# Firm Size, ESG and Investment Behaviour in the Financial Performance of Malaysian Listed Companies

**Irme Rafhan Ibrahim, Norzanah Mat Nor*, Shaliza Alwi**

**Arshad Ayub Graduate Business School, Universiti Teknologi MARA, Shah Alam, Malaysia**

***Corresponding Author**

**DOI: https://doi.org/10.47772/IJRISS.2026.100900943**

**Received: 08 October 2026; Accepted: 13 October 2026; Published: 27 October 2026**

## ABSTRACT

Larger companies may possess greater resources for sustainability implementation, yet organisational scale can also raise coordination costs and reduce the incremental information conveyed by environmental, social and governance (ESG) ratings. This study examines whether firm size conditions the relationship among ESG, investment behaviour and financial performance in Malaysian listed companies. It uses an unbalanced panel of non-financial constituents of the FTSE4Good Bursa Malaysia and FTSE4Good Bursa Malaysia Shariah indices (497 firm-years from 131 firms, 2021–2024), with financial data hand-collected from audited annual reports. Investment behaviour is measured as the signed deviation from an expected-investment benchmark, financial performance as return on equity (ROE), ESG as the FTSE Russell ESG grade and size as the logarithm of total assets. A two-way fixed-effects model with firm-clustered standard errors tests the investment deviation × ESG × size interaction, with conditional effects from the full covariance matrix and equivalence tests for the null. The three-way coefficient is −0.2231 ($p = 0.300$), so the hypothesis is not supported; equivalence to a negligible effect is not established either, so the result is inconclusive rather than evidence that size is irrelevant. The ESG–ROE slope does not differ significantly from zero at any observed firm size. The signed investment deviation is positively associated with ROE ($b = 0.7568$, $p = 0.014$), more strongly in larger firms, a pattern driven by investment above the benchmark. The conclusions hold with return on assets, absolute and signed-split deviations, an industry-year benchmark, an equally spaced ESG grade, random effects and a conditional logit; with lagged ESG the ESG × size term turns positive. The study separates the resource capacity of large firms from demonstrated financial returns, within the bounds of a short panel of ESG-screened firms.

**Keywords:** ESG, firm size, investment deviations, financial performance, Malaysian listed companies

## INTRODUCTION

Corporate sustainability places demands on organisational resources and on the allocation of capital. Companies must collect information, assess environmental and social risks, coordinate implementation and communicate performance to stakeholders. Larger firms may have specialist staff, formal reporting systems and financial capacity to undertake this work. However, greater organisational scale does not automatically imply a stronger accounting return from sustainability activities. Coordination across business units can be costly, established disclosure systems may reduce the additional information supplied by an ESG score, and sustainability investments can yield benefits beyond a short accounting horizon; Eccles et al. (2014), for example, observe the performance advantage of high-sustainability companies over an eighteen-year window.

These competing considerations make firm size more than a routine control variable. Size can potentially alter the relationship between sustainability assessments and shareholder returns. It might also condition whether ESG changes the financial association of investment behaviour. The latter is a more demanding proposition: it asks whether a moderation relationship itself differs with organisational scale. Distinguishing these questions matters because a significant ESG × size term cannot, by itself, establish a size contingency in the investment × ESG relationship.

The broader ESG–performance literature reports heterogeneous associations across measures and settings. Friede et al. (2015) synthesise a large empirical literature and find that non-negative relationships predominate, while also documenting substantial variation. Such aggregate evidence motivates closer examination of organisational conditions rather than a universal financial payoff assumption. Rating interpretation adds another complication. Drempetic et al. (2020) show that firm size and resources for producing ESG information are associated with ESG scores. Their findings concern a different rating system and do not demonstrate bias in the FTSE Russell measure used here, but they identify a relevant distinction between assessed sustainability and the capacity to document it. Berg et al. (2022) further show that ESG ratings differ across providers in scope, measurement and weighting.

Malaysia provides a useful setting for examining these conditional relationships. The Malaysian Code on Corporate Governance 2021 places sustainability within the board's governance responsibilities (Securities Commission Malaysia, 2021). The sampling frame comprises ESG-screened non-financial listed companies, allowing variation in FTSE Russell ESG grades to be examined within that selected population. This design addresses differences among screened firms; it does not compare index membership with non-membership or estimate the effect of governance reform.

The study examines whether firm size changes the ESG conditioning of the investment deviation–ROE relationship, and whether the simpler ESG–size interaction is distinguishable from that three-way relationship. Investment deviations, ESG and size are the focal variables; foreign sales and auditor type are controls. The three-way relationship is not statistically supported, equivalence tests cannot rule out an economically relevant size contingency, and the ESG–ROE slope shows no reliable size gradient; the clearest regularity is that investment above the benchmark accompanies higher ROE, particularly in larger firms.

The article is related to the first author's doctoral thesis (Ibrahim, 2026), which assembled the sample and examined a broader set of firm-specific governance factors. It differs from that work in four respects. First, it isolates a single question that the thesis treated as a subsidiary test: whether organisational scale conditions the ESG moderation of investment behaviour. Second, every financial variable is rebuilt from the audited financial statements in the companies' annual reports and reconciled against its definition before estimation; this reconciliation corrected errors in the database extract used earlier, including the omission of loss-making years from ROE, foreign-sales ratios above one and breaks in the scale of total assets. Third, the specification adds year fixed effects, and inference on conditional effects uses the full coefficient covariance matrix, with confidence bands and equivalence tests for the null hypothesis. Fourth, the results are examined against robustness tests that the thesis does not report: return on assets, absolute and signed-split investment deviations, an industry-year investment benchmark, an equally spaced ESG grade, lagged ESG, random effects and a conditional fixed-effects logit. Where estimates differ from those in the thesis, the estimates reported here supersede them.

The contribution is a disciplined test of the organisational-scale argument. It separates the capacity to implement or report sustainability from evidence that scale changes investment-related returns, and it keeps the higher-order hypothesis distinct from lower-order interactions that involve the same variables.

## LITERATURE REVIEW AND HYPOTHESIS

### Investment behaviour and agency costs

According to agency theory, decision-makers' interests may not always align with shareholders' interests, especially when decision-makers have discretion and information advantages (Jensen & Meckling, 1976). Free cash flow can support investment that expands managerial control without generating commensurate value (Jensen, 1986). Conversely, financing constraints or weak information environments can prevent worthwhile projects from proceeding. Investment discipline therefore concerns departures on both sides of an appropriate allocation benchmark.

