Check for updates

Business Strategy and the Environment

WILEY | Business Strategy and the Environment

RESEARCH ARTICLE OPEN ACCESS

ESG Controversies and Corporate Performance: A Critical Appraisal of Rating Methodologies and Empirical Evidence

Paolo Candio[1] | Massimiliano Kaucic[2]

[1]Department of Economics and Management, University of Trento, Trento, Italy | [2]Department of Economics, Management, Mathematics and Statistical Sciences (DEAMS) "Bruno de Finetti", University of Trieste, Trieste, Italy

Correspondence: Paolo Candio (paolo.candio@unitn.it)

Received: 26 March 2026 | **Revised:** 17 July 2026 | **Accepted:** 4 August 2026

Keywords: corporate performance | corporate social responsibility | ESG controversies | firm value | rating methodologies

ABSTRACT

ESG controversies have become relevant signals in capital markets. However, controversy measurement is not strictly regulated, and evidence on whether and under which conditions such events affect firm-level performance remains fragmented and lacks quality appraisal. Addressing these concerns, this manuscript integrates a document-based comparative assessment of publicly disclosed controversy methodologies used by leading ESG data providers, with a systematic review of empirical research on ESG controversies and corporate performance. The document-based comparison shows that controversy is not a single stable construct. Rather, providers differ significantly in rated object, aggregation rules, materiality perspective and time dimension. Most empirical studies find negative market reactions and higher risk or financing costs following controversy identification. Quality appraisal indicates that the evidence base is generally of moderate quality but unevenly reported, limiting reproducibility and generalizability. Overall, the findings support treating ESG controversy as a consequential but methodologically heterogeneous signal.

1 | Introduction

Environmental, social and governance (ESG) controversies have gained increasing prominence in capital market dynamics (Kweh et al. 2025). ESG controversies can be defined as publicly reported events, incidents or allegations that indicate a firm's actual or perceived involvement in activities that conflict with environmental, social or governance norms and consequently attract stakeholder scrutiny. They differ from ESG risk, which concerns potential future exposure to ESG-related issues; from corporate misconduct, which generally implies proven wrongdoing; from reputational crises, which may arise as consequences of controversies; and from poor ESG performance, which may occur without a specific adverse event. ESG controversies are therefore best understood as public signals of actual or perceived ESG-related failures.

Rating agencies and specialised ESG data providers now frequently complement traditional ESG performance scores with controversy indicators intended to capture adverse incidents, such as environmental accidents, human rights abuses, labour disputes, corruption cases and allegations of greenwashing (OECD 2025; Xue et al. 2023). Such events may provoke immediate stakeholder responses, including investor divestment, activist pressure, consumer backlash and intensified regulatory scrutiny, while also producing longer-term consequences for firms' financing costs, litigation risk, attractiveness to talent, strategic discretion and, ultimately, corporate performance (Bel-Oms 2024).

ESG controversies are inherently heterogeneous. They can differ in severity, verifiability, duration, stakeholder reach and the speed with which information spreads through traditional and

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2026 The Author(s). *Business Strategy and the Environment* published by ERP Environment and John Wiley & Sons Ltd.

social media. For example, an environmental disaster will involve immediate physical damage, fatalities and prolonged reputational consequences, whereas a labour-related controversy will raise questions of supply-chain responsibility and generate a different pattern of stakeholder and regulatory response. Likewise, allegations of greenwashing or emissions manipulation, as in the Volkswagen diesel gate scandal (European Parliamentary Research Service 2016), may be highly salient to investors and regulators even where the underlying harm is mediated through issues of disclosure, compliance and public trust rather than a single visible accident. These differences may help explain why the implications of ESG controversies are unlikely to be uniform across firms or contexts (Muhammad and Tadele 2026). Consequently, empirical assessments of their effects on corporate performance depend heavily on how controversies are operationalised, the time horizon over which outcomes are examined and the extent to which studies address confounding factors that may shape both controversy exposure and firm performance.

A growing body of empirical research has examined the relationship between ESG controversies and corporate performance, yet important gaps remain in the synthesis of this literature. Existing reviews have primarily mapped the field and identified major research themes (Shakil 2024), but they have not systematically examined how ESG controversies are conceptualised and measured across data providers, nor have they formally appraised the quality of the empirical evidence. This omission is particularly important because controversy indicators are constructed using different methodological and conceptual assumptions, potentially affecting the comparability of findings across studies. Consequently, it remains unclear whether differences in reported results reflect genuine variation in the effects of ESG controversies or inconsistencies in how controversies are defined, measured and operationalised.

A further challenge concerns the unregulated measurement of ESG controversies themselves. Existing studies rely on a wide range of proxies, including provider-assigned controversy scores (Treepongkaruna et al. 2024), media-based indicators (Menicacci and Simoni 2024) and narrower issue-specific measures such as greenwashing (Xue et al. 2023). This diversity limits comparability across empirical settings and complicates attempts to synthesise findings. More broadly, the problem reflects the well-documented heterogeneity of ESG ratings, as providers often diverge in their scope, measurement, weighting and aggregation methodologies (Berg et al. 2022). Regulatory developments have also acknowledged these concerns. In the European Union, Regulation (EU) 2024/3005 on the transparency and integrity of ESG rating activities introduces authorisation, disclosure and governance requirements for ESG rating providers under ESMA supervision, with the aim of improving the transparency, comparability and integrity of ESG-related assessments in a market frequently criticised for opacity and ratings divergence.

Beyond synthesizing the literature, this manuscript contributes to the theoretical understanding of ESG controversies by arguing that they should not be treated as a single, homogeneous construct. Drawing on a comparative analysis of leading provider methodologies, the review demonstrates that controversy indicators embody different underlying logics, ranging from reputational risk assessments to double-materiality evaluations and ESG score-adjustment mechanisms. Consequently, empirical studies employing ESG controversy measures may not be examining the same underlying phenomenon, despite using similar terminology. This perspective helps explain inconsistencies in prior findings and highlights the need for greater conceptual precision in future theory development and empirical research.

To address these issues, the manuscript pursues two objectives. First, it undertakes a comparative assessment of the ESG controversy rating methodologies used by major agencies and data providers. Second, it conducts a systematic review of the empirical literature to critically appraise the evidence base, evaluate the methodological strengths and limitations of existing studies and identify priorities for future research.

2 | Methods

2.1 | ESG Controversy Methodologies

A document-based comparative design was used to assess the ESG controversy rating methodologies associated with major ESG data providers. The analytical framework adopted was derived from the analysis conducted by Berg et al. (2022). The comparative assessment was guided by the following six analytical dimensions: taxonomy, measurement, aggregation/weighting, rated object, materiality and time dimension. Taxonomy referred to the structure used to classify controversies; measurement to the criteria used to identify and assess controversy cases; aggregation/weighting to the procedures used to combine controversy information into an overall rating; rated object to the primary phenomenon being assessed; materiality to the principles used to determine the significance of controversies; and time dimension to the rules governing controversy persistence, updates and decay. For the rated object and materiality dimensions, classifications were based primarily on explicit descriptions contained in provider documentation. Where providers did not use identical terminology, classifications were derived from the stated purpose of the methodology, assessment criteria and scoring logic described in the documentation. The analysis therefore focused on documented methodological characteristics rather than inferred organisational intentions.

Those dimensions were retained for comparison, but the focus on controversy systems required extension of the framework. Therefore, an interpretive review of provider documentation was undertaken by two researchers to identify additional dimensions relevant to the comparison of ESG controversy methodologies (Bowen 2009). Where disagreements arose, these were resolved through discussion and, where necessary, adjudication by a third researcher. Accordingly, additional dimensions were added to capture: rated object (i.e., the primary phenomenon being assessed, such as reputational risk or public controversy), materiality (i.e., the principles used to judge which controversies matter) and time dimension (i.e., the rules governing how long controversy effects remain in scores). Provider documentation was independently reviewed by two researchers using a predefined comparison framework based on the six analytical dimensions examined in this

10990836, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/bse.71441 by Massimiliano Kaucic - Universita Di Trieste , Wiley Online Library on [20/08/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

study: taxonomy, measurement, aggregation/weighting, rated object, materiality and time dimension. Relevant information for each provider was extracted into a comparison table and systematically compared across providers. Where differences in interpretation arose, the reviewers discussed the available evidence and sought consensus. Any disagreements that could not be resolved through discussion were adjudicated by a third researcher. This process was designed to improve consistency, transparency and replicability in the comparative assessment of ESG controversy methodologies.

