = \text{book value of equity} / \text{total liabilities}$ is a direct function of the capital structure that LEV measures. Controlling for ROA and LEV in a Z'' regression therefore partials out a substantial share of the dependent variable's own construction, which is consistent with the observed correlations of 0.628 and -0.683 with Z'' .

Because the specification without ROA materially changes the moderation results, this is treated as a genuine specification choice rather than a robustness footnote. Three specifications are therefore reported side by side for both outcomes, and conclusions are drawn only from patterns that are stable across them: (1) the full specification (LEV, SIZE, ROA, GDPG), retained for comparability with the prior literature in which this control set is conventional; (2) without ROA, which removes the mechanical overlap with the ROE decomposition and with the X_3 term of Z'' , and which is treated as the preferred specification for the Z'' models on construction grounds; (3) without ROA and without LEV, which additionally removes the overlap with the X_4 term. A fourth specification replacing ROA and LEV with non-overlapping controls such as sales growth, asset tangibility, and cash holdings would isolate the estimand further, but those variables are not available in the assembled dataset, so that specification is not reported and the gap is stated as a limitation.

Data were analysed using Stata; every estimate reported in this revision was additionally reproduced independently, and the two sets of results agree to seven decimal places. The analysis begins with descriptive statistics, followed by Pearson correlation and the multicollinearity diagnostics described below, panel diagnostics covering residual normality, heteroskedasticity, and autocorrelation, and model specification testing through an F test for individual effects, a Hausman test, and a cluster-robust Mundlak auxiliary-regression test of the random-effects assumption. The main specification uses firm fixed effects and year fixed effects with standard errors clustered at the firm level. In the previous version of this study the two direct-effect models were estimated without clustering while the two moderation models were clustered; all four models are now estimated with firm-clustered standard errors so that inference is uniform across the table. Clustering leaves the two direct-effect coefficients unchanged and both remain insignificant. Country fixed effects are not included alongside firm fixed effects because firms do not change country, so time-invariant country characteristics are already absorbed. Model feasibility is evaluated through the within R-squared and the F test, without imposing a minimum R-squared threshold. The one-tailed testing procedure used in the previous version is withdrawn: deriving one-tailed p-values from two-tailed output whenever the estimated coefficient sign matched the prediction conditions the test on the realised sign and is not defensible as pre-specified directional testing. All inference reported here is therefore two-tailed at conventional thresholds (1%, 5%, 10%). The directional predictions in H1–H4 are retained but are used only for interpretation: a hypothesis is treated as supported when the coefficient is both statistically significant on a two-tailed test and signed in the predicted direction, and a coefficient that is significant with the opposite sign is reported as evidence against the hypothesis. For the moderation models, average marginal effects of ESG Score are computed at specified levels of CONT with 95% confidence intervals, and interaction plots with Johnson–Neyman regions of significance are reported for both outcomes. The sampling procedure follows the purposive-sampling logic set out in Sugiyono (2019). Robustness tests retain the firm fixed effects of the baseline and vary one element at a time: exclusion of ROA, exclusion of ROA and leverage, country $\times$ year fixed effects, a one-year lag, the ASEAN and India sub-samples, the 162-firm sample that precedes the extreme-value screen, random effects, and Driscoll–Kraay standard errors. A separate block replaces the firm fixed effects with country fixed effects in a pooled specification. The conceptual framework of the study is presented in Figure 1.


Figure 1. Conceptual Framework

Notes: Solid arrows are direct paths — ESG Score → ROE (H1, predicted positive) and ESG Score → Altman Z"-Score (H2, predicted positive, i.e. lower distress risk). Dashed arrows are moderation paths, in which ESG Controversies (CONT = 100 – raw LSEG score, so that a higher value denotes greater controversy intensity) conditions each direct path through the ESG × CONT interaction term: predicted negative on ROE (H3) and predicted positive on the Z"-Score (H4). The opposite predicted signs follow Propositions 1 and 2 and are not an artefact of scaling, since a higher CONT always denotes more controversy and a higher Z"-Score always denotes better financial health. All four models include LEV, SIZE, ROA, and GDPG as controls, firm and year fixed effects, and standard errors clustered at the firm level. Source: developed by the authors.

RESULTS AND DISCUSSION

Results

Sample Characteristics and Descriptive Statistics

The final dataset contains 1,168 firm-year observations and 146 firms observed from 2017 to 2024, forming a strongly balanced panel in which each firm contributes eight annual observations. The distribution of observations by year is constant at 146 firm-years per year. By country code, the panel comprises 424 observations for country_id 1, 200 for country_id 2, 216 for country_id 3, 136 for country_id 4, and 192 for country_id 5, with country_id 1 corresponding to India and codes 2 to 5 jointly representing the four ASEAN markets. Descriptive statistics computed on the original variable values before winsorisation are reported in Table 4.

Table 4. Descriptive Statistics of Research Variables

Variable	N	Mean	Median	Std. Dev.	Minimum	Maximum
ROE	1,168	13.732	12.320	11.649	-49.450	99.000
Z"-Score	1,168	6.386	3.830	6.940	-2.869	37.970
ESG Score	1,168	58.859	60.495	16.139	8.164	91.740
ESG Controversies (CONT)	1,168	9.442	0.000	21.357	0.000	98.780
Leverage (LEV)	1,168	0.507	0.508	0.171	0.054	0.876
Firm Size (SIZE)	1,168	22.571	22.581	1.099	19.409	26.070
Return on Assets (ROA)	1,168	0.058	0.049	0.054	-0.164	0.514
GDP Growth (GDPG)	1,168	4.132	5.070	4.162	-9.518	9.690

Source: Stata output, processed by the authors (2026).

ROE has a mean of 13.732 and a median of 12.320, with values ranging from −49.450 to 99.000, indicating considerable heterogeneity in profitability and the presence of loss-making firm-years. The Altman Z"-Score has a mean of 6.386 and a median of 3.830, with a minimum of −2.869 that reflects observations in relatively weak financial condition. ESG Score averages 58.859 on the 0–100 scale. ESG Controversies has a mean of 9.442 but a median of 0.000, and even the seventy-fifth percentile remains at zero while the ninetieth percentile rises to approximately 43.023. The distribution of CONT is therefore heavily concentrated at low values: most firm-years record no material controversy, while a small subset records substantially higher controversy levels. This concentration is important for interpretation because the identification of the moderating effect rests on that minority of high-controversy firm-years. The minimum GDP growth of −9.518 confirms that the observation window includes a period of economic contraction.

Correlation and Multicollinearity

Pearson correlation is used as a preliminary description of the direction and strength of linear relationships and is not used to accept or reject hypotheses, because it does not yet control for time-invariant firm characteristics, year effects, control variables, or the panel error structure. The correlation matrix is presented in Table 5.

Table 5. Pearson Correlation Matrix

Variable	ROE	Z"	ESG	CONT	LEV	SIZE	ROA	GDPG
ROE	1.000							
Z"	0.281*	1.000						
ESG	−0.074*	−0.065*	1.000					
CONT	0.046	−0.006	0.100*	1.000				
LEV	0.068*	−0.683*	0.092*	0.069*	1.000			
SIZE	−0.089*	−0.499*	0.271*	0.236*	0.425*	1.000		
ROA	0.746*	0.628*	−0.080*	0.005	−0.403*	−0.286*	1.000	
GDPG	0.129*	0.093*	0.001	0.054	−0.047	0.025	0.156*	1.000

Source: Stata output, processed by the authors (2026). * denotes significance at the 5% level.