Biddle et al. (2009) connect financial reporting quality with investment efficiency, supporting the relevance of information conditions for allocation decisions. Richardson (2006) examines over-investment of free cash flow,

and Cheng et al. (2014) show that stronger corporate social responsibility performance is associated with lower capital constraints, which links sustainability to the financing side of investment decisions. These contributions motivate a governance interpretation of investment behaviour, but they do not imply that any residual-based measure directly observes project quality. An expected-investment regression provides a statistical benchmark. Its residual can capture omitted opportunities, sector differences and measurement error as well as allocation problems.

Prior studies often label this residual "investment efficiency". This article uses "investment deviation" when interpreting the variable. A positive residual means investment exceeds the benchmark and a negative residual means it falls below it. A larger signed value does not necessarily represent greater efficiency. Moving from a strongly negative residual towards zero differs economically from moving away from zero into a positive residual, even though both movements increase the numerical value. A single linear coefficient averages across these movements, which is why the analysis also reports absolute and signed-split deviations.

**ESG and stakeholder accountability**

Stakeholder theory situates corporate decisions within relationships with employees, investors, customers, communities and other constituencies (Freeman, 1984; Donaldson & Preston, 1995). Sustainability accountability may change the information considered in investment appraisal, the risks management must address and the scrutiny applied to decisions. ESG could consequently condition the association between investment deviations and financial returns rather than operate only as a separate predictor.

Several mechanisms are plausible. Better information about environmental exposure may improve project selection. Stakeholder engagement may identify operational risks or strengthen implementation. External scrutiny may discourage investments justified primarily by managerial discretion. Fatemi et al. (2018) find that ESG strengths are associated with higher firm value and weaknesses with lower value, and that disclosure moderates these associations, which illustrates how the information content of ESG can condition financial outcomes. These possibilities explain why sustainability and capital allocation can be considered jointly. None is directly measured in the available panel, so the regressions cannot establish which mechanism operates.

ESG ratings also compress multiple dimensions into a provider-specific score. Berg et al. (2022) show that disagreement across ratings has several sources. Accordingly, the FTSE Russell ESG grade is interpreted as assessed ESG performance under that provider's framework, not as a complete measure of sustainability accountability. A positive interaction with the grade would be compatible with a governance mechanism, while remaining insufficient to demonstrate the mechanism itself.

**Organisational scale as a competing contingency**

A resource-capacity argument suggests that larger companies can fund specialist expertise, monitoring systems and implementation programmes. Such infrastructure could make ESG practices more relevant to investment appraisal and strengthen the connection between sustainability governance and financial outcomes. Under this argument, the ESG conditioning of the investment–performance relationship becomes more positive as size increases.

A competing argument concerns diminishing incremental information and organisational complexity. Large companies may already operate many controls and disclosures, so an ESG score adds less informational value. Additional levels of management and geographically dispersed operations can further dilute the conversion of information into timely investment decisions. These arguments predict attenuation rather than a stronger positive scale effect. Organisational research on corporate social responsibility points in the same direction: Baumann-Pauly et al. (2013) find that small firms are not necessarily less advanced than large firms in organising CSR and tend to embed it in core operations, whereas large firms are stronger in external communication, and Wickert et al. (2016) attribute large firms' relative emphasis on CSR "talk" over "walk" to the organisational costs of implementation. Drempetic et al. (2020) add a measurement concern: ratings can reflect the resources available for providing ESG information as well as underlying sustainability performance. This concern motivates caution rather than a claim that size-related rating distortion has been demonstrated in Malaysia.

The competing mechanisms justify a non-directional hypothesis. H1: Firm size moderates the ESG conditioning of the relationship between investment deviations and ROE. Its empirical test is the coefficient of investment deviation × ESG × size. The ESG × size term is a subsidiary conditional relationship, interpreted separately and not treated as a substitute test of H1. This distinction preserves the hierarchy of the model and prevents a lower-order result from being read as evidence on the central hypothesis.

# METHODOLOGY

## Design and sample

The study uses a quantitative, observational panel design with the firm-year as the unit of analysis. The sampling frame contains the unique non-financial constituents of the FTSE4Good Bursa Malaysia and FTSE4Good Bursa Malaysia Shariah indices during 2021–2024. After financial-sector exclusions, 133 firms were eligible. One firm, Advanced Synergy Berhad, was excluded because it was delisted and lacked complete data, leaving 132 firms and 528 potential firm-years. The estimation sample contains 497 firm-years from 131 firms after the exclusions described below, so the panel is unbalanced.

The sample spans eleven non-financial Bursa Malaysia sectors. It is a bounded ESG-index sample rather than a random sample of all Malaysian listed companies, and the firms that leave the estimation sample do so because of changes of financial year end, recent listing or missing lagged data rather than at random. The results therefore apply most directly to the screened companies represented here, not to financial institutions, private firms, small enterprises or listed firms outside the indices.

Financial-statement data were hand-collected from the audited consolidated financial statements in each company's annual reports for fiscal years 2019–2024, as filed with Bursa Malaysia. For every firm-year the following group figures were recorded in thousands of ringgit: profit attributable to owners of the parent, equity attributable to owners of the parent, total assets, revenue, purchases of property, plant and equipment from the cash-flow statement, and revenue by geographical location from the segment or revenue-disaggregation note. Revenue and profit were cross-checked against the quarterly results filed with the exchange, and differences larger than three per cent were traced to restatements, discontinued operations or column errors and resolved from the audited statements. A fiscal year is assigned to the calendar year in which it ends. Sector membership follows the Bursa Malaysia classification. Auditor type comes from the original study dataset and was checked against the independent auditors' reports of 44 firms; the 4 firms found to be misclassified were corrected (12 firm-years).

Each firm-year was checked against the variable definitions before estimation. Of the 528 potential firm-years, 5 have no annual report covering the year, because of a change of financial year end or because the firm had not yet listed, and 5 cover a reporting period other than twelve months, so their flows are not comparable with annual observations; these firm-years are excluded (2 of them also have negative equity, for which ROE is undefined). A further 21 firm-years lack an investment deviation because lagged sales growth requires two preceding twelve-month periods. Every remaining firm-year has positive equity attributable to owners, ROE retains losses as negative values, foreign revenue never exceeds total revenue, and each year-on-year change in log total assets larger than 0.7 was confirmed in the audited statements as a merger (Celcom–Digi in 2022) or rapid growth rather than a change of units.

## Variable measurement

Table 1 summarises the measures. ROE is profit attributable to owners of the parent divided by equity attributable to owners of the parent. It measures accounting profitability attributable to equity holders, not market valuation, environmental performance or social welfare, and it is sensitive to leverage and to small equity denominators. Return on assets (ROA), the same profit divided by total assets, is therefore used as an alternative outcome. Continuous variables are winsorised at the first and ninety-ninth percentiles.