Consistent with Berg et al. (2022), four providers with publicly available and clearly defined methodologies for assessing ESG controversies were identified and evaluated: MSCI, Morningstar Sustainalytics, S&P Global and LSEG/Refinitiv. Provider selection was guided by two criteria: market relevance and the availability of sufficiently detailed public documentation describing the methodology used to identify, classify and aggregate ESG controversies. MSCI, Morningstar Sustainalytics, S&P Global and LSEG were selected because they are among the most widely used ESG data providers and publish methodological documentation that permits systematic comparison. Other relevant providers, including Moody's, ISS ESG and RepRisk, were considered. However, publicly available information on their controversy-specific methodologies was either more limited, less directly comparable or embedded within broader ESG assessment frameworks. Restricting the analysis to providers with transparent and sufficiently detailed methodological disclosures was therefore necessary to ensure consistency and replicability of the comparison.

2.2 | Systematic Review Design

2.2.1 | Data Sources and Search Strategy

The review focused only on empirical quantitative studies. The protocol for this was not registered and no ethics approval was required due to the nature of the research. All the details regarding the review protocol followed can be found in the following sections to ensure results reproducibility. A search strategy was developed based on a combination of keywords structured around the two concepts of ESG controversy and corporate performance following the published review protocol by Shakil (2024). A systematic search was conducted in Scopus and EBSCO (Supplementary information). Records published in the English language were included until 10th March 2026. Scopus and EBSCO were selected because they provide extensive coverage of peer-reviewed journals in business, management, finance, sustainability and related social science disciplines. The database selection was also informed by the review protocol proposed by Shakil (2024), thereby supporting methodological consistency with prior evidence syntheses in this field.

2.2.2 | Eligibility Criteria

Studies were eligible if they (a) were peer-reviewed empirical journal articles, (b) operationalised ESG controversies (including negative ESG events, corporate wrongdoing, corporate social irresponsibility, corporate misconduct, greenwashing or comparable constructs) and (c) reported at least one firm-level performance outcome (e.g., market-based performance, accounting performance, financing outcomes, risk or reputation). Although these constructs are not conceptually identical, they were included only where they represented adverse ESG-related events, allegations or behaviours consistent with the broader controversy concept adopted in this review. Differences in operationalisation were retained during data extraction and synthesis, allowing variation in controversy measurement to be considered when interpreting findings. Conceptual papers, reviews, editorials, studies without an explicit controversy construct and papers focusing solely on ESG performance/disclosure/ratings without a negative-event component or without a firm-level outcome were therefore excluded.

2.2.3 | Study Selection

Records were merged into a single screening database and deduplicated using DOI where available and otherwise via normalised title matching. Screening proceeded in two stages: (1) title-and-abstract screening and (2) full-text eligibility assessment. Both stages were conducted independently by the two reviewers. Where disagreements regarding study eligibility arose, these were resolved through discussion and, where necessary, adjudication by a third reviewer. The principal reasons for exclusion at the full-text stage were the absence of an explicit ESG controversy construct, the lack of a firm-level performance outcome, nonempirical study design or a primary focus on ESG performance, disclosure or ratings without a controversy component. A PRISMA-style flow diagram summarises the search, screening and inclusion process for the review.

2.2.4 | Data Extraction and Quality Assessment

For each included study, data were extracted on geographic and industry context, empirical design, ESG controversy measure(s), corporate performance outcome(s), sample size and unit of analysis, study period and reported direction of results. The direction of findings was classified as negative, mixed or conditional but overall adverse or null/insignificant. Extracted study characteristics are reported in Table A1. Study-level quality appraisal results for both the JBI and STROBE assessments are reported in Table B1.

Methodological quality was assessed using the JBI Critical Appraisal Checklist for Analytical Cross-Sectional Studies (Joanna Briggs Institute 2017). This checklist was used to evaluate clarity of inclusion criteria, description of the study setting and participants, validity and reliability of exposure and outcome measurement, use of objective measurement criteria, identification of confounding factors, strategies used to address confounding and appropriateness of the statistical analysis. Reporting quality was also assessed using the STROBE checklist for observational studies (von Elm et al. 2007), which was used to evaluate the completeness and transparency of reporting across the main sections of each article, with particular attention to study design, setting, participant selection, variables, data sources and measurement, bias, study size, statistical methods, handling of missing data, descriptive and

10990836, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/bse.71441 by Massimiliano Kaucic - Universita Di Trieste , Wiley Online Library on [20/08/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

outcome data, main results, limitations, generalisability and funding. For both tools, summary scores were calculated as the proportion of applicable items satisfied, with items rated as not applicable excluded from the denominator. Although both instruments were originally developed for observational research in health-related fields, they were considered appropriate for the purposes of this review because they assess broad aspects of methodological and reporting quality that remain relevant across empirical disciplines, including measurement validity, treatment of confounding factors, transparency of reporting and analytical appropriateness. The tools were therefore used to support structured comparison across studies rather than to evaluate discipline-specific econometric sophistication.

3 | Results

3.1 | ESG Controversy Methodologies

Across the agencies reviewed, ESG controversy ratings are designed to convert adverse events into a comparable company-level signal, but they differ in what they classify, how they assess severity and how they combine incidents into an overall rating. All major providers move beyond a simple list of incidents. MSCI (2024), Morningstar Sustainalytics (2025), S&P Global (2025) and LSEG (2022) each use formal taxonomies and explicit rules to translate case information into a rating outcome, but the methods remain heterogeneous (Table 1).

Morningstar Sustainalytics adopts the broadest analytical framework, separating impact from risk and combining harm to stakeholders or the environment with financially oriented exposure. Its model therefore captures both external effects and firm-level risk more clearly than the others. MSCI takes a case-based approach, evaluating each controversy through severity, company role and case status, with particular weight placed on the firm's involvement and on whether the issue remains unresolved. S&P Global follows a different route, opening controversies through tests of significance and responsibility and then converting them into modifiers that affect existing ESG criterion scores. LSEG relies on a more economical method, based mainly on weighted controversy counts across topic areas, with the resulting measure reducing the broader ESG score. Although all four agencies quantify controversy, they are not measuring exactly the same underlying construct.

Under MSCI and Sustainalytics, a worst-case logic means that the most severe case or indicator can determine the overall company assessment, making the rating highly sensitive to a single major event. S&P Global instead uses a cumulative multiplier system, allowing several controversies to combine through downward adjustments to affected criteria. LSEG is less centred on diagnosing individual events and instead integrates controversy information into the ESG Combined score through weighted interaction with the base ESG assessment. In this respect, MSCI and Sustainalytics treat controversy more as a signal with independent force, whereas S&P Global and LSEG position it more clearly as an adjustment within a broader ESG framework.

MSCI's methodology places particular emphasis on reputational and brand risk associated with alleged involvement in harmful activities, while remaining anchored in international norms. Sustainalytics adopts the widest scope, assessing both external impact and firm-level risk exposure. At S&P Global, the focus remains closer to financially relevant score effects, even though its impact assessment also refers to stakeholders and the environment. LSEG's methodology primarily incorporates controversy information through its interaction with ESG performance measurement, using controversies mainly as a modifier of the broader ESG assessment rather than as a standalone controversy rating.

Severe controversies can remain relevant for years in MSCI's system because active cases are monitored continuously. Sustainalytics links the initial incident to later updates and reviews cases periodically. S&P Global applies a formal depreciation schedule unless new information reactivates the case. LSEG mainly links controversies to the latest closed fiscal period, with some carry-over where legal or regulatory consequences continue. Depending on the provider, controversy may therefore be treated as a short-term event, an ongoing risk or a reputational issue that gradually loses force.

3.2 | Critical Appraisal of the Evidence

Following a PRISMA-guided public health review workflow, 4421 records were identified from Scopus and EBSCO, 1281 duplicates were removed, 3140 records were screened, 62 reports were sought for retrieval, 61 full-text reports were assessed for eligibility because one report could not be retrieved, and 33 studies were included in the qualitative synthesis (Figure 1).