The correlation between ESG Score and ROE is −0.074 and the correlation between ESG Score and the Z"-Score is −0.065, both significant at the five percent level. Both correlations are weak and negatively signed, so they provide no descriptive indication in support of H1 or H2. The correlation between ESG Score and CONT is 0.100, indicating a positive but low association. Among the explanatory variables, the largest correlations are between leverage and firm size (0.425) and between leverage and ROA (−0.403); none approaches 0.80. Multicollinearity requires more careful treatment than a single variance inflation factor (VIF) table, because a VIF reported from a fixed-effects model requires an explicit statement of how it was obtained. VIFs are not defined for the within estimator in the usual way, since the firm effects are swept out rather than estimated as regressors, and the values previously reported — firm-level values between 1.33 and 1.50, 9.17 for GDP growth, 9.48 for the year-2020 dummy, and a mean of 2.63 — were obtained from a pooled ordinary least squares auxiliary regression that included the year dummies, and are therefore diagnostics for a model that is not the estimated specification. Read that way, the year-2020 dummy and GDP growth inflate each other almost by construction: the COVID-19 contraction of 2020 is the single largest common shock in the window, so a year dummy for 2020 and a growth rate that collapses in 2020 across all five countries carry nearly the same information.

Three diagnostics are therefore reported in place of the single pooled VIF table. First, within-VIFs computed on firm-demeaned data, which is the transformation the estimator

actually applies and therefore the collinearity relevant to the reported standard errors: these are 1.66 for ESG Score, 1.22 for leverage, 1.66 for firm size, and 1.14 for ROA, with 8.45 for GDP growth and 9.81 for the year-2020 dummy. Second, VIFs excluding the year dummies, which isolate collinearity among the substantive regressors from collinearity between GDPG and the year effects: every value falls to between 1.11 and 1.49 and the mean VIF falls from 2.63 to 1.33, with the GDP growth VIF falling from 9.17 to 1.11. This localises the issue precisely in the overlap with the year effects rather than in any relationship among the substantive regressors. Third, the condition number of the within regressor matrix, which does not depend on the choice of left-hand-side variable in an auxiliary regression, is 6.4, far below the value of 30 conventionally treated as indicating a collinearity problem.

GDP growth is retained in the baseline because it varies at the country-year level while the year effects are common across countries, so the two are not collinear by construction: GDPG carries the cross-country differences in macroeconomic conditions within each year, which the year effects cannot absorb, and its high pooled VIF reflects the 2020 shock specifically rather than general redundancy. Three alternative macroeconomic specifications were nevertheless estimated and the moderation results assessed for stability across all four. In (b) GDPG is dropped and the year fixed effects retained; in (c) GDPG is replaced by country $\times$ year fixed effects, which absorb every country-year macroeconomic shock non-parametrically, including GDPG itself, so that identification comes entirely from variation within country-year, and which is therefore the most demanding specification and the cleanest answer to the collinearity concern; in (d) GDPG is retained and the year fixed effects are replaced by a linear time trend. The moderation results are stable across all four. On ROE the interaction is -0.001316 ($p = 0.009$) in the baseline, -0.001323 ($p = 0.008$) in (b), -0.001373 ($p = 0.011$) in (c), and -0.001423 ($p = 0.004$) in (d). On the Z"-Score it is 0.000899 ($p < 0.001$) in the baseline, 0.000909 ($p < 0.001$) in (b), 0.000636 ($p = 0.003$) in (c), and 0.000911 ($p < 0.001$) in (d). The proximity of GDPG to the year effects therefore has no material bearing on the results the hypotheses concern.

Panel Diagnostics and Model Specification

Panel diagnostics and specification tests are reported in Table 6. The residuals of both outcome models depart from normality, heteroskedasticity is present across firms, and first-order autocorrelation is detected. In a large panel, non-normal residuals do not invalidate estimator consistency, but the presence of heteroskedasticity and serial correlation makes robust inference necessary, which supports the use of standard errors clustered at the firm level in the main specification and Driscoll–Kraay standard errors as a sensitivity check.

Table 6. Panel Diagnostic and Model Specification Tests

Test	Outcome	Statistic	p-value	Conclusion
Skewness/kurtosis normality	ROE	Adj. $\chi^2(2) = 315.03$	0.0000	Residuals not normally distributed
Skewness/kurtosis normality	Z"	Adj. $\chi^2(2) = 254.87$	0.0000	Residuals not normally distributed
Modified Wald heteroskedasticity	ROE	$\chi^2(146) = 144,047.53$	0.0000	Heteroskedasticity present
Modified Wald heteroskedasticity	Z"	$\chi^2(146) = 93,853.82$	0.0000	Heteroskedasticity present
Wooldridge autocorrelation	ROE	$F(1,145) = 8.196$	0.0048	Autocorrelation present
Wooldridge autocorrelation	Z"	$F(1,145) = 20.141$	0.0000	Autocorrelation present
F test: all firm effects = 0	ROE	$F(145,1010) = 6.58$	0.0000	Firm effects significant
F test: all firm effects = 0	Z"	$F(145,1010) = 26.79$	0.0000	Firm effects significant
Hausman FE vs RE	ROE	$\chi^2(12) = 15.65$	0.2077	RE not statistically rejected

Test	Outcome	Statistic	p-value	Conclusion
Hausman FE vs RE	Z"	$\chi^2(12) = 79.56$	0.0000	FE supported
Mundlak test (cluster-robust)	ROE	$\chi^2(7) = 9.64$	0.2098	RE assumption not rejected
Mundlak test (cluster-robust)	Z"	$\chi^2(7) = 25.20$	0.0007	FE required

Source: Stata output, processed by the authors (2026).

The F test rejects the null that all firm effects are zero for both outcomes (ROE: $F(145,1010) = 6.58$, $p < 0.001$; Z": $F(145,1010) = 26.79$, $p < 0.001$), so pooled ordinary least squares is inadequate as the main specification. The choice between fixed and random effects requires more careful treatment. The standard Hausman test rejects random effects for the Z"-Score ($\chi^2(12) = 79.56$, $p < 0.001$) but not for ROE ($\chi^2(12) = 15.65$, $p = 0.2077$). Two features of these outputs limit their evidential weight. First, both report that the $V_b - V_B$ difference matrix is not positive definite; this is not a numerical curiosity to be noted and set aside, because it means the estimated variance difference violates the condition the test requires, so the reported statistic and p-value are unreliable in either direction, including the non-rejection for ROE. Second, and more fundamentally, the Hausman test assumes that the random-effects estimator is efficient under the null; that assumption fails here, because the panel exhibits heteroskedasticity (modified Wald $\chi^2(146) = 144,047.53$ for ROE and 93,853.82 for Z", both $p < 0.001$) and first-order serial correlation (Wooldridge $F(1,145) = 8.196$, $p = 0.0048$ for ROE and $F(1,145) = 20.141$, $p < 0.001$ for Z"). Under these conditions the conventional Hausman test is not valid, which is the standard explanation for a non-positive-definite difference matrix.