Investment deviations are residuals from an expected-investment regression of capital expenditure (purchases of property, plant and equipment) scaled by lagged total assets on lagged sales growth, adapted from Biddle et al.

(2009). The baseline first stage is estimated by pooled ordinary least squares across the sample firm-years (intercept 0.0404; slope on lagged sales growth 0.0124), so the residual measures a departure from a common benchmark and has a mean of zero by construction. Pooling across sectors is a limitation because firms with similar sales growth may face different asset requirements and investment cycles. A robustness version therefore adds sector-by-year intercepts to the first stage, benchmarking each firm against sector peers in the same year; the sector-year cells are too small for separate slope estimation. Two further transformations separate the direction of the deviation: the absolute deviation, and a signed split into over-investment (positive residuals) and under-investment (the magnitude of negative residuals).

Firm size is the natural logarithm of total assets. Foreign sales are revenue from customers outside Malaysia divided by revenue with a disclosed geographical split, bounded between zero and one; firms stating that they operate wholly or predominantly in Malaysia without a breakdown are coded zero. 11 firms report revenue only by region (for example Europe or Asia Pacific) without separating Malaysia; for them foreign sales are set to a firm-constant value of zero, which the firm fixed effects absorb, and an indicator for these firms is added in the pooled logit. Auditor type equals one when the independent auditors' report is signed by a Big Four firm and zero otherwise. ESG is the annual FTSE Russell ESG rating as disclosed by Bursa Malaysia in four grading bands, from one star for the lowest quartile of rated companies to four stars for the highest. In the data the rating takes four values (1.20, 1.85, 3.05, 4.35), one per band; because the values are not equally spaced, an equally spaced rank from one to four is used as a robustness check. A banded rating varies within a firm only when the firm moves between bands, which 84 of the 131 firms do at least once. The one-year lag of the rating is also used in robustness tests. FTSE Russell does not publish the environmental, social and governance pillar scores underlying the grade, so pillar-level tests are not possible with public data. Investment deviations, ESG and size are mean-centred before their products are constructed. Centring makes reference values interpretable; it does not eliminate endogeneity or guarantee that collinearity is immaterial.

**Table 1. Measurement of variables**

| Variable | Definition |
|---|---|
| ROE | Profit attributable to owners of the parent ÷ equity attributable to owners of the parent |
| ROA | Profit attributable to owners of the parent ÷ total assets (robustness outcome) |
| Investment deviation I | Signed residual from a pooled regression of capital expenditure ÷ lagged total assets on lagged sales growth; mean-centred. Robustness: absolute value; over/under split; sector-year benchmark |
| ESG E | Annual FTSE Russell ESG rating in four Bursa Malaysia grading bands; mean-centred. Robustness: equally spaced rank (1–4); one-year lag |
| Size S | Natural logarithm of total assets; mean-centred |
| Foreign sales | Revenue from customers outside Malaysia ÷ revenue with a geographical split (0–1) |
| Auditor | 1 if the auditors' report is signed by a Big Four firm; 0 otherwise |

Note: All financial data are group figures from audited annual reports, in thousands of ringgit.

**Specification and conditional interpretation**

Let I, E and S denote mean-centred investment deviations, ESG and size. The baseline model is:

$$ROE_{it} = \alpha_i + \tau_t + \beta1\, I_{it} + \beta2\, E_{it} + \beta3\, S_{it} + \beta4\, I_{it} E_{it} + \beta5\, I_{it} S_{it} + \beta6\, E_{it} S_{it} + \beta7\, I_{it} E_{it} S_{it} + \gamma1\, FSALES_{it} + \gamma2\, AUDITOR_{it} + \varepsilon_{it}.$$

The firm effects $\alpha_i$ absorb stable observed and unobserved differences across companies, and the year effects $\tau_t$ absorb shocks common to all firms in a year, such as the post-pandemic recovery and commodity-price cycles. All main effects and constituent two-way products are included, so $\beta7$ tests the higher-order relationship net of those simpler terms. Standard errors are clustered by firm, which permits heteroscedasticity and arbitrary within-firm dependence; with 131 clusters, inference uses the small-sample adjustment of the within estimator and t-statistics with 130 degrees of freedom (Cameron & Miller, 2015; Petersen, 2009). A specification with firm effects only is reported for comparison.

The hypothesis concerns β7. The conditional investment slope is β1 + β4E + β5S + β7ES. Its change with ESG is β4 + β7S. Thus, a significant β6 does not test whether size changes ESG moderation of investment behaviour. Instead, the conditional ESG slope is β2 + β4I + β6S + β7IS. The size gradient of this slope is β6 + β7I, reducing to β6 only when investment deviation is at its sample mean (Aiken & West, 1991). Because these conditional effects are linear combinations of coefficients, their standard errors are computed exactly from the full coefficient covariance matrix, which is reported in Appendix Table A1. Conditional effects are plotted with 95% confidence bands across the observed range of firm size, and Johnson–Neyman values identify where they differ significantly from zero (Bauer & Curran, 2005; Brambor et al., 2006; Hainmueller et al., 2019).

The Breusch–Pagan Lagrange multiplier statistic is 237.56 ($p < 0.001$), rejecting pooled estimation in favour of a panel estimator. Because the conventional Hausman test is not valid under clustered errors, fixed and random effects are compared with its cluster-robust Mundlak form, which tests whether the firm means of the time-varying regressors are jointly zero (Hausman, 1978; Mundlak, 1978): $\chi^2(9) = 10.60$ ($p = 0.304$). The test does not reject the random-effects restriction, so a random-effects estimator would be consistent and more efficient if that restriction holds. Fixed effects are nevertheless retained as the baseline: the hypothesis concerns within-firm changes, and a non-rejection in a four-year panel is weak evidence that persistent firm characteristics such as business model and governance are unrelated to ESG grades, size and investment. Random-effects estimates are reported as a robustness check. Mean and maximum variance inflation factors are 1.31 and 1.77. These diagnostics inform estimator choice and inference; they do not resolve reverse causality or time-varying confounding.