The evidence base is mainly finance-oriented, with most studies focusing on listed firms rather than on a single industry. Cross-industry samples make up about three quarters of the reviewed studies, while banking and financial services represent the largest sector-specific group (Klepczarek et al. 2026; Dou et al. 2025; Gidage et al. 2025; Agnese et al. 2024; Mahyoub et al. 2024). The remaining studies examine more specific settings, such as grocery retail subsidiaries (Nuruzzaman et al. 2024), manufacturing (Yin and Xu 2025), technology firms (Horobet et al. 2025) and supply-chain–related firms (Kweh et al. 2025). Geographically, the literature includes both single-country and cross-country contexts, including India (Mendiratta 2024; Mendiratta et al. 2026), China (Ma and Ma 2025; Yin and Xu 2025), France (Brinette et al. 2024), South Africa (Nyakurukwa and Seetharam 2023), South Korea (Kim et al. 2025), MENA countries (Abu Khalaf et al. 2025), Europe (Klepczarek et al. 2026; Hampl and Vagnerova Linnertova 2024) and broader international samples (Gutknecht 2024; Frost et al. 2022; Gidage et al. 2025; De Vincentiis 2024). Sample sizes also differ widely, ranging from 20 South African listed firms (Nyakurukwa and Seetharam 2023) and 54 Asia-Pacific banks (Mahyoub et al. 2024) to 231 European nonfinancial firms (Hampl and Vagnerova Linnertova 2024), 1205 firm-year observations for French firms (Brinette et al. 2024), 5670 US firm-year observations (Mendiratta and Kumar 2025), 10,212 firms in emerging

10990836, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/bse.71441 by Massimiliano Kaucic - Universita Di Trieste , Wiley Online Library on [20/08/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 1 | Comparative summary of publicly disclosed ESG controversy methodologies.

Agency	Taxonomy	Measurement	Aggregation/ Weighting	Rated object	Materiality	Time dimension	Potential empirical implications
MSCI	3 pillars, 5 subpillars, 28 themes; controversy cases can be event-based or ongoing situations.	Case-level scoring based on severity, company role and status; uses public documents, media, NGOs and related sources.	Overall company score determined by the lowest-scoring case; hierarchy aggregates cases to themes, subpillars, pillars and company level.	Reputational and brand risk associated with alleged involvement in adverse-impact activities.	Primarily reputational risk, anchored in international norms and standards.	Active cases monitored continuously; severe cases may remain relevant for several years.	High sensitivity to severe individual controversies; major events may influence ratings for extended periods.
Morningstar Sustainalytics	10 Controversy Indicators derived from 41 event indicators across ESG topics.	News-triggered screening model assessing separate Impact and Risk dimensions.	Highest impact or risk determines event category; highest indicator determines company rating.	Stakeholder/ environmental impact and firm-level risk exposure.	Explicit double-materiality approach.	Ongoing updates and periodic reviews.	Strong sensitivity to highly material incidents; captures both impact and risk effects.
S&P Global	39 Media and stakeholder analysis thematic tags.	Cases initiated through responsibility, materiality and timing tests.	Rules-based multiplier matrix adjusts affected ESG criteria; multiple cases may compound.	Controversy information incorporated as an overlay to ESG assessment outcomes.	Primarily financially relevant materiality.	Formal depreciation schedule; cases may be reactivated.	Effects depend partly on interaction between controversy events and broader ESG performance.
LSEG/ Refinitiv	23 controversy topics covering environmental, social and governance issues.	Media-reported controversies counted and benchmarked against industry peers.	Weighted controversy scores combined with ESG scores to generate ESGC.	Controversy overlay used to modify ESG performance assessment.	Primarily investor-focused perspective.	Recent controversies reflected in latest fiscal period; effects may persist with legal/regulatory consequences.	Greater sensitivity to media visibility and recurring or persistent issues.

10990836, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/bse.71441 by Massimiliano Kaucic - Universita Di Trieste , Wiley Online Library on [20/08/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

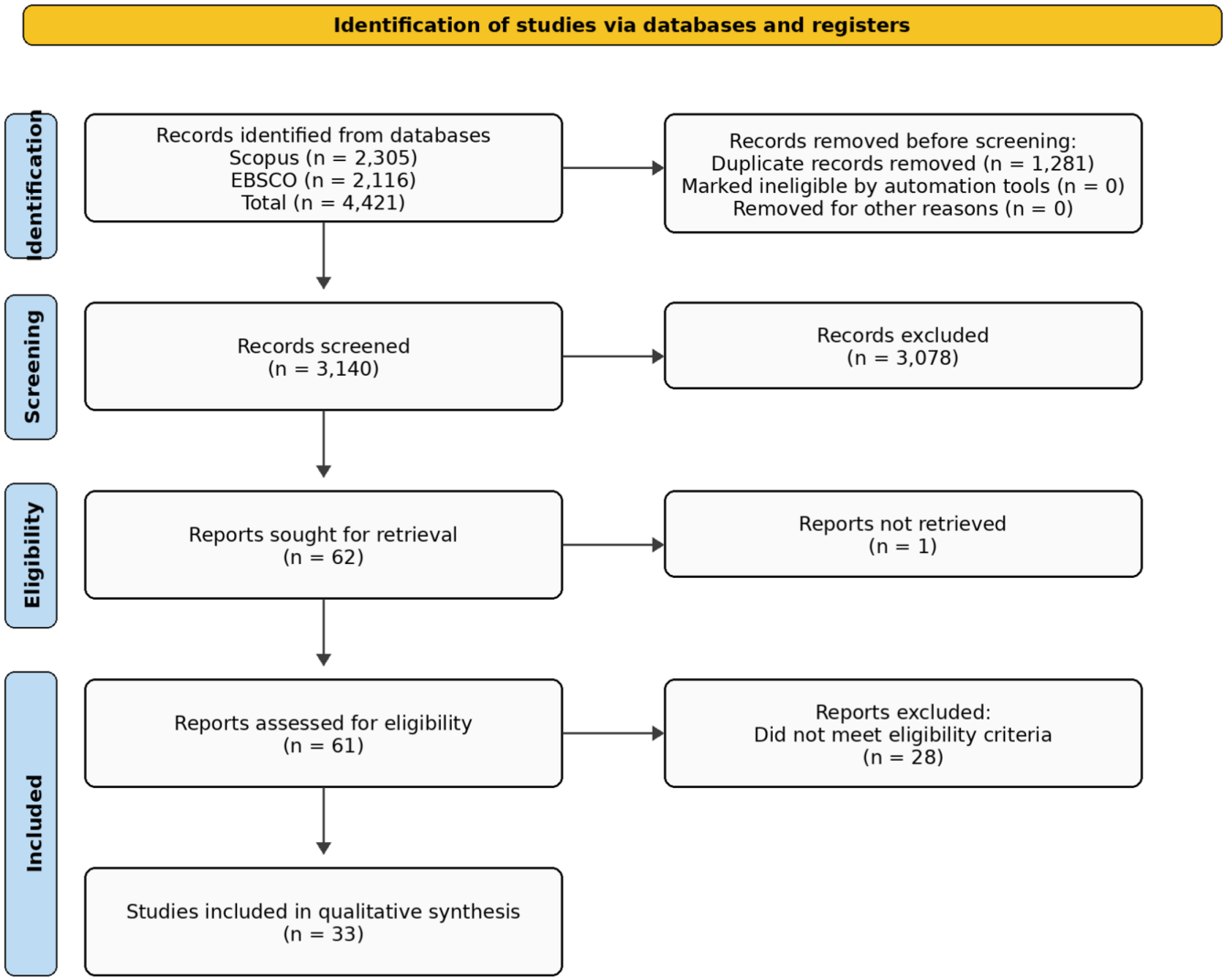

FIGURE 1 | PRISMA diagram.

markets (Brighi et al. 2025) and 15,039 firm-year observations in Chinese manufacturing (Yin and Xu 2025). Some studies also use event- or news-based units of analysis, most notably 111,251 news items covering 3202 firms across 20 countries (Gutknecht 2024).

In methodological terms, the literature is dominated by general empirical archival studies rather than clearly distinct quasi-experimental designs. Eighteen studies (54.5%) are best classified as empirical observational analyses without a more specific design label. Four studies (12.1%) use panel regression models (Brighi et al. 2025; Esposito De Falco et al. 2025; Yin and Xu 2025; Mahyoub et al. 2024), five (15.2%) apply event-study logic (Engelen et al. 2025; Gutknecht 2024; Kim et al. 2025; Nicolas et al. 2024; De Vincentiis 2024) and three (9.1%) use instrumental-variable or 2SLS approaches (Mendiratta et al. 2026; Treepongkaruna et al. 2024; Wu et al. 2023). One study each uses stochastic frontier analysis (Klepczarek et al. 2026), difference-in-differences (Dou et al. 2025) and GMM estimation (Agnese et al. 2024). Where time periods are reported, most studies cover multiyear windows of about 6–10 years, such as 2008–2015 (Mendiratta 2024), 2012–2020 (Nuruzzaman et al. 2024), 2012–2018 (Mendiratta and Kumar 2025), 2018–2023 (Yin and Xu 2025) and 2014–2021 (Horobet et al. 2025), although some extend much further, including 2002–2021 (Brighi et al. 2025) and 1995–2020 (Derrien et al. 2025).