A cluster-robust Mundlak auxiliary-regression test is therefore used in its place: the random-effects specification is augmented with the firm-level means of all time-varying regressors, and the joint significance of those means is tested with standard errors clustered at the firm level. Rejection indicates correlation between the regressors and the firm effects, that is, that fixed effects are required, and the test remains valid under heteroskedasticity and serial correlation. It reproduces the pattern of the Hausman results on a valid basis. For the Z"-Score it rejects decisively ($\chi^2(7) = 25.20$, $p = 0.0007$), so fixed effects are required. For ROE it does not reject ($\chi^2(7) = 9.64$, $p = 0.2098$), so the random-effects assumption is not contradicted for that outcome.

Independently of the test outcome, firm fixed effects are appropriate for the ROE model on identification grounds, and this is argued here rather than asserted as an a priori preference. ESG Score is highly persistent within firms: it is driven largely by the maturity of a firm's disclosure infrastructure, its industry, its ownership structure, and its size, which are precisely the characteristics that also determine profitability. A random-effects estimator assumes these firm effects are uncorrelated with ESG Score, which is implausible in this setting and would load cross-firm differences in disclosure capacity onto the ESG coefficient, whereas fixed effects identify the association from within-firm changes over time, which is the variation the hypotheses are about. Because the Mundlak test does not reject random effects for ROE, the random-effects specification was nevertheless estimated in full for both outcomes with clustered standard errors, and it is reported in Table 10. The substantive conclusions do not change: under random effects the ROE interaction is -0.001391 ($p = 0.004$) against -0.001316 ($p = 0.009$) under fixed effects, and the Z" interaction is 0.000820 ($p < 0.001$) against 0.000899 ($p < 0.001$). Both direct ESG coefficients remain insignificant under random effects (ROE $p = 0.519$; Z" $p = 0.802$). All four models are jointly significant with Prob > F of 0.0000. The regression results are presented in Table 7.

Hypothesis Testing

Table 7. Fixed Effects Regression Results

Variable	Model 1 ROE	Model 2 ROE + Moderation	Model 3 Z''-Score	Model 4 Z''-Score + Moderation
ESG / cESG	0.000508 (0.984)	-0.001222 (0.961)	-0.011000 (0.457)	-0.009793 (0.495)
CONT / cCONT	–	0.010310 (0.177)	–	-0.006569 (0.034)
ESG × CONT	–	-0.001316 (0.009)	–	0.000899 (0.000)
LEV	17.239820 (0.000)	17.209580 (0.000)	-18.933230 (0.000)	-18.918610 (0.000)
SIZE	1.987025 (0.214)	1.840603 (0.242)	-0.290439 (0.709)	-0.192845 (0.802)
ROA	187.627500 (0.000)	188.183800 (0.000)	24.221500 (0.000)	23.853590 (0.000)
GDPG	-0.078366 (0.411)	-0.074721 (0.433)	0.096314 (0.005)	0.093772 (0.006)
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
N	1,168	1,168	1,168	1,168
Within R ²	0.6351	0.6380	0.3593	0.3704
F-statistic	35.38	32.62	11.52	10.36
Prob > F	0.0000	0.0000	0.0000	0.0000

Source: Stata output, processed by the authors. Figures in parentheses are two-tailed p-values. Standard errors are clustered at the firm level in all four models; the previous version clustered Models 2 and 4 only, and the direct-effect coefficients in Models 1 and 3 are unchanged by the correction. F-statistics are the cluster-robust model F tests. *** p < 0.01, ** p < 0.05, * p < 0.10.

All inference below is two-tailed. The coefficient of ESG Score in the ROE model is 0.000508 with p = 0.984: the sign is positive as predicted, but the coefficient is far from significant, so H1 is not supported. In the Z''-Score model the coefficient of ESG Score is -0.011000 with p = 0.457: the coefficient is insignificant and additionally carries the opposite sign to the one H2 requires, so H2 is not supported. The interaction coefficient in the ROE model is -0.001316 with p = 0.009, which is significant at the one percent level with the predicted negative sign, so H3 is supported in the baseline specification. The interaction coefficient in the Z''-Score model is +0.000899 with p < 0.001, which is significant with the predicted positive sign, so H4 is supported in the baseline specification. Table 8 sets out each decision so that it can be traced mechanically from the reported coefficient and p-value: a hypothesis is recorded as supported only when both conditions hold, namely p < 0.05 on a two-tailed test and a coefficient sign matching the prediction. The stability of the two supported decisions across alternative specifications is assessed separately in Table 10.

Table 8. Hypothesis Testing Decisions

Hypothesis	Parameter	Model	Coefficient	Std. error	P (two-tailed)	Predicted sign	Observed sign	Sign matches	p < 0.05	Decision
H1: ESG Score → ROE	β_1	1	0.000508	0.025303	0.984	+	+	Yes	No	Not supported
H2: ESG Score → Z''-Score	β_1	3	-0.011000	0.014748	0.457	+	-	No	No	Not supported

H3: ESG × CONT → ROE	β_3	2	−0.001316***	0.000495	0.009	−	−	Yes	Yes	Supported
H4: ESG × CONT → Z''-Score	β_3	4	0.000899***	0.000228	<0.001	+	+	Yes	Yes	Supported

Source: Stata output, processed by the authors. All p-values are two-tailed. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Decisions are mechanical: H1 fails the significance condition; H2 fails both the significance and the sign condition; H3 and H4 satisfy both. The supported decisions apply to the baseline specification only.

Because the direct ESG coefficient is insignificant in both models while the interaction is significant, the substantive result is contained in the conditional effects rather than in any single coefficient. Average marginal effects of ESG Score are therefore reported for both outcomes at the same set of controversy levels, with 95 percent confidence intervals, and are plotted in Figures 2 and 3. Both moderation models are estimated with ESG Score and CONT mean-centred (mean CONT ≈ 9.39), so the reported ESG coefficient is the marginal effect at the sample average controversy level. The marginal effects are $\partial \text{ROE} / \partial \text{ESG} = \beta_1 + \beta_3(\text{CONT} - 9.39)$ from Model 2 and $\partial Z'' / \partial \text{ESG} = \beta_1 + \beta_3(\text{CONT} - 9.39)$ from Model 4, and are reported in Table 9.

Table 9. Average Marginal Effect of ESG Score at Selected Controversy Levels

Controversy level	CONT	$\partial \text{ROE} / \partial \text{ESG}$	95% CI	p	$\partial Z'' / \partial \text{ESG}$	95% CI	p
No controversy (also the median and the 75th percentile)	0.00	+0.01113	[−0.0409, 0.0632]	0.673	−0.01824	[−0.0465, 0.0100]	0.204
Sample mean	9.39	−0.00122	[−0.0505, 0.0481]	0.961	−0.00979	[−0.0381, 0.0185]	0.495
90th percentile	43.02	−0.04547*	[−0.0983, 0.0074]	0.091	+0.02045	[−0.0128, 0.0537]	0.226
95th percentile	63.64	−0.07258**	[−0.1365, −0.0087]	0.026	+0.03899**	[0.0001, 0.0779]	0.050
Maximum	88.04	−0.10469**	[−0.1863, −0.0231]	0.012	+0.06093**	[0.0138, 0.1081]	0.012

Source: authors' estimation (Stata margins, verified by independent replication). Point estimates are computed directly from the Model 2 and Model 4 coefficients; confidence intervals and p-values are obtained from the estimated variance–covariance matrix.