A non-significant β7 cannot by itself show that the three-way relationship is negligible (Lakens, 2017). The null is therefore assessed with two one-sided tests (TOST) at $\alpha = 0.05$, which is equivalent to asking whether the 90% confidence interval for β7 lies inside $\pm\Delta$ (Lakens et al., 2018). The smallest effect size of interest $\Delta$ is set on a standardised scale: a three-way coefficient large enough that a joint one-standard-deviation increase in ESG and size changes the ROE response to a one-standard-deviation investment deviation by 0.10 standard deviations of ROE, that is $\Delta = 0.10 \cdot SD(ROE)/[SD(I) \cdot SD(E) \cdot SD(S)]$. Bounds of 0.05 and 0.20 standard deviations are reported for sensitivity, together with the smallest bound for which equivalence would be supported.

**Robustness and logit specifications**

Each robustness test changes one element of the baseline: ROA replaces ROE; the absolute deviation or the over- and under-investment split replaces the signed deviation; the industry-year benchmark replaces the pooled benchmark; an equally spaced rank of the ESG grade replaces the band values; the one-year lag of ESG replaces its contemporaneous value, with the contemporaneous model re-estimated on the same reduced sample for comparison; and random effects with year dummies replace the fixed effects. As a check on functional form, ROE is also dichotomised at the sample median (High_ROE). A pooled logit with year dummies and firm-clustered standard errors uses between- and within-firm variation. A conditional (firm fixed-effects) logit with year dummies and firm-clustered standard errors uses only within-firm changes and so matches the identifying variation of the linear model (Chamberlain, 1980); firms whose High_ROE status never changes contribute no information, and regressors that barely vary within firms are not identified.

Dichotomisation changes the estimand and discards information about differences in the magnitude of ROE. Logit interaction coefficients describe relationships on the log-odds scale, and interactions in predicted probabilities require separate calculation (Ai & Norton, 2003). Because the conditional logit does not estimate the firm effects, probability-scale marginal effects are not identified; the comparison with the linear model is therefore confined to the sign and significance of the log-odds coefficients.

# RESULTS

**Descriptive profile and data reconciliation**

Table 2 reports descriptive statistics after winsorisation. Every statistic was recomputed from the source fields and reconciled to the estimation sample of 497 firm-years. Mean ROE is 0.1152 (SD 0.1593), ranging from −0.2841 to 0.8689; 53 firm-years record a loss. Mean ROA is 0.0613. The investment deviation has a mean of −0.0001, close to zero as its construction implies, and a standard deviation of 0.0440; 192 firm-years lie above

the benchmark and 305 below it. Foreign sales average 0.3115 of revenue and lie between zero and one. Mean ESG is 3.3241 on the band values, and mean log total assets is 21.9206.

The reconciliation matters for interpretation. ROE computed from audited components preserves losses, so the lower tail of Table 2 reflects genuine loss-making firm-years rather than a transformed ratio. Bounding foreign sales by total revenue makes the coefficient interpretable as the effect of a change in revenue share, and screening log total assets for unit breaks matters because the fixed-effects estimator identifies the size terms entirely from within-firm changes in size. The upper tail of ROE is genuine as well: a few firms that distribute nearly all their earnings, such as a brewer and a mobile network operator, carry small equity bases, which is one reason ROA is examined as an alternative outcome.

**Table 2. Descriptive statistics after winsorisation**

| Variable | Mean | SD | Minimum | Maximum |
|---|---|---|---|---|
| ROE | 0.1152 | 0.1593 | −0.2841 | 0.8689 |
| ROA | 0.0613 | 0.0792 | −0.0947 | 0.4236 |
| Investment deviation | −0.0001 | 0.0440 | −0.0465 | 0.1818 |
| ESG | 3.3241 | 1.0550 | 1.2000 | 4.3500 |
| Size ln assets | 21.9206 | 1.5447 | 18.8594 | 25.1290 |
| Foreign sales | 0.3115 | 0.3527 | 0.0000 | 0.9954 |
| Big Four auditor | 0.6720 | 0.4699 | 0.0000 | 1.0000 |

Note: n = 497 firm-years; 131 firms. Continuous variables winsorised at the 1st and 99th percentiles; statistics are on the raw (uncentred) scale.

**Two-way fixed effects and the central hypothesis**

Table 3 reports the baseline two-way fixed-effects model. The three-way coefficient β7 is −0.2231 (SE 0.2145; t = −1.04; p = 0.300), so H1 is not supported at the 5 per cent level. Whether a non-significant coefficient can be read as evidence of a negligible relationship is examined directly by the equivalence tests reported below.

The ESG × size coefficient β6 is −0.0077 (SE 0.0066; p = 0.244). Because the model includes the three-way term, β6 is the size gradient of the ESG–ROE slope only at the mean investment deviation; the conditional-effects analysis below evaluates the slope across the size distribution. The investment × ESG coefficient (β4 = −0.4598; p = 0.184) is not significant at mean size either.

The investment terms carry the clearest signal. At mean ESG and mean size, the investment deviation coefficient β1 is 0.7568 (p = 0.014): a one-standard-deviation rise in the deviation (0.044) is associated with ROE higher by 0.033, or 0.21 standard deviations. The investment × size coefficient β5 is 0.4068 (p = 0.032), so the association strengthens with firm size. Splitting the deviation by sign locates the association: above-benchmark investment has a coefficient of 0.9635 (p = 0.020) and the magnitude of below-benchmark investment −0.0372 (p = 0.959), and the size interaction is concentrated in the above-benchmark component (0.5805, p = 0.009; below-benchmark 0.3591, p = 0.421).

Among the controls, foreign sales has a coefficient of 0.0127 (p = 0.883) and Big Four auditor −0.0615 (p = 0.114); under firm effects the auditor coefficient is identified only from the 3 firms that changed between Big Four and other auditors, and omitting it leaves β6 (−0.0072) and β7 (−0.2204) virtually unchanged. The within $R^2$ is 0.0843, and the clustered Wald test of all slopes gives F(12, 130) = 1.35 (p = 0.201), so the regressors jointly explain little of the within-firm variation in ROE. The year effects are jointly significant only at the 10 per cent level (F(3, 130) = 2.58; p = 0.057). With firm effects only, β6 is −0.0073 (p = 0.271) and β7 is −0.2006 (p = 0.352).

**Table 3. Two-way fixed effects with firm-clustered standard errors (dependent variable: ROE)**

| Term | Coefficient | Robust SE | t |
|---|---|---|---|
| Investment deviation I | 0.7568** | 0.3022 | 2.50 |

| | | | |
|---|---|---|---|
| ESG E | 0.0005 | 0.0079 | 0.07 |
| Size S | 0.0105 | 0.1006 | 0.10 |
| I × E | −0.4598 | 0.3442 | −1.34 |
| I × S | 0.4068** | 0.1876 | 2.17 |
| E × S | −0.0077 | 0.0066 | −1.17 |
| I × E × S | −0.2231 | 0.2145 | −1.04 |
| Foreign sales | 0.0127 | 0.0861 | 0.15 |
| Big Four auditor | −0.0615 | 0.0386 | −1.59 |
| Constant | 0.1772*** | 0.0442 | 4.01 |

Note: Firm and year fixed effects; standard errors clustered by firm (small-sample factor G/(G − 1) × (N − 1)/(N − K)); t-tests with 130 degrees of freedom. * p < 0.10; ** p < 0.05; *** p < 0.01. n = 497; firms = 131. Within $R^2$ = 0.0843. The coefficient covariance matrix is reported in Appendix Table A1.