ESG controversy is most often measured through media- or news-based negative ESG events or through vendor-based controversy scores, especially Refinitiv (Mendiratta 2024; Klepczarek et al. 2026; Abu Khalaf et al. 2025; Kim et al. 2025; Mendiratta and Kumar 2025; Hampl and Vagnerova Linnertova 2024; Nicolas et al. 2024). The most common outcome variables are firm value and valuation, followed by financing costs and accounting-based performance. Overall, the reported findings are more often negative than neutral. Clearly negative results, together with mixed but mainly adverse findings, make up most of the sample, especially in studies on firm value, profitability and financing costs (Nuruzzaman et al. 2024; Abu Khalaf et al. 2025; Brighi et al. 2025; Ma and Ma 2025; Yin and Xu 2025; Agnese et al. 2024; Hampl and Vagnerova Linnertova 2024; Mahyoub et al. 2024; Horobet et al. 2025; Derrien et al. 2025). In contrast, null or statistically insignificant findings are less common and are more likely to appear when the effects of controversy depend on moderators, event framing or estimation choices (Engelen et al. 2025; Mendiratta 2024; Mendiratta

10990836, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/bse.71441 by Massimiliano Kaucic - Universita Di Trieste , Wiley Online Library on [20/08/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

and Kumar 2025; Bang et al. 2023; Hanson and Sharpe 2025; Nyakurukwa and Seetharam 2023).

Mean JBI performance was high (84.1%), with a median of 87.5%, and 26 of the 33 reviewed studies scored at least 0.80. Detailed study-level appraisal results are provided in Table B1. The strongest areas relate to description of the study setting and sample, use of established outcome measures, identification of confounders, strategies to address confounding and the appropriateness of the statistical analysis. This pattern is visible in several stronger studies, including Nuruzzaman et al. (2024), Frost et al. (2022), Abu Khalaf et al. (2025), Brighi et al. (2025), Gidage et al. (2025), Yin and Xu (2025) and Mahyoub et al. (2024), which combine clearly defined measures with multivariate modelling and sensitivity analyses. The main methodological weakness concerns sample-definition criteria. Many papers rely on large archival datasets but do not fully specify their inclusion and exclusion rules, making JBI item 1 the least consistently satisfied criterion. Lower overall JBI scores in studies such as Fauser and Gruener (2020), Bang et al. (2023) and Dorfleitner and Zhang (2024) therefore appear to reflect either limited methodological transparency or a weaker fit between the JBI cross-sectional checklist and some finance-oriented empirical designs.

Mean STROBE evaluation was 77.5%, with a median of 82.8%, indicating weaknesses in reporting quality. Although most studies report their objectives, variables, main statistical methods and key findings reasonably well, fewer provide a clear justification of study size, explicit treatment of missing data, participant-flow information or a detailed discussion of generalisability and funding. In other words, the literature is often stronger in analytical execution than in documentation. Better reporting can be seen in studies such as Gutknecht (2024), Brighi et al. (2025), Gidage et al. (2025), Mendiratta and Kumar (2025) and Mahyoub et al. (2024), where the design, variables, estimation strategy and limitations are presented relatively clearly. Weaker reporting is more evident in studies such as Mendiratta (2024), Fauser and Gruener (2020), Bang et al. (2023) and Dorfleitner and Zhang (2024), where sample construction, treatment of missing data and the limits of external validity are less clearly reported. Overall, the appraisal suggests that this evidence base is supported by reasonably strong quantitative analysis but uneven reproducibility. Analytical decisions are often visible, but reporting on sample selection, attrition, missingness and the conditions under which findings can be generalised is less consistent. This distinction matters because it suggests that the main limitation of the literature is not simply weak evidence, but incomplete reporting of how the evidence was produced and how far it can be generalised.

4 | Discussion

This article critically assesses the rating methodologies used by major agencies and data providers to evaluate ESG controversies, alongside the quality of the empirical evidence on the effects of such controversies on corporate performance. The findings indicate that ESG controversy ratings are not directly comparable across providers, as they reflect different conceptualisation and measurement assumptions. This noncomparability can be summarised across the following three dimensions: the conceptual object being measured, the architecture used to identify and score controversies, as well as the aggregation and temporal logic through which controversies affect company-level assessments. Across providers, controversy ratings variously capture reputational exposure, stakeholder harm, financial materiality or adjustments to broader ESG performance scores. They also differ in their reliance on case-level severity assessments, event indicators, media-based counts, peer benchmarking or rules-based score adjustments. These differences mean that controversy ratings are not simply alternative measures of the same phenomenon, but reflect provider-specific assumptions about what controversies represent and how they should be weighted over time. At the same time, although the empirical literature generally points to an adverse association between controversies and corporate performance, evidence quality is weakened by suboptimal reporting standards and limited reproducibility, which hinder causal inference, cross-study comparison and consolidation of the evidence base.

The marked heterogeneity in ESG controversy methodologies documented here underscores the need for greater standardisation in data provision. Recent regulatory developments, including the EU regulation on ESG rating providers (European Union 2024), constitute an important step in this direction. However, further harmonisation is needed if controversy metrics are to become genuinely comparable, robust and reliable across agencies and providers. Pending such advances, this review offers a practical basis for selecting the ESG data providers most appropriate for empirical analysis. Provider choice should be aligned with the decision-making context, research objective, industry under examination and theoretical mechanism of interest. For example, studies of legitimacy loss, public scrutiny or reputational penalties may be best aligned with providers that emphasise media visibility and reputational risk. Research on stakeholder harm, social or environmental impacts or double materiality should prioritise methodologies that distinguish impact from firm-level risk. Studies of financial outcomes, investor responses or capital-market penalties should favour providers that incorporate financial materiality, regulatory exposure or firm-value implications. Finally, where the research question concerns how controversies affect broader ESG assessments, providers that integrate controversies as an adjustment to ESG scores may be most appropriate. Making this alignment explicit would improve construct validity, support clearer interpretation and reduce the risk of treating methodologically distinct controversy measures as interchangeable. Beyond the question of standardisation, there is a need for clearer theoretical and normative foundations regarding how controversies should be defined, classified and interpreted. This practical alignment also points to a broader theoretical issue: The field needs clearer foundations for defining, classifying and interpreting ESG controversies, while recognising their socially constructed and multidimensional nature (Bel-Oms 2024; Kweh et al. 2025).

A key theoretical implication emerges from the study findings. The divergence observed across ESG controversy methodologies suggests that controversy ratings do not merely differ in measurement but reflect different conceptualisations of what constitutes a controversy and why it matters. Building on this interpretation, ESG controversy measures can be understood as

10990836, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/bse.71441 by Massimiliano Kaucic - Universita Di Trieste , Wiley Online Library on [20/08/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

reflecting the following four broad measurement logics: reputational, stakeholder-impact, financial-materiality and ESG-adjustment logics. The reputational logic views controversies mainly as signals of public scrutiny, media attention and potential threats to corporate legitimacy. The stakeholder-impact logic treats controversies as evidence of harm, conflict or unmet expectations among affected stakeholder groups. The financial-materiality logic interprets controversies as indicators of ESG-related risks with potential implications for firm value, financing costs, operational performance or regulatory exposure. Finally, the ESG-adjustment logic uses controversies to qualify, penalise or contextualise broader assessments of firms' ESG performance. This typology reinforces the view that controversy ratings should not be treated as interchangeable indicators of a single construct. Future empirical studies should therefore align provider choice with the theoretical mechanism being examined.