Setting each marginal effect to zero gives the controversy level at which the estimated association changes sign. For ROE the marginal effect of ESG Score turns from positive to negative at CONT ≈ 8.5 , slightly below the sample mean. For the Z''-Score it turns from negative to positive at CONT ≈ 20.3 , which lies between the 75th percentile (zero) and the 90th percentile (43.02) of the controversy distribution. The Johnson–Neyman regions, that is the ranges over which each marginal effect is statistically distinguishable from zero at the five per cent level, are narrower than the sign-change points suggest and are shown as shaded regions in Figures 2 and 3: the ROE marginal effect is significant only for CONT > 51.7 and the Z'' marginal effect only for CONT > 63.5 . Both thresholds lie above the 90th percentile of the controversy distribution, and the second sits essentially at the 95th percentile (63.64).

The marginal-effects results sharpen, and also constrain, the reading of H3 and H4. On the profitability channel the estimated association of ESG Score with ROE is positive among controversy-free firm-years (+0.0111) and negative among high-controversy firm-years (−0.0455 at the 90th percentile and −0.1047 at the maximum). On the solvency channel the pattern is reversed: −0.0182 among controversy-free firm-years and +0.0205 at the 90th

percentile, rising to +0.0609 at the maximum. Three cautions attach to both statements and are carried through the discussion. First, the point at which each association becomes statistically distinguishable from zero lies far into the upper tail of the controversy distribution, so the significant regions are populated by a small minority of firm-years and the estimates there are correspondingly imprecise; the confidence bands in Figures 2 and 3 should be read before the point estimates. Second, the interaction coefficient is estimated over the whole distribution and is significant even though the marginal effect is not significant at the mean, which is the expected pattern when an interaction is identified from the tail rather than the centre. Third, these are conditional associations recovered from an interaction term, not effects of moving a given firm along the controversy distribution.

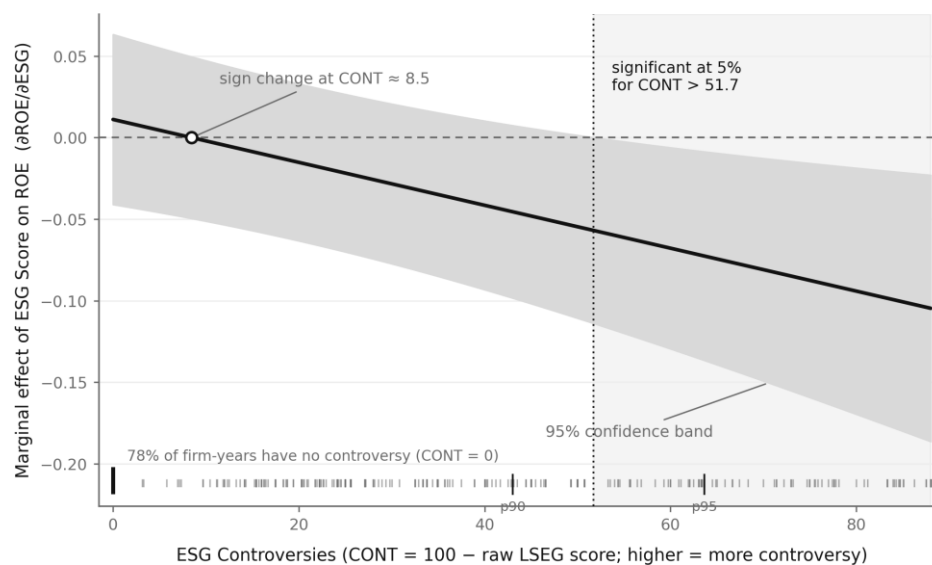


Figure 2. Marginal effect of ESG Score on return on equity across the ESG Controversies distribution, with 95% confidence band. The shaded region marks the Johnson–Neyman range in which the marginal effect is significant at the 5% level (CONT > 51.7). The rug below the axis shows the distribution of observations; 78.4% of firm-years record CONT = 0. Source: authors' estimation from Model 2.

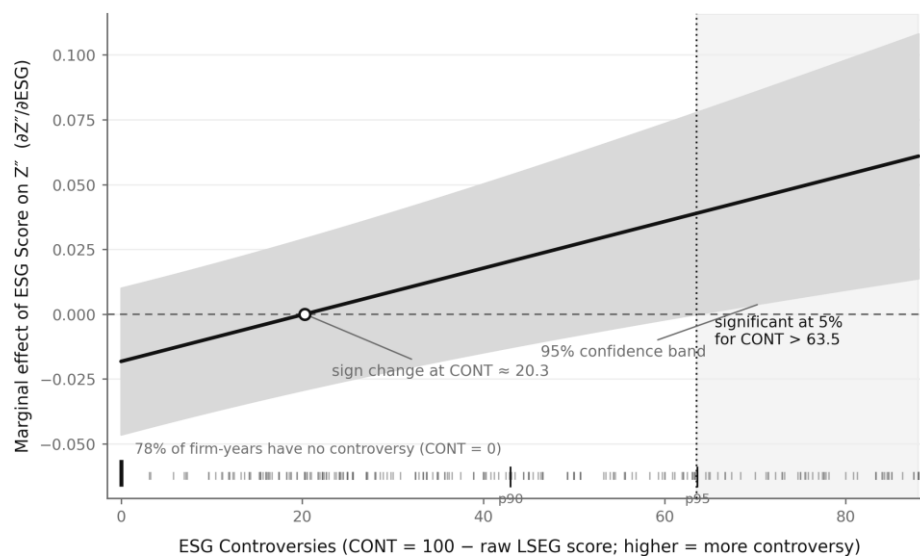


Figure 3. Marginal effect of ESG Score on the Altman Z''-Score across the ESG Controversies distribution, with 95% confidence band. The shaded region marks the Johnson–Neyman range in which the marginal effect is significant at the 5% level (CONT > 63.5). A higher Z''-Score denotes better financial health. Source: authors' estimation from Model 4.

Robustness tests were conducted to assess the stability of the findings across estimators, timing, controls, sub-samples, and error structures. The results are summarised in Table 10.

Table 10. Summary of Robustness Tests

Specification	ESG $\times$ CONT $\rightarrow$ ROE	ESG $\times$ CONT $\rightarrow$ Z"-Score
Panel A. Firm fixed effects retained; one element varied at a time		
Baseline: firm FE + year FE, clustered	-0.001316*** (0.009)	0.000899*** (<0.001)
Without ROA	-0.000744 (0.433)	0.000972*** (<0.001)
Without ROA and leverage	-0.000751 (0.427)	0.000952*** (0.001)
Country $\times$ year fixed effects (absorb GDPG)	-0.001373** (0.011)	0.000636*** (0.003)
One-year lag (N = 1,022)	-0.000510 (0.325)	0.000936*** (0.002)
ASEAN sub-sample (N = 744)	-0.000649 (0.346)	0.000527 (0.125)
India sub-sample (N = 424)	-0.001700** (0.025)	0.000724*** (0.009)
162 firms, before the extreme-value screen (N = 1,296)	0.000063 (0.952)	0.000980*** (0.001)
Random effects, clustered	-0.001391*** (0.004)	0.000820*** (<0.001)
Driscoll–Kraay standard errors	-0.001316* (0.069)	0.000899 (0.050)
Panel B. Firm fixed effects replaced by country fixed effects (pooled OLS)		
Pooled OLS + country and year FE, clustered	-0.001418* (0.064)	0.000126 (0.816)
Pooled OLS, without ROA	-0.000727 (0.665)	0.000291 (0.615)
Pooled OLS, ASEAN sub-sample	-0.001447 (0.263)	0.000192 (0.752)
Pooled OLS, India sub-sample	-0.001567* (0.092)	-0.000532 (0.498)

Source: authors' estimation. Each cell reports the ESG $\times$ CONT interaction coefficient with the two-tailed p-value in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Standard errors are clustered at the firm level except in the Driscoll–Kraay row. Panel A varies one element at a time while holding the estimator fixed; Panel B changes the estimator itself and is reported separately for that reason. The Driscoll–Kraay row is taken from the xtsc output of the original estimation; every other row was re-estimated for this revision.