**Conditional effects with confidence intervals**

Table 4 and Figure 1 report the conditional effects implied by the baseline model, with standard errors computed from the coefficient covariance matrix. At the mean investment deviation, the ESG–ROE slope is 0.0124 (95% CI −0.0084 to 0.0333) one standard deviation below mean size, 0.0005 (−0.0151 to 0.0162) at mean size and −0.0114 (−0.0409 to 0.0180) one standard deviation above it. The Johnson–Neyman analysis finds no value within the observed size range at which the slope differs significantly from zero. Panel (b) of Figure 1 plots the quantity that H1 concerns, the ESG moderation of the investment slope (β4 + β7S), across firm size; its 95% confidence band includes zero throughout the observed size range.

**Table 4. Conditional effects at mean investment deviation with 95% confidence intervals**

| Size reference | Centred size S | ESG slope (SE) | 95% CI | β4 + β7S (SE) |
|---|---|---|---|---|
| One SD below mean | −1.5447 | 0.0124 (0.0105) | −0.0084 to 0.0333 | −0.1152 (0.1318) |
| Sample mean | 0.0000 | 0.0005 (0.0079) | −0.0151 to 0.0162 | −0.4598 (0.3442) |
| One SD above mean | 1.5447 | −0.0114 (0.0149) | −0.0409 to 0.0180 | −0.8045 (0.6626) |

Note: ESG slope = β2 + β6S at I = 0; β4 + β7S is the ESG moderation of the investment slope. Standard errors from the clustered covariance matrix (Appendix Table A1).

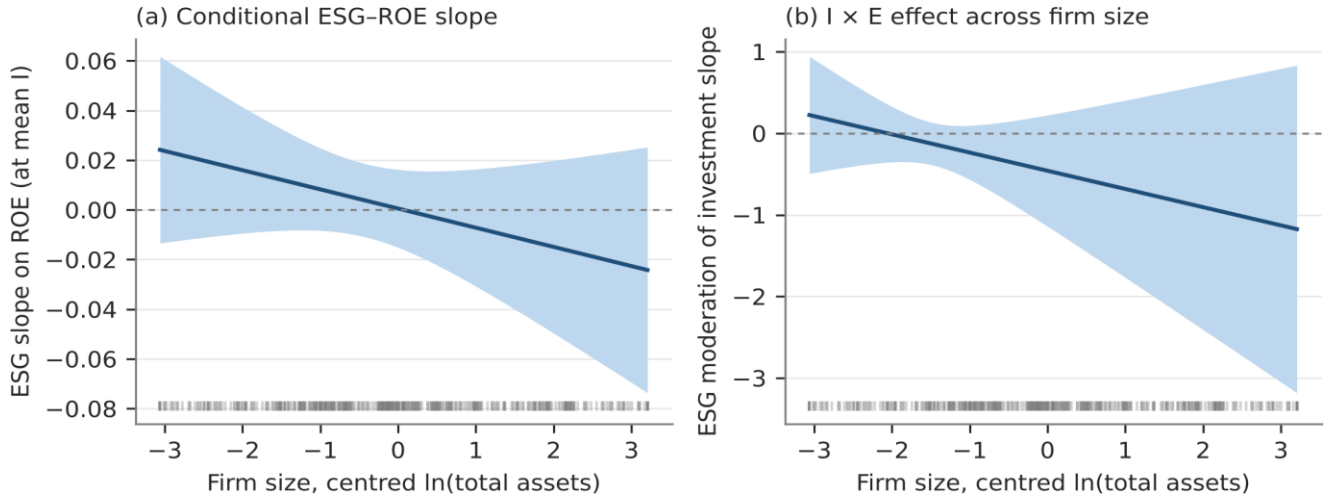



**Figure 1. Conditional effects across firm size (two-way fixed effects, 95% confidence bands)**

Note: Panel (a) plots ∂ROE/∂E = β2 + β6S at mean investment deviation; panel (b) plots $\partial^2$ROE/∂I∂E = β4 + β7S. Bands are 95% confidence intervals from the clustered covariance matrix; tick marks show the observed distribution of firm size.

**Equivalence tests for the three-way relationship**

Table 5 reports two one-sided tests for β7. With the smallest effect of interest set at 0.10 standard deviations of ROE, Δ = 0.2221 and the 90% confidence interval for β7 is −0.5784 to 0.1322, so equivalence is not established (p_TOST = 0.502). The smallest bound the data can support is 0.5784, which corresponds to 0.26 standard deviations of ROE. The non-significant three-way coefficient should therefore be read as inconclusive: the data cannot rule out a size-dependent ESG moderation of a magnitude that would be economically relevant.

**Table 5. Two one-sided equivalence tests for the three-way coefficient β7**

| Smallest effect of interest | Δ | 90% CI for β7 | p (TOST) | Equivalent |
|---|---|---|---|---|
| 0.05 SD of ROE | 0.1110 | −0.5784 to 0.1322 | 0.699 | No |
| 0.10 SD of ROE | 0.2221 | −0.5784 to 0.1322 | 0.502 | No |
| 0.20 SD of ROE | 0.4442 | −0.5784 to 0.1322 | 0.152 | No |

Note: Δ = k·SD(ROE)/[SD(I)·SD(E)·SD(S)] for k = 0.05, 0.10 and 0.20. Equivalence is declared when both one-sided tests reject at α = 0.05, i.e. when the 90% confidence interval lies inside ±Δ.

**Robustness**

Table 6 summarises the robustness tests; each row re-estimates the baseline model with one change. The three-way term is not significant at the 5 per cent level in any of the 11 specifications; the smallest p-value, 0.074, belongs to the under-investment component of the signed split. With the contemporaneous grade, the ESG × size coefficient is negative and insignificant in all 10 specifications. With the one-year lag of ESG, which restricts the sample to 381 firm-years in 2022–2024, it is 0.0107 (p = 0.008). The contemporaneous grade on that same sample gives −0.0045 (p = 0.412), so the difference reflects the timing of the rating rather than the shorter window.