The empirical evidence on ESG controversies would be strengthened by greater use of quasi-experimental designs, sector-specific analyses and replication across alternative data providers. Although most studies report adverse associations between controversies and corporate performance, the predominance of observational designs limits causal interpretation. Reverse causality, omitted-variable bias and other endogeneity concerns remain plausible, while instrumental variables, difference-in-differences designs and related approaches are still relatively uncommon. Accordingly, the literature currently provides stronger evidence of association than causation. Future research should also distinguish more clearly between short-term and longer-term effects. Event studies primarily capture immediate market reactions to controversy-related information, whereas panel and archival studies examine outcomes such as profitability, financing costs, efficiency and firm value over longer horizons. These outcomes may reflect different mechanisms and adjustment processes and should not be interpreted as evidence of a single controversy-performance relationship.

Industry context also requires greater attention. The materiality and consequences of controversies are likely to vary across sectors because firms face different stakeholder expectations, regulatory environments and ESG-related risks. Environmental controversies may be especially salient in resource-intensive industries, governance failures in highly regulated sectors and labour or supply-chain issues in manufacturing, retail and technology. Greater use of sector-specific or multilevel designs would help clarify when and where controversies are most consequential. Overall, future studies should examine whether capital markets penalise firms involved in ESG controversies, which types of controversies are most salient to investors and other stakeholders, and how quickly these signals are incorporated into share prices, firm value, financing costs and broader financial performance. Replication using measures from different ESG data providers would also help determine whether observed effects reflect underlying economic relationships or differences in data construction.

This study contributes to the literature on ESG controversies and, more broadly, on CSR by providing a systematic assessment and quality appraisal of an evidence base that remains fragmented, particularly with respect to reporting standards and reproducibility. The systematic search covered the major relevant databases and, although it may not have captured the full published literature, the evidence reviewed is representative of current research and timely given recent regulatory developments. However, the review is based primarily on narrative and documentary analyses, and future work could use more quantitative synthesis methods where sufficient comparability exists.

The narrative synthesis has limitations. Classifying findings as negative, mixed or null necessarily simplifies a diverse evidence base, and variation in measures, outcomes, designs, time horizons and reporting practices limited the feasibility of meta-analysis. As reporting becomes more standardised, future reviews may benefit from effect-direction plots and other structured synthesis approaches recommended for heterogeneous evidence bases (Konnyu et al. 2024), as well as meta-summary techniques, meta-analysis or meta-regression where appropriate. A further limitation concerns the quality appraisal tools employed. Although JBI and STROBE provide structured approaches for assessing methodological and reporting quality, they were not designed specifically for finance and management research and may not fully capture the strengths of econometric studies. The resulting scores should therefore be interpreted as indicators of methodological and reporting transparency rather than definitive measures of study quality. Finally, although the review covers four major providers with publicly documented methodologies, other relevant actors, including Moody's, ISS ESG and RepRisk, were not included because comparable information was not consistently available in the public domain. The findings should therefore be interpreted as representative of leading transparent controversy-rating frameworks rather than as an exhaustive assessment of all ESG controversy providers.

5 | Conclusions

ESG controversies represent a complex and multifaceted phenomenon, and significant challenges remain in understanding their implications for corporate performance and their signalling role for capital markets and other stakeholders. This review shows that ESG controversy cannot be regarded as a uniform construct. Instead, controversy measures are shaped by distinct methodological approaches and conceptual foundations, including reputational risk assessments, stakeholder-impact evaluations and ESG score-adjustment frameworks. Consequently, studies employing different controversy measures may capture different dimensions of the phenomenon despite using similar terminology. The evidence reviewed generally indicates a negative association between ESG controversies and outcomes such as firm value, financing costs and financial performance. However, the literature is characterised by substantial methodological variation, limitations in causal identification and inconsistencies in reporting quality. As such, the current evidence provides stronger support for association than for causation and should be interpreted accordingly.

These findings carry important implications for multiple stakeholder groups. Investors should carefully consider how controversy indicators are constructed before incorporating them into investment decisions. Regulators may wish to promote greater transparency and comparability in controversy

10990836, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/bse.71441 by Massimiliano Kaucic - Universita Di Trieste , Wiley Online Library on [20/08/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

assessment methodologies, particularly as ESG ratings continue to attract regulatory attention. For corporate managers, the findings highlight the potential financial and reputational costs of controversy events and the importance of robust governance, risk management and stakeholder engagement practices. For researchers, the review underscores the need for clearer conceptualisation, stronger causal research designs and closer examination of how measurement choices influence empirical findings.

Overall, advancing this field requires not only further empirical investigation but also greater conceptual clarity, transparency and consistency in the measurement of ESG controversies. Addressing these issues will enhance the ability of researchers, practitioners and policymakers to understand the conditions under which ESG controversies arise and the extent to which they affect corporate outcomes.

Author Contributions

Paolo Candio: conceptualization, investigation, writing – original draft, methodology, writing – review and editing, software, formal analysis, resources. **Massimiliano Kaucic:** conceptualization, methodology, validation, writing – review and editing.

Acknowledgements

Open access publishing facilitated by Universita degli Studi di Trento, as part of the Wiley - CRUI-CARE agreement.

References

Abu Khalaf, B., M. S. Alqahtani, and M. S. Al-Naimi. 2025. "ESG Controversies and the Financial Performance of MENA Firms: The Moderating Role of Board Characteristics." *Sustainability* 17, no. 11: 5055. https://doi.org/10.3390/su17115055.

Agnese, P., M. Cerciello, R. Oriani, and S. Taddeo. 2024. "ESG Controversies and Profitability in the European Banking Sector." *Finance Research Letters* 61: 105042. https://doi.org/10.1016/j.frl.2024.105042.

Bang, J., D. Ryu, and R. I. Webb. 2023. "ESG Controversy as a Potential Asset-Pricing Factor." *Finance Research Letters* 58: 104315. https://doi.org/10.1016/j.frl.2023.104315.

Bel-Oms, I. 2024. "Determinants and Consequences of Environmental, Social and Governance Controversies: Status Quo and Limitations of Recent Empirical Quantitative Research." *Corporate Social Responsibility and Environmental Management* 32: 2412–2437. https://doi.org/10.1002/csr.3071.

Berg, F., J. F. Kölbel, and R. Rigobon. 2022. "Aggregate Confusion: The Divergence of ESG Ratings." *Review of Finance* 26, no. 6: 1315–1344. https://doi.org/10.1093/rof/rfac033.

Bowen, G. A. 2009. "Document Analysis as a Qualitative Research Method." *Qualitative Research Journal* 9, no. 2: 27–40. https://doi.org/10.3316/QRJ0902027.

Brighi, P., A. C. F. Della Bina, and V. Venturelli. 2025. "Firm Value and Risk: How Relevant Are ESG Factors and ESG Controversies?" *Journal of Financial Reporting and Accounting*: 1–46. https://doi.org/10.1108/JFRA-12-2024-0953.

Brinette, S., F. D. Sonmez, and P. Sener. 2024. "ESG Controversies and Firm Value: Moderating Role of Board Gender Diversity and Board Independence." *IEEE Transactions on Engineering Management* 71: 1–12. https://doi.org/10.1109/TEM.2023.3236667.

De Vincentiis, P. 2024. "ESG News, Stock Volatility and Tactical Disclosure." *Research in International Business and Finance* 68: 102187. https://doi.org/10.1016/j.ribaf.2023.102187.

Derrien, F., P. Krüger, A. Landier, and T. Yao. 2025. "ESG News, Future Cash Flows, and Firm Value." *Journal of Finance* 80, no. 6: 3499–3554. https://doi.org/10.1111/jofi.13498.

Dorfleitner, G., and R. Zhang. 2024. "ESG News Sentiment and Stock Price Reactions: A Comprehensive Investigation via BERT." *Schmalenbach Journal of Business Research* 76: 197–244. https://doi.org/10.1007/s41471-024-00185-3.

Dou, J., N. Mirza, M. Umar, and A. Horobet. 2025. "ESG Uncertainties and Valuation Implications: Evidence From the EU Banking Sector." *Research in International Business and Finance* 76: 102872. https://doi.org/10.1016/j.ribaf.2025.102872.

Engelen, P.-J., L. Enneking, A. van Baar, and J. van Erp. 2025. "Are Firms Penalized for Their Involvement in Human Rights Violations?" *Corporate Social Responsibility and Environmental Management* 32: 8353–8368. https://doi.org/10.1002/csr.70139.

Esposito De Falco, S., G. Coniglio, E. Dalipi, and V. Giannetti. 2025. "The Moderating Role of the Legal Context Between ESG Controversies, Economic Performance and Board of Directors." *Corporate Social Responsibility and Environmental Management* 33: 3208–3224. https://doi.org/10.1002/csr.70316.