Reading the robustness evidence requires one methodological point to be settled first. In the previous version of this study every robustness specification other than Driscoll–Kraay replaced the firm fixed effects with country fixed effects in a pooled regression. Those specifications therefore changed two things at once — the control set or the sample, and the estimator — so they could not isolate the sensitivity of the results to either. In this revision the robustness tests retain the firm fixed effects of the baseline and vary one element at a time, and the pooled specifications are reported separately as a distinct block. This change alters the conclusions materially, and in opposite directions for the two moderation hypotheses. All p-values below are two-tailed.

For the Z"-Score interaction the evidence is considerably stronger than the previous version reported. The coefficient is positive in every specification that retains firm fixed effects. It is significant at the one per cent level in eight of the ten and at the five per cent level in a ninth: the baseline (0.000899, $p < 0.001$), without ROA (0.000972, $p < 0.001$), without ROA and leverage (0.000952, $p = 0.001$), under country $\times$ year fixed effects (0.000636, $p = 0.003$), in the one-year lag model (0.000936, $p = 0.002$), in the India sub-sample (0.000724, $p = 0.009$), on the 162-firm sample that precedes the extreme-value screen (0.000980, $p = 0.001$), under random effects (0.000820, $p < 0.001$), and under Driscoll–Kraay standard errors (0.000899, $p = 0.050$, the five per cent case). The only specification in which it is not significant at conventional levels is the ASEAN sub-sample, where the coefficient keeps its sign and roughly half its magnitude (0.000527, $p = 0.125$). H4 therefore survives the removal of ROA, which was the specific concern raised about the previous version, and it survives the most demanding specification in the study, in which country $\times$ year fixed effects absorb every macroeconomic

shock non-parametrically. What it does not survive is the removal of the firm fixed effects: in every pooled specification with country fixed effects the coefficient falls by an order of magnitude and loses significance (0.000126, $p = 0.816$ in the baseline pooled model). That contrast is informative rather than disconfirming, and is discussed below.

For the ROE interaction the position is the reverse, and weaker than the previous version reported. The coefficient is negative in thirteen of the fourteen specifications, so its direction is stable, but it is significant at the five per cent level in only four of them: the baseline (-0.001316 , $p = 0.009$), under country $\times$ year fixed effects (-0.001373 , $p = 0.011$), in the India sub-sample (-0.001700 , $p = 0.025$), and under random effects (-0.001391 , $p = 0.004$). It is not significant when ROA is excluded (-0.000744 , $p = 0.433$), when ROA and leverage are both excluded (-0.000751 , $p = 0.427$), in the one-year lag model (-0.000510 , $p = 0.325$), or in the ASEAN sub-sample (-0.000649 , $p = 0.346$). The most consequential failure is the last row of the table: on the 162-firm sample that precedes the extreme-value screen the coefficient changes sign and is indistinguishable from zero ($+0.000063$, $p = 0.952$). This is not an artefact of one winsorisation choice, since the same result obtains when that sample is winsorised at the 2.5th and 97.5th or at the 5th and 95th percentiles ($p = 0.709$ and $p = 0.215$ respectively). H3 therefore depends on the exclusion of the sixteen firms that fail the extreme-value screen, and it is not robust to the inclusion of firms with return on equity above 100 per cent or with negative book equity.

Two conclusions follow, and they are the opposite of what the previous version reported. H4 is robust: the positive interaction on the Z"-Score holds across control sets, macroeconomic specifications, timing, the extreme-value screen, the estimator, and one of the two sub-samples. Its confinement to specifications with firm fixed effects is informative rather than a weakness, because it identifies where the association lives: the evidence comes from changes in the same firm's condition as its controversy exposure varies over time, and is absorbed by design in specifications that compare across firms. H3 is fragile: although the negative direction is stable, statistical significance requires both the full control set and the exclusion of extreme-profitability firms, and neither requirement is innocuous, since ROA overlaps mechanically with the dependent variable and the extreme-value screen removes sixteen firms on grounds that are defensible for estimation but are not part of the hypothesis. H3 is accordingly reported as directionally consistent but not robust, and the implications drawn from it are qualified throughout. H1 receives no support in any specification, and H2 is not supported at conventional two-tailed levels anywhere, the closest cases being the one-year lag model and the India sub-sample.

Discussion

ESG Score and Financial Performance

The results show no significant association between ESG Score and financial performance as proxied by ROE. An increase in the sustainability score of non-financial firms in the five Asian emerging markets has therefore not been accompanied by measurably higher equity returns within the same period. From the perspective of stakeholder theory, benefits arising from relationships with employees, customers, suppliers, regulators, and capital providers may require time before any economic consequence becomes measurable, while ESG implementation can create up-front costs and trade-offs that offset any immediate profitability gain. The finding reinforces the literature showing that the ESG-performance relationship is contextual and heterogeneous: Rao et al. (2023) demonstrate that pillar effects differ along the distribution of profitability, and Narula et al. (2024) report that ESG effects vary across emerging markets. The result differs from Alareeni and Hamdan (2020), which confirms that ESG benefits cannot be assumed to be uniform across regions, periods, and firm characteristics.

ESG Score and Financial Distress

No significant association between ESG Score and financial health is found in the main model. Because the Z'' -Score moves inversely to distress risk, support for H2 would require a positive and significant coefficient; the coefficient obtained is negative and insignificant, so there is no evidence of an average association between a higher ESG Score and better financial health once fixed firm heterogeneity and year effects are absorbed. This differs from Boubaker et al. (2020), Shi et al. (2023), and Zhou et al. (2025), which generally report that corporate social responsibility or ESG is associated with lower distress risk. The difference may reflect the heterogeneity of emerging market contexts, industry composition, and the fact that ESG Score measures sustainability policy and performance rather than the capacity to service obligations. As shown below, the moderation results provide a further explanation: the association between ESG Score and financial health is conditional on the controversy level, whereas the direct-effect model estimates a single average association over a sample in which 78.4 per cent of firm-years record no controversy.

ESG Controversies as a Moderator of Financial Performance

H3 is supported in the baseline specification but is not robust, and the discussion below is qualified accordingly. The negative and significant interaction coefficient indicates that a higher controversy level is associated with a less favourable ESG–ROE relationship, and the marginal-effects analysis locates the change of sign at $CONT \approx 8.5$, just below the sample mean. The direction of the interaction is stable across thirteen of the fourteen specifications estimated, but statistical significance requires both the full control set and the exclusion of the sixteen firms that fail the extreme-value screen; on the 162-firm sample the coefficient changes sign and is indistinguishable from zero. The reading that the evidence supports is therefore a directional one. It is consistent with stakeholder theory, in which controversies damage stakeholder relationships and reduce the capacity of the firm to convert ESG practice into economic benefit; with signalling theory, in which controversies act as a counter-signal that reduces the credibility of the positive ESG signal; and with legitimacy theory, in which controversies reveal a legitimacy gap between disclosure and conduct. The result aligns in direction with Elamer and Boulhaga (2024) and differs from Rahmadini and Hartanti (2025), who found no significant moderation in the Indonesian setting; the present study, which also finds no significant moderation in its ASEAN sub-sample, is consistent with theirs on that point rather than in conflict with it. Because H1 is not significant, the finding should not be read as controversies eroding a proven positive effect. The accurate statement is that higher controversy levels are associated with a progressively less favourable relationship between sustainability performance and profitability in the estimation sample, that this association is not established with confidence, and that it is particularly sensitive to how extreme-profitability firms are treated.