**Table 6. Robustness of the focal coefficients (two-way fixed effects, firm-clustered SE)**

| Specification | N | E × S (SE) | p | Three-way term (SE) | p |
|---|---|---|---|---|---|
| Two-way FE baseline | 497 | −0.0077 (0.0066) | 0.244 | −0.2231 (0.2145) | 0.300 |
| Firm FE only | 497 | −0.0073 (0.0066) | 0.271 | −0.2006 (0.2149) | 0.352 |
| DV = ROA | 497 | −0.0016 (0.0032) | 0.630 | −0.0622 (0.1138) | 0.586 |
| \|Investment deviation\| | 497 | −0.0075 (0.0062) | 0.228 | −0.4287 (0.2836) | 0.133 |
| Signed split: over-investment | 497 | −0.0080 (0.0062) | 0.205 | −0.5022 (0.3246) | 0.124 |
| Signed split: under-investment | 497 | −0.0080 (0.0062) | 0.205 | −0.4889 (0.2710) | 0.074 |
| Industry-year investment benchmark | 497 | −0.0066 (0.0066) | 0.315 | −0.1209 (0.2256) | 0.593 |
| ESG grade as rank (1–4) | 497 | −0.0103 (0.0083) | 0.215 | −0.3042 (0.2726) | 0.267 |
| Lagged ESG (t−1) | 381 | 0.0107 (0.0040) | 0.008 | 0.2215 (0.1709) | 0.197 |
| Contemporaneous ESG, lagged-ESG sample | 381 | −0.0045 (0.0055) | 0.412 | −0.1948 (0.1743) | 0.266 |
| Random effects with year dummies | 497 | −0.0055 (0.0052) | 0.292 | −0.2311 (0.1989) | 0.247 |

Note: The three-way term is I × E × S, or (over- or under-investment) × E × S in the split model. Each specification includes all constituent terms, foreign sales, auditor type, and firm and year fixed effects, except the firm-FE-only row (no year effects) and the random-effects row (random firm effects, year dummies). The lagged-ESG rows cover 2022–2024. Standard errors clustered by firm; p-values from t(G − 1).

**Logit specifications**

Table 7 reports the logit models for High_ROE. In the pooled logit, the three-way coefficient is −1.754 (p = 0.435) and the ESG × size coefficient is 0.033 (p = 0.693). The conditional logit uses the 52 firms whose High_ROE status changes during the window (203 firm-years); in it the three-way coefficient is −0.952 (p = 0.829) and the ESG × size coefficient is 0.066 (p = 0.714). The investment deviation coefficient is positive in both models (pooled 18.661, p < 0.001; conditional 18.976, p = 0.014), matching the linear model. Log-odds

coefficients are not comparable in magnitude with the linear estimates, and the comparison is confined to sign and significance.

**Table 7. Logit models for above-median ROE (High_ROE)**

| Term | Pooled logit | SE | Conditional logit | SE |
|---|---|---|---|---|
| Investment deviation I | 18.661*** | 4.510 | 18.976** | 7.746 |
| ESG E | 0.069 | 0.141 | 0.245 | 0.290 |
| Size S | −0.295*** | 0.096 | 1.729 | 2.042 |
| I × E | −4.900 | 3.918 | −3.952 | 6.560 |
| I × S | 5.499** | 2.344 | 3.341 | 6.234 |
| E × S | 0.033 | 0.084 | 0.066 | 0.179 |
| I × E × S | −1.754 | 2.246 | −0.952 | 4.398 |
| Foreign sales | 0.743 | 0.493 | 2.642 | 3.481 |
| Big Four auditor | −0.040 | 0.349 | not identified | — |
| Region-only geography | −0.715 | 0.622 | absorbed | — |

Note: High_ROE = 1 when ROE exceeds the sample median. Both models include year dummies; standard errors clustered by firm. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$. Pooled logit: n = 497, pseudo $R^2 = 0.121$. Conditional logit: n = 203 firm-years from 52 firms with within-firm variation in High_ROE; auditor type is not identified because it changes within too few of these firms. The region-only geography indicator is firm-constant and is absorbed in the conditional logit.

# DISCUSSION

### Scale does not establish stronger investment-related ESG returns

The unsupported three-way hypothesis challenges a straightforward resource-capacity prediction. Within this panel and specification, size does not reliably change the ESG conditioning of the investment deviation–ROE relationship. Larger organisations may have more sustainability infrastructure, but the estimated product term does not demonstrate that this infrastructure produces a stronger accounting-return association for investment behaviour.

An offset between capacity and complexity is a possible explanation. Greater resources can coexist with coordination costs, pre-existing information systems and slower implementation. Another possibility is measurement: signed benchmark deviations and composite ESG ratings may not isolate the mechanisms through which sustainability influences project selection. The four-year window may also be too short to observe financial consequences of investment decisions. These explanations remain interpretations because no measures of implementation capacity, organisational complexity or project-level outcomes are included.

The finding adds a boundary to agency-based reasoning. Information and monitoring may matter for allocation, as suggested by Biddle et al. (2009) and Richardson (2006), without their joint association with ROE varying detectably by firm size. Benlemlih and Bitar (2018) report an association between greater CSR involvement and lower investment inefficiency in US firms. Their outcome concerns investment efficiency itself, whereas this study examines whether size conditions an ESG interaction in a profitability equation; the three-way result therefore does not directly contradict their finding. Nor does it show that size is irrelevant. A non-significant coefficient cannot establish the absence of an economically meaningful relationship (Lakens, 2017), and the equivalence tests leave open three-way effects as large as 0.26 standard deviations of ROE, beyond the 0.10 bound set as the smallest effect of interest. With 131 firms, four years and an ESG measure that changes only when a firm crosses a grading band, the design has limited power for a three-way interaction, and the hypothesis is better described as unresolved than as rejected.

### ESG and size: no reliable gradient

At mean investment deviation, the ESG–ROE slope is close to zero at every observed firm size, and the ESG × size coefficient is small and imprecise. The estimate obtained earlier from an unreconciled database extract

(Ibrahim, 2026) suggested a significantly negative ESG × size gradient, that is, weaker ESG–ROE associations in larger firms. That pattern does not survive the reconstruction of the variables from audited statements. With the rating lagged by one year the coefficient turns positive and significant. A gradient whose sign depends on data construction and on the timing of the rating cannot support a claim that ESG pays less, or more, in large firms.