European Parliamentary Research Service. 2016. "The Volkswagen Emissions Scandal and Its Aftermath." European Parliament.

European Union. 2024. "Regulation (EU) 2024/3005 of the European Parliament and of the Council of 27 November 2024 on the Transparency and Integrity of Environmental, Social and Governance (ESG) Rating Activities, and Amending Regulations (EU) 2019/2088 and (EU) 2023/2859 (Text With EEA Relevance).". Official Journal of the European Union, L 2024/3005. http://data.europa.eu/eli/reg/2024/3005/.

Fauser, D. V., and A. Gruener. 2020. "Corporate Social Irresponsibility and Credit Risk Prediction: A Machine Learning Approach." *Credit and Capital Markets* 53, no. 4: 513–554. https://doi.org/10.3790/ccm.53.4.513.

Frost, T., L. Li, A. Tsang, and M. Yu. 2022. "Corporate Social Irresponsibility and Firm Value: International Evidence From Media Coverage." *Asia-Pacific Journal of Financial Studies* 51: 401–430. https://doi.org/10.1111/ajfs.12375.

Gidage, M., S. Bhide, A. Kolte, B. A. Gyamfi, and S. A. Asongu. 2025. "Greenwashing and Bank Financial Performance in Developing Markets: The Role of Boardroom Diversity and ESG Controversies." *Business Strategy & Development* 8: e70260. https://doi.org/10.1002/bsd2.70260.

Gutknecht, D. 2024. "Media coverage of corporate wrongdoing: International evidence on the stock market reaction and the buffering effect of prior corporate social performance." *Corporate Social Responsibility and Environmental Management* 31, no. 4: 2670–2698. https://doi.org/10.1002/csr.2700.

Hampl, F., and D. Vagnerova Linnertova. 2024. "How Do ESG Controversies Moderate the nexus Between ESG Performance and Cost of Capital? Evidence From European Listed Companies." *Managerial Finance* 50, no. 10: 1727–1746. https://doi.org/10.1108/MF-12-2023-0762.

Hanson, N., and S. Sharpe. 2025. "Can Sustainability (ESG) Controversies Be Offset With Advertising? An Empirical Investigation Into Advertising, Negative ESG, and Firm Value." *Journal of Risk and Financial Management* 18, no. 2: 86. https://doi.org/10.3390/jrfm18020086.

Horobet, A., A. Smedoiu-Popoviciu, R. Oprescu, L. Belascu, and A. Pentescu. 2025. "Seeing Through the Haze: Greenwashing and the Cost of Capital in Technology Firms." *Environment, Development and*

10990836, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/bse.71441 by Massimiliano Kaucic - Universita Di Trieste , Wiley Online Library on [20/08/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Sustainability 27: 21651–21682. https://doi.org/10.1007/s10668-024-04817-w.

Joanna Briggs Institute. 2017. "Checklist for Analytical Cross Sectional Studies." https://jbi.global/sites/default/files/2019-05/JBI_Critical_Appraisal-Checklist_for_Analytical_Cross_Sectional_Studies2017_0.pdf.

Jucá, M. N., P. Domadenik Muren, A. Valentinčič, and R. Ichev. 2024. "The Impact of ESG Controversies on the Financial Performance of Firms: An Analysis of Industry and Country Clusters." *Borsa Istanbul Review* 24, no. 6: 1305–1315. https://doi.org/10.1016/j.bir.2024.08.001.

Kim, H. J., T.-W. Ahn, and J. Yi. 2025. "ESG Controversies and Firm Value: Moderating Role of ESG Performance." *Asia-Pacific Journal of Financial Studies* 54: 228–269. https://doi.org/10.1111/ajfs.70001.

Klepczarek, E., A. Wieczorek, and W. Wojciechowski. 2026. "The Hidden Cost of Shame: ESG Controversies and Cost Efficiency in Banks – An SFA Approach." *Corporate Social Responsibility and Environmental Management*: 1–21. https://doi.org/10.1002/csr.70542.

Konnyu, K. J., J. M. Grimshaw, T. A. Trikalinos, N. M. Ivers, D. Moher, and I. J. Dahabreh. 2024. "Evidence Synthesis for Complex Interventions Using meta-Regression Models." *American Journal of Epidemiology* 193, no. 2: 323–338. https://doi.org/10.1093/aje/kwad184.

Kweh, Q. L., I. W. K. Ting, W.-M. Lu, J. Asif, and H. T. M. Le. 2025. "Environmental, Social, and Governance (ESG) and ESG Controversies: Opportunities or Challenges in Achieving Firm Efficiency?" *Natural Resources Forum* 50: 1163–1177. https://doi.org/10.1111/1477-8947.70019.

LSEG. 2022. "ESG scores methodology." https://www.lseg.com/content/dam/ftse-russell/en_us/documents/other/lseg-esg-scores-methodology.pdf.

Ma, S., and T. Ma. 2025. "ESG Controversies and Firm Investment Efficiency: Impact and Mechanism Examination." *Risks* 13, no. 4: 67. https://doi.org/10.3390/risks13040067.

Mahyoub, M., R. Ja'afar, and N. L. Ab Ghani. 2024. "ESG Controversies and Banking Performance: The Moderating Effect of Board Activity." *Asian Economic and Financial Review* 14, no. 12: 895–913. https://doi.org/10.55493/5002.v14i12.5232.

Mendiratta, A. 2024. "The Impact of Corporate Social Irresponsibility Media Coverage on Firm Performance." *Journal of Advances in Management Research* 21: 1–19. https://doi.org/10.1108/JAMR-08-2022-0170.

Mendiratta, A., and A. Kumar. 2025. "Does Industry Munificence Matter? ESG Controversies and Firm Value." *Journal of Advances in Management Research*. 22: 819–832. https://doi.org/10.1108/JAMR-08-2024-0326.

Mendiratta, A., S. Singh, S. S. Yadav, and A. Mahajan. 2026. "Relationship Between Firm Value, Environment, Social, and Governance Strengths and Controversies." *Global Finance Journal* 70: 101238. https://doi.org/10.1016/j.gfj.2026.101238.

Menicacci, L., and L. Simoni. 2024. "Negative Media Coverage of ESG Issues and Corporate Tax Avoidance." *Sustainability Accounting, Management and Policy Journal* 15, no. 7: 1–33. https://doi.org/10.1108/SAMPJ-01-2023-0024.

Morningstar Sustainalytics. 2025. "Controversy Ratings Methodology (Version 2.0)." https://www.sustainalytics.com/docs/knowledgehublibraries/default-document-library/morningstar-sustainalytics_-controversy-ratings_-version-2-0_-sep-2025.pdf?sfvrsn=a2f4f2_3MSCI ESGResearch (2024).

MSCI. 2024. "ESG controversies and global norms methodology." https://www.msci.com/documents/1296102/34424357/ESG-Research-Controversies-Methodology.pdf.

Muhammad, H., and H. Tadele. 2026. "Corporate Governance and ESG Controversies: A Quantile Regression Approach." *Corporate Governance* 26, no. 3: 596–615. https://doi.org/10.1108/CG-11-2024-0616.

Nicolas, M. L. D., A. Desroziers, F. Caccioli, and T. Aste. 2024. "ESG Reputation Risk Matters: An Event Study Based on Social Media Data." *Finance Research Letters* 59: 104712. https://doi.org/10.1016/j.frl.2023.104712.

Nuruzzaman, N., E. E. Makarius, D. Mukherjee, and A. Gaur. 2024. "Mncs' Corporate Social Irresponsibility and Foreign Subsidiary Performance." *Global Strategy Journal* 14: 509–541. https://doi.org/10.1002/gsj.1502.

Nyakurukwa, K., and Y. Seetharam. 2023. "Investor Reaction to ESG News Sentiment: Evidence From South Africa." *Economia* 24: 68–85. https://doi.org/10.1108/ECON-09-2022-0126.

OECD. 2025. *Behind ESG Ratings: Unpacking Sustainability Assessment*. OECD Publishing.

Pinto-Gutiérrez, C., P. Neudörfer, and S. Isaías. 2026. "ESG Controversies, Greenwashing, and Value-Destroying Acquisitions." *Journal of Sustainable Finance & Investment* 16, no. 2: 595–620. https://doi.org/10.1080/20430795.2025.2581647.