ESG Controversies as a Moderator of Financial Distress

H4 is supported and is the more robust of the two moderation results. The positive and significant interaction on the Altman Z'' -Score indicates that the estimated association between ESG Score and financial health is more favourable among firm-years with higher recorded controversy exposure, with the change of sign at $CONT \approx 20.3$. The coefficient keeps its sign and significance when ROA is removed, when ROA and leverage are both removed, under country $\times$ year fixed effects, in the one-year lag model, in the India sub-sample, on the sample that precedes the extreme-value screen, under random effects, and under Driscoll–Kraay standard errors; it is not significant only in the ASEAN sub-sample, and in the pooled specifications that replace firm fixed effects with country fixed effects. That last pattern locates the association: it is identified from changes within the same firm as its controversy exposure

varies over time, and is absorbed by design when firms are compared with one another. The account that follows is a theoretical interpretation of that statistical pattern and not a mechanism measured in this dataset: the study observes an interaction coefficient, and does not observe creditor decisions, loan pricing, covenant terms, credit ratings, bond spreads, or any other direct indicator of how capital providers process ESG information. No claim is therefore made that creditors do assign greater weight to the ESG Score. What can be said is that the observed pattern is consistent with two propositions developed in the theoretical framework, and inconsistent with a reading in which ESG Score carries the same informational weight regardless of controversy context. Under the signal-salience interpretation (Proposition 2), non-financial risk is not salient to downside-oriented evaluators in the absence of an incident, so an ESG Score would be treated as routine disclosure; when an incident makes such risk salient, evaluators would require information distinguishing firms able to contain it, and the ESG Score would acquire diagnostic value. Under the complementary legitimacy-repair interpretation, a controversy opens a legitimacy gap, and the environmental policies, safety procedures, grievance channels, compliance functions, and board oversight that a high ESG Score proxies are the organisational instruments available for containment. Both interpretations are consistent with Broadstock et al. (2021) and Chiaramonte et al. (2022), who find greater resilience among stronger ESG performers during periods of stress, and with Elamer and Boulhaga (2024), who show that the consequences of controversies depend on the governance mechanisms a firm already possesses. Competing interpretations cannot be excluded with these data: the same pattern would arise if firms that are already financially stronger are both better able to sustain high ESG ratings and better able to absorb the financial consequences of an incident, in which case the interaction reflects unobserved firm capacity rather than the informational role of the score. Firm fixed effects absorb the time-invariant part of such capacity but not its time-varying part. A further caution follows from the marginal-effects analysis: although the interaction is estimated precisely, the marginal effect of ESG Score on the Z'' -Score is itself distinguishable from zero only above $CONT \approx 63.5$, which is essentially the 95th percentile of the controversy distribution (63.64), so the conditional effect is individually significant only in the sparsely populated upper tail.

The contribution of this study is the evidence that the estimated interaction sign of a single moderating construct differs between a profitability outcome and a solvency outcome within the same panel. The two halves of that contribution do not rest on equally firm ground, and the difference is stated rather than smoothed over. On the solvency channel the positive interaction is robust across control sets, macroeconomic specifications, timing, the extreme-value screen, and the estimator. On the profitability channel the negative interaction is directionally consistent but statistically fragile. The proposed explanation — that continuously attentive, upside-oriented evaluators discount a signal contradicted by events, while intermittently attentive, downside-oriented evaluators weight the signal only once an event makes the underlying risk salient — is a theoretical account offered to organise the finding, and is advanced as a proposition for future testing rather than as a demonstrated mechanism. This pattern also offers a possible reconciliation of the null result for H2 with the interaction result for H4: if the association between ESG Score and the Z'' -Score changes sign across the controversy distribution and 78.4 per cent of firm-years record no controversy, an average association estimated without an interaction term would be expected to be small and imprecise, which is what Model 3 returns. The insignificance of H2 is therefore consistent with the absence of an average association rather than with the irrelevance of sustainability performance to financial health.

CONCLUSION AND RECOMMENDATIONS

Conclusion

This study examined the association between ESG Score and two financial outcomes — financial performance, proxied by return on equity, and financial health, proxied by the Altman Z"-Score — together with the moderating role of ESG Controversies, using 1,168 firm-year observations from 146 non-financial firms in Indonesia, Malaysia, the Philippines, Thailand, and India over 2017–2024. The design is observational and the estimates are conditional associations; the language below is chosen accordingly. ESG Score shows no statistically significant association with either outcome: the coefficient on ROE is 0.00051 ($p = 0.984$) and on the Z"-Score is -0.01100 ($p = 0.457$), both two-tailed, so H1 and H2 are not supported. In the sample and period examined, changes in a firm's sustainability score were not accompanied by measurable changes in equity-based profitability or in financial health once time-invariant firm heterogeneity and common year effects are absorbed.

The two moderation results are supported in the baseline specification, but they are not equally secure and the conclusion distinguishes them rather than reporting a single verdict. The $\text{ESG} \times \text{Controversies}$ interaction is negative and significant on ROE ($\beta = -0.001316$, $p = 0.009$), supporting H3, and positive and significant on the Altman Z"-Score ($\beta = 0.000899$, $p < 0.001$), supporting H4. Stated in terms of the measured dependent variable: a higher controversy level is associated with a more positive estimated relationship between ESG Score and the Z"-Score, and since a higher Z"-Score corresponds to lower distress risk, this is a positive interaction on financial health rather than a demonstration that ESG infrastructure causes better financial health during controversies. H4 is robust: the positive interaction holds when ROA is removed, when ROA and leverage are both removed, under country $\times$ year fixed effects, in a one-year lag model, in the India sub-sample, on the 162-firm sample that precedes the extreme-value screen, under random effects, and under Driscoll–Kraay standard errors, failing only in the ASEAN sub-sample and in pooled specifications that replace the firm fixed effects. H3 is not robust: its negative direction is stable across specifications, but significance requires both the full control set and the exclusion of the sixteen extreme-value firms, and the coefficient changes sign when those firms are restored. Both interactions are identified from the 21.6 per cent of firm-years in which a controversy is recorded, and the marginal effect of ESG Score is itself statistically distinguishable from zero only in the upper tail of that distribution. Subject to those qualifications, the pattern documented here is that a single moderating construct is associated with interaction terms of opposite sign on a profitability outcome and a solvency outcome within the same panel. The credibility-discount and signal-salience accounts developed in the theoretical framework are offered as interpretations consistent with that pattern; neither creditor behaviour nor market pricing is observed in this dataset, so neither account is tested directly. The practical upshot, stated as an association rather than a prescription, is that the financial relevance of a sustainability score in these markets appears to depend on the controversy context in which it is read and on which financial outcome is being assessed, which is the answer this study offers to the four research questions posed in the introduction.