The absence of a reliable gradient is compatible with offsetting forces. Larger firms may convert sustainability capacity into returns more readily, as the resource-capacity argument predicts, while a high grade conveys less additional information where firms already maintain extensive disclosure and control systems. Drempetic et al. (2020) interpret score–size relations in terms of reporting capacity, but their results cannot identify which force dominates here. The provider differences documented by Berg et al. (2022), and the coarseness of a four-band grade, underline the need to distinguish an association with a rating from a direct measure of sustainability effectiveness.

The absence of a significant positive ESG main effect is consequential. The model does not first establish an overall financial premium and then show its attenuation. Rather, it describes a conditional relationship centred on particular investment and size values. This distinction guards against converting conditional estimates into a universal business case for ESG. Even if no short-term accounting premium can be shown, stakeholder accountability is still relevant because ROE does not reflect environmental or social outcomes. Malaysian evidence also underscores the importance of specifying the organisational contingency. Mohammad and Wasiuzzaman (2021) examine ESG disclosure and firm performance with competitive advantage as a moderator. Their positive conditional evidence is not a direct comparison with the present size interaction: disclosure, provider-assessed ESG performance and competitive advantage represent different constructs. Read together, the studies support a context-dependent interpretation rather than a common financial payoff across all ESG measures and firm characteristics.

**Investment above the benchmark**

The most stable result concerns investment rather than ESG. Firm-years with capital expenditure above the sales-growth benchmark report higher ROE, and the association is stronger in larger firms; shortfalls below it show no association. The positive investment coefficient also appears in both logit models and under random effects. It does not imply that over-investment creates value. A benchmark based only on lagged sales growth omits current profitability and investment opportunities, so a positive residual partly identifies firms that invest because current returns are high (Biddle et al., 2009; Richardson, 2006), and larger firms may face fewer financing constraints in acting on such opportunities. The signed residual is therefore better read as an opportunity-related deviation than as a measure of inefficiency, which supports describing it as an investment deviation rather than investment efficiency.

**Implications for managers and investors**

Managers should evaluate sustainability-linked capital allocation using project-specific information rather than assuming organisational scale guarantees stronger returns. Investment appraisal can record expected financial outcomes, material sustainability risks, implementation responsibilities and subsequent performance. These are practical implications of the information and accountability framework, not interventions whose effectiveness is estimated in this study.

Investors should distinguish a company's ability to produce ESG documentation from evidence about operational and financial results. The findings encourage assessment of rating methodology and material exposures alongside company size. They do not justify favouring small or large firms on ESG grounds or using any interaction coefficient as a trading rule. The sample is screened, the estimates are conditional and the available evidence does not establish out-of-sample predictive performance. Khan et al. (2016) distinguish sustainability issues that are financially material within an industry from those that are not, and report different performance associations across those categories. This supports examining material exposures alongside a composite grade and motivates issue-level extensions of the present model, which would require pillar or issue scores that are not publicly available for these firms.

Foreign sales show no reliable association with ROE once firm effects are removed; the coefficient is identified only from within-firm changes in revenue mix, which are modest over four years. A revenue-proportion proxy also does not identify export capability, diversification quality or access to particular markets, so the variable serves as a control rather than a basis for policy conclusions.

**Implications for sustainability governance research**

Social science research can benefit from separating organisational capacity, reporting visibility and realised outcomes. Size is an imperfect proxy for capacity; an ESG score is an imperfect proxy for accountability; and ROE is a narrow financial outcome. Treating all three as complete measures of their underlying concepts risks overstating the empirical contribution. A stronger future design would observe resources dedicated to sustainability, internal decision processes and the timing of project outcomes directly.

The study also illustrates why model hierarchy and data verification matter. A lower-order interaction answers a different question from a three-way interaction, even when both involve ESG and size, and Brambor et al. (2006) emphasise retaining constituent terms and evaluating conditional marginal relationships when interpreting multiplicative models. Applied here, that principle requires distinguishing the ESG × size relationship at mean investment deviation from the investment × ESG × size hypothesis, and reporting uncertainty for combined slopes before making claims across the size distribution. The revision of the ESG × size estimate after the variables were rebuilt from audited statements adds a practical lesson: interaction terms in short panels are sensitive to unit and definition errors in database extracts, so reconciling variables with source documents is part of the inferential work rather than a clerical step.

**Limitations And Future Research**

The estimates are observational associations. Firm and year fixed effects remove stable firm differences and common annual shocks, but reverse causality and time-varying omitted variables remain; the positive investment association in particular may reflect profitable firms investing more. The cluster-robust Mundlak test did not reject random effects, and the random-effects estimates lead to the same conclusions, but neither estimator resolves endogeneity, and a four-year panel is too short for credible dynamic-panel instruments.

The sample is restricted to ESG-screened, non-financial listed firms, and firm-years are lost where financial year ends changed or lagged data were unavailable. Selection into the indices and continued observability can limit generalisability. A longer panel including non-members could distinguish within-firm changes, selection into screening and between-firm differences. Such an extension should use a design appropriate to its research question rather than assume that index membership is an exogenous treatment.

Measurement remains a constraint. The investment deviation is a residual from a parsimonious benchmark that omits other determinants of investment opportunities, and the industry-year version relies on small sector-year cells that permit sector-year intercepts but not sector-specific slopes. ESG is a four-band grade from a single provider: it changes within a firm only when the firm crosses a band boundary, its pillar scores are not public, and Berg et al. (2022) show that conclusions can differ across providers, so replication with continuous scores, pillar scores or alternative ratings would strengthen the evidence. Foreign sales cannot be measured for the 11 firms that report revenue only by region. ROE and ROA are accounting measures that do not capture environmental or social outcomes.

The equivalence conclusions depend on the chosen smallest effect of interest. The bounds are defined on a standardised scale for transparency, and readers with a different view of an economically meaningful effect can compare the reported 90% confidence interval with their own threshold. Independent validation would require new observations, such as an extension of the panel beyond 2024 or a comparison sample of non-index firms.

## CONCLUSION

This study examines organisational scale as a contingency in the relationship among ESG, investment deviations and financial performance in 131 Malaysian ESG-screened non-financial listed firms over 2021–2024, using

financial data rebuilt from audited annual reports. The investment deviation × ESG × size hypothesis is not supported ($\beta 7 = -0.2231$, $p = 0.300$), and equivalence tests cannot exclude a size contingency of up to 0.26 standard deviations of ROE, so the question remains open rather than closed. The ESG–ROE slope does not differ from zero at any observed firm size, and the ESG × size coefficient ($-0.0077$, $p = 0.244$) is not robust to the timing of the rating. Investment above the sales-growth benchmark is associated with higher ROE, more strongly in larger firms, a pattern best read as opportunity-driven investment rather than evidence of efficient allocation.