S&P Global. 2025. "S&P Global ESG scores methodology." https://portal.s1.spglobal.com/survey/documents/spglobal_esg_scores_methodology.pdf.

Shakil, M. H. 2024. "Environmental, Social and Governance Controversies: A Bibliometric Review and Research Agenda." *Finance Research Letters* 70: 106325. https://doi.org/10.1016/j.frl.2024.106325.

Treepongkaruna, S., K. Kyaw, and P. Jiraporn. 2024. "ESG Controversies, Corporate Governance, and the Market for Corporate Control." *Journal of Sustainable Finance & Investment* 14, no. 4: 815–842. https://doi.org/10.1080/20430795.2024.2334253.

von Elm, E., D. G. Altman, M. Egger, et al. 2007. "The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: Guidelines for Reporting Observational Studies." *Journal of Clinical Epidemiology* 61, no. 4: 344–349. https://doi.org/10.1016/j.jclinepi.2007.11.008.

Wu, Z., S. Lin, T. Chen, C. Luo, and H. Xu. 2023. "Does Effective Corporate Governance Mitigate the Negative Effect of ESG Controversies on Firm Value?" *Economic Analysis and Policy* 80: 1772–1793. https://doi.org/10.1016/j.eap.2023.11.018.

Xue, R., H. Wang, Y. Yang, M. K. Linnenluecke, K. Jin, and C. W. Cai. 2023. "The Adverse Impact of Corporate ESG Controversies on Sustainable Investment." *Journal of Cleaner Production* 427: 139237. https://doi.org/10.1016/j.jclepro.2023.139237.

Yin, J., and J. Xu. 2025. "ESG Controversies, Intellectual Capital and Firm Financial Performance: Evidence From Chinese Manufacturing Sector." *Journal of Intellectual Capital* 26: 1208–1232. https://doi.org/10.1108/JIC-10-2024-0324.

Appendix A: Supplementary Information

Database Search Strategies

Scopus

TITLE-ABS-KEY(

'environmental social and governance controvers*'

OR 'esg controvers*"

OR "environmental social and governance scandal'

OR 'esg scandal'

OR 'environmental social and governance reputation risk'

OR 'esg reputation risk'

OR 'corporate scandal*'

10990836, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/bse.71441 by Massimiliano Kaucic - Universita Di Trieste , Wiley Online Library on [20/08/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

OR ‘corporate misconduct*’

OR ‘corporate misdoing*’

OR ‘corporate fraud’

OR ‘corporate crime’

OR ‘corporate social irresponsibility’

OR ‘negative esg’

)

AND PUBYEAR < 2027

AND (LIMIT-TO (LANGUAGE, ‘English’))

AND (LIMIT-TO (SRCTYPE, ‘j’))

AND (LIMIT-TO (DOCTYPE, ‘ar’))

EBSCOhost

(TI ‘environmental social and governance controvers*’

OR AB ‘environmental social and governance controvers*’

OR TI ‘esg controvers*’

OR AB ‘esg controvers*’

OR TI ‘environmental social and governance scandal’

OR AB ‘environmental social and governance scandal’

OR TI ‘esg scandal’

OR AB ‘esg scandal’

OR TI ‘corporate scandal*’

OR AB ‘corporate scandal*’

OR TI ‘corporate misconduct*’

OR AB ‘corporate misconduct*’

OR TI ‘corporate misdoing*’

OR AB ‘corporate misdoing*’

OR TI ‘corporate fraud’

OR AB ‘corporate fraud’

OR TI ‘corporate crime’

OR AB ‘corporate crime’

OR TI ‘corporate social irresponsibility’

OR AB ‘corporate social irresponsibility’

OR TI ‘negative esg’

OR AB ‘negative esg’)

Filters:

- Published Date: January 2003–March 10, 2026.

- Source Type: Academic Journals.

- Document Type: Article.

- Language: English

AND PD(≥01/01/2003 AND ≤03/10/2026).

10990836, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/bse.71441 by Massimiliano Kaucic - Universita Di Trieste , Wiley Online Library on [20/08/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE A1 | Data extraction.

Author/s	Year	Region	Industry	Sample size/ units	Study horizon	Design type	Controversy measure	Corporate performance	Direction of results
Abu Khalaf et al. 2025	2025	MENA	Cross-industry MENA listed firms	Firms from 10 MENA countries (exact observation count not captured in extracted text)	2010–2023	Empirical (unspecified)	Refinitiv controversies/ scores	Accounting performance	Mixed/conditional (overall adverse)
Agnese et al. 2024	2024	Europe/EU	Banking/financial services	European banks (exact observation count not captured in extracted text)	2015–2022	GMM	Refinitiv controversies/ scores	Firm value/ valuation	Negative
Bang et al. 2023	2023	United States	Cross-industry listed firms	5421 firms	2002–2020	Empirical (unspecified)	Refinitiv controversies/ scores	Firm value/ valuation	Null/insignificant
Brighi et al. 2025	2025	Emerging markets	Cross-industry emerging-market listed firms	10,212 companies/ firms	2002–2021	Panel regression	ESG controversies (generic)	Firm value/ valuation	Mixed/conditional (overall adverse)
Brinette et al. 2024	2024	France	Cross-industry French listed firms	1205 firm-year observations; 119 French companies	2007–2021	Empirical (unspecified)	Media/news-based negative ESG events	Firm value/ valuation	Negative
De Vincentiis 2024	2024	Global/ multicountry	Blue-chip cross-industry listed firms	199 blue-chip listed firms	2015–2020	Event study	Media/news-based negative ESG events	Cost of capital	Mixed/conditional (overall adverse)
Derrien et al. 2025	2025	Global/ multicountry	Cross-industry listed firms	Analyst-firm ESG-news sample (exact observation count not captured in extracted text)	1995–2020	Empirical (unspecified)	Media/news-based negative ESG events	Cost of capital	Negative
Dou et al. 2025	2025	Europe/EU	Banking/financial services	European banks (exact observation count not captured in extracted text)	2018–2021	Difference-in-differences	Media/news-based negative ESG events	Cost of capital	Mixed/conditional (overall adverse)

(Continues)

10990836, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/bse.71441 by Massimiliano Kaucic - Universita Di Trieste , Wiley Online Library on [20/08/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE A1 | (Continued)

Author/s	Year	Region	Industry	Sample size/ units	Study horizon	Design type	Controversy measure	Corporate performance	Direction of results
Engelen et al. 2025	2025	Not stated	Cross-industry multinational firms/human-rights incidents	Event-level accusations of corporate human-rights violations (exact N not reported in extracted text)	Event-window design; underlying study years not explicit in extracted text	Event study	Media/news-based negative ESG events	Firm value/ valuation	Null/insignificant
Esposito De Falco et al. 2025	2025	United States	Cross-industry US firms	Exact sample size not captured in extracted text	2021–2023	Panel regression	Media/news-based negative ESG events	Firm value/ valuation	Mixed/conditional (overall adverse)
Fauser and Gruener 2020	2020	Not stated	Cross-industry corporate sample (sector not clearly stated)	Not retrievable from parsed PDF text	Not retrievable from parsed PDF text	Empirical (unspecified)	Media/news-based negative ESG events	Accounting performance	Mixed/conditional (overall adverse)
Frost et al. 2022	2022	Global/ multicountry	Cross-industry international firms	Firms from 43 countries (exact observation count not captured in extracted abstract)	Multi-year international archival window; exact years not captured in extracted text	Empirical (unspecified)	Refinitiv controversies/ scores	Firm value/ valuation	Mixed/conditional (overall adverse)
Gidage et al. 2025	2025	Emerging markets	Banking/financial services	110 listed banks across 18 developing countries	2012/2013–2021/2022	Empirical (unspecified)	Media/news-based negative ESG events	Accounting performance	Negative
Gutknecht 2024	2024	Global/ multicountry	Cross-industry international listed firms	111,251 news items covering 3202 firms in 20 countries	2007–2021	Event study	Media/news-based negative ESG events	Cost of capital	Mixed/conditional (overall adverse)
Hampl and Vagnerova Linnertova 2024	2024	Europe/EU	Nonfinancial European listed firms	231 European non-financial companies	Balanced panel; exact span not captured in extracted text	Empirical (unspecified)	Refinitiv controversies/ scores	Cost of capital	Negative

(Continues)

10990836, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/bse.71441 by Massimiliano Kaucic - Universita Di Trieste , Wiley Online Library on [20/08/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE A1 | (Continued)