Limitations

This study has several limitations. It is observational, so the results indicate empirical associations rather than absolute causality; firm and year fixed effects, clustered standard errors, lag models, and robustness tests reduce but do not eliminate reverse causality and omitted variable bias. The sample depends on the availability of complete ESG and financial data in LSEG Workspace, and firms with ESG coverage tend to be larger and more transparent, which may introduce selection bias and limit generalisation. The aggregate ESG Score may also mask differences among the environmental, social, and governance pillars. ESG

Controversies are measured through a transformation of a score built from publicly reported events, so firms with greater media exposure may be more likely to be recorded as experiencing controversy, and the distribution of CONT is heavily concentrated at low values, so identification of the moderating effect rests on a small subset of firm-years. The observation window includes the COVID-19 pandemic, whose effects may differ across countries and industries despite the use of year fixed effects. Finally, the evidence for the fourth hypothesis comes from specifications with firm fixed effects, which locates the association in within-firm variation rather than limiting the support for it, and the study does not distinguish the type, severity, or duration of controversies, nor does it observe the remediation steps firms take after an incident, so the legitimacy repair mechanism is inferred from a consistent statistical pattern rather than measured directly. Three further limitations follow from the sample construction and are stated explicitly because they bear directly on how firmly the results should be held. First, firms were retained only where they had complete ESG and financial coverage in all eight years, so the estimation sample is a balanced panel of continuously rated firms, from which any firm that entered distress severely enough to be suspended or delisted is absent by construction; the distress results should be read as applying to that population. Second, sixteen firms were removed by an extreme-value screen ($|ROE| > 100\%$, $Z'' > 40$, or leverage > 1 in any year). That screen is defensible for estimation, since such observations reflect near-zero or negative book equity rather than economic profitability, but it is not part of any hypothesis, and the support for H3 does not survive its removal: on the 162-firm sample the ROE interaction changes sign and becomes insignificant. The profitability-channel result should therefore be regarded as conditional on that screen. Third, identification of both interactions rests on the 21.6 per cent of firm-years with a recorded controversy, and the marginal effect of ESG Score is statistically distinguishable from zero only above the 90th percentile of the controversy distribution in both models, so the precision of the conditional estimates in the sparse upper tail is limited.

Recommendations

Future studies should separate the environmental, social, and governance pillar scores in order to identify which dimension is most relevant to profitability and to distress risk, and should use alternative distress measures such as the Ohlson O-Score, probability of default, or interest coverage as comparators for the Altman Z'' -Score. The role of ESG Controversies can be analysed in greater detail by type, severity, frequency, and duration, complemented by data on remediation, which would allow the signal salience and legitimacy repair mechanisms to be tested directly rather than inferred. Longer observation windows and stronger identification designs, such as dynamic panels, instrumental variables, propensity-score approaches, or event-based designs around the emergence of controversies, would strengthen inference, while heterogeneity analyses by environmentally sensitive industry, ownership structure, governance quality, firm size, leverage, and country would clarify the boundaries of the findings.

To move the mechanisms proposed in this study from interpretation to measurement, future work should obtain creditor-side and market-pricing outcomes — loan spreads, covenant structure, credit rating changes, bond yields — alongside the accounting outcomes used here, and should model controversy type, severity, duration, and observed remediation rather than a single aggregate index. Event-based designs around the onset of a controversy, dynamic panel or instrumental-variable strategies, and analysis on unbalanced panels that retain distressed and delisted firms would each address a limitation this design could not.

Implications for practice are separated below into two categories. Evidence-based implications follow from the estimated associations reported in this study, with their qualifications attached. Prescriptive recommendations are practice suggestions that are

consistent with the findings but rest additionally on theory, on prior literature, and on normative judgement; they are not established by this study's design and are labelled as such.

For investors and analysts — evidence-based: in this sample the estimated association between ESG Score and equity profitability varies with controversy exposure, being positive among controversy-free firm-years and negative above a controversy level of approximately 8.5, with the negative interaction directionally stable across specifications, so an ESG Score read without the accompanying controversy record describes a different empirical object at different points of the controversy distribution. Prescriptive: analysts may wish to treat ESG Score and event-based controversy data as joint inputs rather than as substitutes; this study evaluates no particular screening rule or investment strategy and implies no performance claim for one.

For creditors — evidence-based: the estimated association between ESG Score and the Altman Z"-Score is more favourable at higher controversy levels in specifications with firm fixed effects, with the sign change located at approximately $CONT \approx 20.3$. This result is robust across control sets, macroeconomic specifications, timing, the extreme-value screen, and the estimator, failing only in the ASEAN sub-sample; it is nevertheless identified from the 21.6 per cent of firm-years in which a controversy is recorded, and the marginal effect is distinguishable from zero only in the upper tail. Prescriptive: the finding is suggestive that ESG information may have greater discriminatory value within the group of controversy-exposed borrowers than across borrowers generally, but this study observes no lending decisions, credit ratings, or pricing, so it provides no basis for a change in credit assessment practice, and the absence of any observed creditor-side variable argues against acting on it without corroborating evidence.

For firms — evidence-based: the study finds no significant average association between ESG Score and either profitability or financial health, and finds that a higher controversy level is associated with a less favourable ESG–profitability relationship; it does not establish that building ESG infrastructure improves financial outcomes, and it does not observe whether firms with high scores took different actions during or after an incident. Prescriptive: on the basis of the theoretical framework rather than of a causal estimate, managing incident risk alongside disclosure quality is a defensible strategic posture, since a score contradicted by events is, on the evidence here, associated with a weaker relationship to profitability; this is a plausibility argument and not a demonstrated return on ESG investment.

For regulators and exchanges — evidence-based: the economic information content of ESG data in this sample varies with controversy context, which is descriptive evidence that policy disclosure and incident disclosure are not interchangeable inputs for report users. Prescriptive: disclosure frameworks covering incidents, violations, remediation actions, and recovery mechanisms alongside policies and targets would be consistent with that pattern, though this study evaluates no disclosure regime and cannot estimate the effect of adopting one.

Author Contributions

All authors made significant contributions to this study. These contributions included research conceptualisation, development of the theoretical framework, data collection, data analysis, interpretation of results, manuscript preparation, scholarly revision, and final approval of the manuscript for publication. All authors confirm that they have read and approved the final version of the article.