The contribution is a verified and fully reported test of the organisational-scale argument. Greater scale may supply sustainability resources, but the evidence does not show that it changes investment-related ESG returns. Assessments of sustainability and capital allocation should report the higher-order test together with its conditional effects and equivalence bounds, distinguish ratings from implementation and recognise the financial and population boundaries of the evidence.

## REFERENCES

1. Ai, C., & Norton, E. C. (2003). Interaction terms in logit and probit models. Economics Letters, 80(1), 123–129. https://doi.org/10.1016/S0165-1765(03)00032-6
2. Aiken, L. S., & West, S. G. (1991). Multiple regression: Testing and interpreting interactions. SAGE Publications.
3. Bauer, D. J., & Curran, P. J. (2005). Probing interactions in fixed and multilevel regression: Inferential and graphical techniques. Multivariate Behavioral Research, 40(3), 373–400. https://doi.org/10.1207/s15327906mbr4003_5
4. Baumann-Pauly, D., Wickert, C., Spence, L. J., & Scherer, A. G. (2013). Organizing corporate social responsibility in small and large firms: Size matters. Journal of Business Ethics, 115(4), 693–705. https://doi.org/10.1007/s10551-013-1827-7
5. Benlemlih, M., & Bitar, M. (2018). Corporate social responsibility and investment efficiency. Journal of Business Ethics, 148(3), 647–671. https://doi.org/10.1007/s10551-016-3020-2
6. Berg, F., Kölbel, J. F., & Rigobon, R. (2022). Aggregate confusion: The divergence of ESG ratings. Review of Finance, 26(6), 1315–1344. https://doi.org/10.1093/rof/rfac033
7. Biddle, G. C., Hilary, G., & Verdi, R. S. (2009). How does financial reporting quality relate to investment efficiency? Journal of Accounting and Economics, 48(2–3), 112–131. https://doi.org/10.1016/j.jacceco.2009.09.001
8. Brambor, T., Clark, W. R., & Golder, M. (2006). Understanding interaction models: Improving empirical analyses. Political Analysis, 14(1), 63–82. https://doi.org/10.1093/pan/mpi014
9. Cameron, A. C., & Miller, D. L. (2015). A practitioner's guide to cluster-robust inference. Journal of Human Resources, 50(2), 317–372. https://doi.org/10.3368/jhr.50.2.317
10. Chamberlain, G. (1980). Analysis of covariance with qualitative data. The Review of Economic Studies, 47(1), 225–238. https://doi.org/10.2307/2297110
11. Cheng, B., Ioannou, I., & Serafeim, G. (2014). Corporate social responsibility and access to finance. Strategic Management Journal, 35(1), 1–23. https://doi.org/10.1002/smj.2131
12. 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
13. Drempetic, S., Klein, C., & Zwergel, B. (2020). The influence of firm size on the ESG score: Corporate sustainability ratings under review. Journal of Business Ethics, 167(2), 333–360. https://doi.org/10.1007/s10551-019-04164-1
14. Eccles, R. G., Ioannou, I., & Serafeim, G. (2014). The impact of corporate sustainability on organizational processes and performance. Management Science, 60(11), 2835–2857. https://doi.org/10.1287/mnsc.2014.1984
15. Fatemi, A., Glaum, M., & Kaiser, S. (2018). ESG performance and firm value: The moderating role of disclosure. Global Finance Journal, 38, 45–64. https://doi.org/10.1016/j.gfj.2017.03.001
16. Freeman, R. E. (1984). Strategic management: A stakeholder approach. Pitman.

17. Friede, G., Busch, T., & Bassen, A. (2015). ESG and financial performance: Aggregated evidence from more than 2000 empirical studies. Journal of Sustainable Finance & Investment, 5(4), 210–233. https://doi.org/10.1080/20430795.2015.1118917
18. Hainmueller, J., Mummolo, J., & Xu, Y. (2019). How much should we trust estimates from multiplicative interaction models? Simple tools to improve empirical practice. Political Analysis, 27(2), 163–192. https://doi.org/10.1017/pan.2018.46
19. Hausman, J. A. (1978). Specification tests in econometrics. Econometrica, 46(6), 1251–1271. https://doi.org/10.2307/1913827
20. Ibrahim, I. R. (2026). The moderating role of ESG in the relationship between firm-specific factors and business performance of Malaysian public listed companies [Doctoral thesis]. Universiti Teknologi MARA.
21. Jensen, M. C. (1986). Agency costs of free cash flow, corporate finance, and takeovers. American Economic Review, 76(2), 323–329.
22. Jensen, M. C., & Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. Journal of Financial Economics, 3(4), 305–360. https://doi.org/10.1016/0304-405X(76)90026-X
23. Khan, M., Serafeim, G., & Yoon, A. (2016). Corporate sustainability: First evidence on materiality. The Accounting Review, 91(6), 1697–1724. https://doi.org/10.2308/accr-51383
24. Lakens, D. (2017). Equivalence tests: A practical primer for t tests, correlations, and meta-analyses. Social Psychological and Personality Science, 8(4), 355–362. https://doi.org/10.1177/1948550617697177
25. Lakens, D., Scheel, A. M., & Isager, P. M. (2018). Equivalence testing for psychological research: A tutorial. Advances in Methods and Practices in Psychological Science, 1(2), 259–269. https://doi.org/10.1177/2515245918770963
26. Mohammad, W. M. W., & Wasiuzzaman, S. (2021). Environmental, social and governance (ESG) disclosure, competitive advantage and performance of firms in Malaysia. Cleaner Environmental Systems, 2, 100015. https://doi.org/10.1016/j.cesys.2021.100015
27. Mundlak, Y. (1978). On the pooling of time series and cross section data. Econometrica, 46(1), 69–85. https://doi.org/10.2307/1913646
28. Petersen, M. A. (2009). Estimating standard errors in finance panel data sets: Comparing approaches. Review of Financial Studies, 22(1), 435–480. https://doi.org/10.1093/rfs/hhn053
29. Richardson, S. (2006). Over-investment of free cash flow. Review of Accounting Studies, 11(2–3), 159–189. https://doi.org/10.1007/s11142-006-9012-1
30. Securities Commission Malaysia. (2021). Malaysian Code on Corporate Governance 2021. Securities Commission Malaysia.
31. Wickert, C., Scherer, A. G., & Spence, L. J. (2016). Walking and talking corporate social responsibility: Implications of firm size and organizational cost. Journal of Management Studies, 53(7), 1169–1196. https://doi.org/10.1111/joms.12209.