Author/s	Year	Region	Industry	Sample size/ units	Study horizon	Design type	Controversy measure	Corporate performance	Direction of results
Hanson and Sharpe 2025	2025	Not stated	Cross-industry listed firms	Firms with at least one negative ESG event (exact observation count not captured in extracted text)	1995–2019	Empirical (unspecified)	ESG controversies (generic)	Firm value/ valuation	Null/insignificant
Horobet et al. 2025	2025	Global/ multicountry	Technology sector	Global technology-firm sample (exact observation count not captured in extracted text)	2014–2021	Empirical (unspecified)	Media/news-based negative ESG events	Cost of capital	Negative
Jucá et al. 2024	2024	Global/ multicountry	Cross-industry firms with industry/country clusters	7500 observations	Multi-year panel; exact years not captured in extracted text	Empirical (unspecified)	Refinitiv controversies/ scores	Efficiency/ productivity	Mixed/conditional (overall adverse)
Kim et al. 2025	2025	South Korea	Cross-industry Korean listed firms/event-study setting	Exact sample size not captured in extracted text	2011–2017	Event study	Refinitiv controversies/ scores	Firm value/ valuation	Mixed/conditional (overall adverse)
Klepczarek et al. 2026	2026	Europe/EU	Banking/financial services	Exact sample size not captured in extracted text	Multi-year European banking panel; exact years not captured in extracted text	Stochastic frontier (SFA)	Refinitiv controversies/ scores	Firm value/ valuation	Mixed/conditional (overall adverse)
Kweh et al. 2025	2025	Not stated	Supply-chain/ cross-industry firms	Exact sample size not captured in extracted text	2011–2020	Empirical (unspecified)	Refinitiv controversies/ scores	Cost of capital	Null/insignificant
Ma and Ma 2025	2025	China	Cross-industry Chinese listed firms	Exact sample size not captured in extracted text	Multi-year Chinese listed-firm panel; exact years not captured in extracted text	Empirical (unspecified)	Media/news-based negative ESG events	Firm value/ valuation	Mixed/conditional (overall adverse)
Mahyoub et al. 2024	2024	Asia-Pacific	Banking/financial services	54 listed banks	2013–2022	Panel regression	MSCI controversies/ ratings	Cost of debt/ financing	Negative

(Continues)

10990836, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/bse.71441 by Massimiliano Kaucic - Universita Di Trieste , Wiley Online Library on [20/08/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE A1 | (Continued)

Author/s	Year	Region	Industry	Sample size/ units	Study horizon	Design type	Controversy measure	Corporate performance	Direction of results
Mendiratta 2024	2024	India	Cross-industry Indian listed firms	1103 media counts; 693 firm-year observations	2008–2015	Empirical (unspecified)	Refinitiv controversies/ scores	Accounting performance	Null/insignificant
Mendiratta and Kumar 2025	2025	Not stated	Cross-industry US firms	5670 firm-year observations	2012–2018	Empirical (unspecified)	Refinitiv controversies/ scores	Firm value/ valuation	Mixed/conditional (overall adverse)
Mendiratta et al. 2026	2026	India	Cross-industry Indian listed firms	688 observations of Indian firms	2008–2018	Instrumental variables/2SLS	Media/news-based negative ESG events	Firm value/ valuation	Mixed/conditional (overall adverse)
Nicolas et al. 2024	2024	United States	Cross-industry listed firms/ social-media event sample	Exact sample size not captured in extracted text	2016–2022	Event study	Refinitiv controversies/ scores	Cost of capital	Negative
Nuruzzaman et al. 2024	2024	Not stated	Grocery retail/ multinational retail subsidiaries	335 subsidiaries of 42 multinational grocery retailers; 2185 subsidiary-year observations	2012–2020	Empirical (unspecified)	Media/news-based negative ESG events	Firm value/ valuation	Negative
Nyakurukwa and Seetharam 2023	2023	South Africa	Cross-industry South African listed firms	20 JSE-listed companies	2015–2021	Empirical (unspecified)	Media/news-based negative ESG events	Accounting performance	Null/insignificant
Pinto-Gutiérrez et al. 2026	2025	Not stated	Missing PDF/not appraised	Missing PDF/ not appraised	Missing PDF/not appraised				
Treepongkaruna et al. 2024	2024	Not stated	Cross-industry firms in the market for corporate control	Exact sample size not captured in extracted text	Multi-year archival window; exact years not captured in extracted text	Instrumental variables/2SLS	ESG controversies (generic)	Firm value/ valuation	Negative
Wu et al. 2023	2023	China	Cross-industry listed firms	Exact sample size not captured in extracted text	2018–2022	Instrumental variables/2SLS	ESG controversies (generic)	Cost of capital	Null/insignificant
Yin and Xu 2025	2025	China	Manufacturing (Chinese manufacturing sector)	3102 companies; 15,039 firm-year observations	2018–2023	Panel regression	ESG controversies (generic)	Firm value/ valuation	Negative

10990836, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/bse.71441 by Massimiliano Kaucic - Universita Di Trieste , Wiley Online Library on [20/08/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE B1 | Study-Level JBI and STROBE Scores.

Study	Design	JBI yes	JBI applicable	JBI score (%)	STROBE items reported	STROBE applicable	STROBE score (%)
Abu Khalaf et al. (2025)	Panel regression	7	8	87.5	24	29	82.8
Agnese et al. (2024)	Panel regression	7	8	87.5	24	29	82.8
Bang et al. (2023)	Other/unclear	5	8	62.5	19	29	65.5
Brighi et al. (2025)	Panel regression	8	8	100.0	26	29	89.7
Brinette et al. (2024)	Panel regression	7	8	87.5	22	29	75.9
De Vincentiis (2024)	Other/unclear	5	8	62.5	19	29	65.5
Derrien et al. (2025)	Other/unclear	7	8	87.5	20	29	69.0
Dorfleitner and Zhang (2024)	Event study	5	8	62.5	17	29	58.6
Dou et al. (2025)	Panel regression	7	8	87.5	22	29	75.9
Engelen et al. (2025)	Event study	7	8	87.5	24	29	82.8
Esposito De Falco et al. (2025)	Panel regression	7	8	87.5	22	29	75.9
Fauser and Gruener (2020)	Machine learning	5	8	62.5	17	29	58.6
Frost et al. (2022)	Other/unclear	7	8	87.5	24	29	82.8
Gidage et al. (2025)	Other/unclear	7	8	87.5	26	29	89.7
Gutknecht (2024)	Panel regression	7	8	87.5	26	29	89.7
Hampl and Vagnerova Linnertova (2024)	Panel regression	7	8	87.5	24	29	82.8
Hanson and Sharpe (2025)	Other/unclear	5	8	62.5	19	29	65.5
Horobet et al. (2025)	Other/unclear	5	8	62.5	22	29	75.9
Jucá et al. (2024)	Other/unclear	5	8	62.5	20	29	69.0
Kim et al. (2025)	Event study	7	8	87.5	24	29	82.8
Klepczarek et al. (2026)	Stochastic frontier analysis	7	8	87.5	22	29	75.9
Kweh et al. (2025)	Other/unclear	7	8	87.5	22	29	75.9
Ma and Ma (2025)	Panel regression	7	8	87.5	22	29	75.9
Mahyoub et al. (2024)	Panel regression	7	8	87.5	26	29	89.7
Mendiratta (2024)	Panel regression	7	8	87.5	17	29	58.6
Mendiratta and Kumar (2025)	Other/unclear	7	8	87.5	26	29	89.7
Mendiratta et al. (2026)	Causal inference (IV/PSM/etc.)	8	8	100.0	24	29	82.8
Nicolas et al. (2024)	Event study	7	8	87.5	24	29	82.8
Nuruzzaman et al. (2024)	Other/unclear	8	8	100.0	24	29	82.8
Nyakurukwa and Seetharam (2023)	Event study	7	8	87.5	22	29	75.9
Treepongkaruna et al. (2024)	Causal inference (IV/PSM/etc.)	8	8	100.0	24	29	82.8
Wu et al. (2023)	Causal inference (IV/PSM/etc.)	8	8	100.0	24	29	82.8
Yin and Xu (2025)	Panel regression	7	8	87.5	24	29	82.8

10990836, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/bse.71441 by Massimiliano Kaucic - Universita Di Trieste , Wiley Online Library on [20/08/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License