REFERENCES

- Alareeni, B. A., & Hamdan, A. (2020). ESG impact on performance of US S&P 500-listed firms. *Corporate Governance: The International Journal of Business in Society*, 20(7), 1409–1428. <https://doi.org/10.1108/CG-06-2020-0258>
- Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance*, 23(4), 589–609. <https://doi.org/10.2307/2978933>
- Altman, E. I. (2000). Predicting financial distress of companies: Revisiting the Z-score and ZETA models [Working paper]. Stern School of Business, New York University.
- Boubaker, S., Cellier, A., Manita, R., & Saeed, A. (2020). Does corporate social responsibility reduce financial distress risk? *Economic Modelling*, 91, 835–851. <https://doi.org/10.1016/j.econmod.2020.05.012>
- Broadstock, D. C., Chan, K., Cheng, L. T. W., & Wang, X. (2021). The role of ESG performance during times of financial crisis: Evidence from COVID-19 in China. *Finance Research Letters*, 38, Article 101716. <https://doi.org/10.1016/j.frl.2020.101716>
- Buallay, A. (2019). Is sustainability reporting (ESG) associated with performance? Evidence from the European banking sector. *Management of Environmental Quality: An International Journal*, 30(1), 98–115. <https://doi.org/10.1108/MEQ-12-2017-0149>
- Chiaromonte, L., Dreassi, A., Girardone, C., & Piserà, S. (2022). Do ESG strategies enhance bank stability during financial turmoil? Evidence from Europe. *The European Journal of Finance*, 28(12), 1173–1211. <https://doi.org/10.1080/1351847X.2021.1964556>
- Connelly, B. L., Certo, S. T., Ireland, R. D., & Reutzel, C. R. (2011). Signaling theory: A review and assessment. *Journal of Management*, 37(1), 39–67. <https://doi.org/10.1177/0149206310388419>
- Darsono, D., Ratmono, D., Tujori, A., & Clarisa, T. Y. (2025). The relationship between ESG, financial performance, and cost of debt: The role of independent assurance. *Cogent Business & Management*, 12(1), Article 2437137. <https://doi.org/10.1080/23311975.2024.2437137>
- Deegan, C. (2002). Introduction: The legitimising effect of social and environmental disclosures – A theoretical foundation. *Accounting, Auditing & Accountability Journal*, 15(3), 282–311. <https://doi.org/10.1108/09513570210435852>
- Deegan, C. M. (2019). Legitimacy theory: Despite its enduring popularity and contribution, time is right for a necessary makeover. *Accounting, Auditing & Accountability Journal*, 32(8), 2307–2329. <https://doi.org/10.1108/AAAJ-08-2018-3638>
- Donaldson, T., & Preston, L. E. (1995). The stakeholder theory of the corporation: Concepts, evidence, and implications. *Academy of Management Review*, 20(1), 65–91. <https://doi.org/10.5465/amr.1995.9503271992>
- Elamer, A. A., & Boulhaga, M. (2024). ESG controversies and corporate performance: The moderating effect of governance mechanisms and ESG practices. *Corporate Social Responsibility and Environmental Management*, 31(4), 3312–3327. <https://doi.org/10.1002/csr.2749>
- Elkington, J. (1997). Cannibals with forks: The triple bottom line of 21st century business. Capstone.
- Freeman, R. E. (1984). *Strategic management: A stakeholder approach*. Pitman.
- Jensen, M. C. (2001). Value maximization, stakeholder theory, and the corporate objective function. *Journal of Applied Corporate Finance*, 14(3), 8–21. <https://doi.org/10.1111/j.1745-6622.2001.tb00434.x>
- Landi, G. C., Iandolo, F., Renzi, A., & Rey, A. (2022). Embedding sustainability in risk management: The impact of environmental, social, and governance ratings on corporate financial risk. *Corporate Social Responsibility and Environmental Management*, 29(4), 1096–1107. <https://doi.org/10.1002/csr.2256>

- Le, L. T. (2024). Impact of environmental, social and governance practices on financial performance: Evidence from listed companies in Southeast Asia. *Cogent Business & Management*, 11(1), Article 2379568. <https://doi.org/10.1080/23311975.2024.2379568>
- LSEG. (2024). Environmental, social and governance scores from LSEG. London Stock Exchange Group. <https://www.lseg.com/en/data-analytics/sustainable-finance/esg-scores>
- Lohmann, C., Möllenhoff, S., & Lehner, S. (2025). On the relationship between financial distress and ESG scores. *Corporate Social Responsibility and Environmental Management*, 32(5), 6377–6401. <https://doi.org/10.1002/csr.70033>
- Mitchell, R. K., Agle, B. R., & Wood, D. J. (1997). Toward a theory of stakeholder identification and salience: Defining the principle of who and what really counts. *Academy of Management Review*, 22(4), 853–886. <https://doi.org/10.5465/amr.1997.9711022105>
- Narula, R., Rao, P., Kumar, S., & Matta, R. (2024). ESG scores and firm performance: Evidence from emerging market. *International Review of Economics & Finance*, 89(Part A), 1170–1184. <https://doi.org/10.1016/j.iref.2023.08.024>
- Nirino, N., Santoro, G., Miglietta, N., & Quaglia, R. (2021). Corporate controversies and company's financial performance: Exploring the moderating role of ESG practices. *Technological Forecasting and Social Change*, 162, Article 120341. <https://doi.org/10.1016/j.techfore.2020.120341>
- Purnanandam, A. (2008). Financial distress and corporate risk management: Theory and evidence. *Journal of Financial Economics*, 87(3), 706–739. <https://doi.org/10.1016/j.jfineco.2007.04.003>
- Rahmadini, S., & Hartanti, D. (2025). ESG, financial performance, and the moderating role of ESG controversies. *Jurnal Dinamika Akuntansi*, 17(1), 34–49. <https://doi.org/10.15294/jda.v17i1.17301>
- Rao, A., Dagar, V., Sohag, K., Dagher, L., & Tanin, T. I. (2023). Good for the planet, good for the wallet: The ESG impact on financial performance in India. *Finance Research Letters*, 56, Article 104093. <https://doi.org/10.1016/j.frl.2023.104093>
- Ross, S. A. (1977). The determination of financial structure: The incentive-signalling approach. *The Bell Journal of Economics*, 8(1), 23–40. <https://doi.org/10.2307/3003485>
- Shi, Y., Li, X., & Asal, M. (2023). Impact of sustainability on financial distress in the air transport industry: The moderating effect of Asia–Pacific. *Financial Innovation*, 9(1), Article 97. <https://doi.org/10.1186/s40854-023-00506-1>
- Sihotang, D. G., & Siregar, S. V. (2025). The effect of ESG performance on firm value and financial distress with ESG controversies as a moderating variable. *Petra International Journal of Business Studies*, 8(2), 257–270. <https://doi.org/10.9744/petrajbs.8.2.257-270>
- Song, Y., Li, R., Zhang, Z., & Sahut, J.-M. (2024). ESG performance and financial distress prediction of energy enterprises. *Finance Research Letters*, 65, Article 105546. <https://doi.org/10.1016/j.frl.2024.105546>
- Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics*, 87(3), 355–374. <https://doi.org/10.2307/1882010>
- Suchman, M. C. (1995). Managing legitimacy: Strategic and institutional approaches. *Academy of Management Review*, 20(3), 571–610. <https://doi.org/10.5465/amr.1995.9508080331>
- Sugiyono. (2019). *Metode penelitian kuantitatif, kualitatif, dan R&D* (2nd ed.). Alfabeta.
- Yadav, A., & Asongu, S. A. (2025). The role of ESG performance in moderating the impact of financial distress on company value: Evidence of wavelet-enhanced quantile regression

- with Indian companies. *Business Strategy and the Environment*, 34(3), 2782–2798.
<https://doi.org/10.1002/bse.4118>
- Zheng, C., Khan, M. A. M., Rahman, M. M., Moudud-Ul-Huq, S., & Rahman, M. S. (2025). ESG performance and financial stability: A bibliometric and meta-analysis. *SAGE Open*, 15(4), Article 21582440251381772. <https://doi.org/10.1177/21582440251381772>
- Zhou, F., Chen, H., Tang, S., Lee, C., & Lin, C.-T. (2025). Integrating ESG into financial distress models: The role of financial report quality and innovation in the Chinese market. *International Review of Economics & Finance*, 101, Article 104135. <https://doi.org/10.1016/j.iref.2025.104135>