

ORIGINAL ARTICLE OPEN ACCESS

The Effects of Country Governance Quality on Corporate Sustainability and Ethical Behaviour

Paulo Morais Francisco^{1,2} 

¹School of Economics, University of Algarve, Faro, Portugal | ²Center for Advanced Studies in Management and Economics (CEFAGE-UA)lg

Correspondence: Paulo Morais Francisco (pmfrancisco@ualg.pt)

Received: 22 December 2025 | **Revised:** 22 December 2025 | **Accepted:** 4 May 2026

Funding: Fundação para a Ciência e a Tecnologia, Grant/Award Number: UIDB/04007/2020

Keywords: business ethics | controversies | corporate sustainability | country governance | ESG performance

ABSTRACT

Using institutional theory, we examine how country governance affects two ESG outcomes: ESG performance and ESG controversies. With Refinitiv/LSEG data for ~146,000 firm-years in 86 countries (2002–2023) and World Bank WGI, we apply a Mundlak within/between decomposition to test complementarity versus substitution (performance) and prevention versus detection (controversies). Better governance is strongly associated with higher ESG performance—overall and across E, S and G—especially cross-country. Yet governance also predicts more reported controversies, consistent with detection/visibility rather than worse conduct. Instrumental Variable and DiD tests corroborate these results.

JEL Classification: D02, M14, Q56, C23

1 | Introduction

The debate around corporate sustainability increasingly hinges on how we measure responsible behaviour. Under the broad label ‘environmental, social and governance (ESG)’, two conceptually distinct constructs are often conflated. On the one hand, ESG performance refers to programmatic policies, practices and disclosed outcomes across ESG pillars. On the other hand, ESG controversies capture discrete, event-based incidents such as environmental accidents, labour disputes, product recalls, bribery scandals or governance breakdowns. Refinitiv/LSEG’s framework makes this distinction explicit by providing both a performance-based ESG Score and a separate ESG Controversies metric derived from event screening of public sources. The coexistence of these two measures means a firm may report strong sustainability programmes yet still face reputationally and financially damaging incidents—or conversely record few controversies despite limited sustainability disclosure. For management and governance scholars, this bifurcation raises a core question: which features of a country’s institutional

governance shape both systematic ESG performance and the incidence or visibility of ESG controversies?

Institutions shape how firms behave: companies are embedded in national systems that set incentives and constraints (Campbell 2007; Aguilera and Jackson 2010; Hall and Soskice 2001). Within these systems, corporate social responsibility (CSR) takes different forms. In liberal market economies, CSR is typically explicit—voluntary programmes and disclosures—whereas in coordinated market economies it is more implicit, embedded in laws and collective institutions (Matten and Moon 2008; Jackson and Apostolakou 2010), though many countries exhibit hybrids of both. Cross-country evidence confirms that corporate social performance varies systematically with political and social institutions (Ioannou and Serafeim 2012). Accordingly, the quality of national governance is an important determinant of firms’ ESG performance.

To capture governance quality in a way that is globally comparable and time-varying, we use the World Bank’s worldwide governance indicators (WGI). The WGI summarise six dimensions—voice and

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2026 The Author(s). *European Financial Management* published by John Wiley & Sons Ltd.

accountability (VA), political stability (PV), government effectiveness (GE), regulatory quality (RQ), rule of law (RL) and control of corruption (CC) (Kaufmann et al. 2011). The indicators aggregate information from over 30 data sources (comprising hundreds of underlying variables) to measure how power is exercised, how rules are enforced, and how citizens participate in governance. Their cross-national coverage and long-time span (1996–present) make them well-suited to linking institutional context to firm-level ESG outcomes.

In environments where public institutions are capable, coherent and predictable—captured by the WGI—two contrasting forces can operate on firms' ESG performance. A complementarity view holds that strong institutions lower compliance frictions and raise the expected costs of poor conduct, thereby fostering the adoption and persistence of sustainability practices. A substitution view, by contrast, suggests that when stakeholder protections are already embedded in formal institutions, firms face less pressure to pursue explicit, voluntary ESG programmes and disclosures; as a result, measured ESG performance—which often emphasises disclosed policies and programmes—may appear lower even when underlying conduct is responsible (Matten and Moon 2008; Jackson and Apostolakou 2010).

A parallel tension arises for ESG controversies, used here as a proxy for unethical behaviour. On the one hand, robust institutional governance should prevent incidents by increasing sanction certainty, clarifying expectations and supporting routine oversight—consistent with the deterrence-and-compliance logic implicit in RL and CC (Kaufmann et al. 2011). On the other hand, those same institutions can heighten detection and visibility: stronger legal processes, credible audits and whistleblowing, administrative capacity, mandated transparency and active media and civil society improve the discovery, verification and publicisation of misconduct. Empirical work shows that improved monitoring and freer information flows can increase reported wrongdoing even when underlying behaviour does not worsen (Dyck et al. 2010; Olken 2007; Reinikka and Svensson 2004).

These mechanisms have direct implications for corporate governance and management research. Negative ESG incidents typically trigger sharper and more immediate valuation penalties than positive sustainability announcements generate gains (Krüger 2015; Flammer 2013), underscoring the financial relevance of controversies as outcomes distinct from ESG performance. If institutional governance reforms both reduce underlying incidents and enhance transparency, paradoxical patterns may emerge in the data: fewer serious failures over time yet more recorded controversies in high-governance settings with strong detection capacity; greater adoption of sustainability practices where enforcement is strong yet lower measured ESG scores where institutionalised protections substitute for explicit, voluntary programmes.

Within this framework, this study addresses three research gaps. First, while prior work links national institutions to corporate responsibility, few studies use pillar-level ESG performance (E, S, G) as granular proxies for corporate sustainability rather than broad CSR indices; we do so on a large global panel. Second, research rarely treats ESG controversies as outcomes distinct from performance with their own institutional

drivers; we model controversies alongside performance to separate sustainability from unethical behaviour. Third, the literature seldom disentangles mechanisms: we formalise and test complementarity versus substitution in the governance–performance link and prevention versus detection in the governance–controversy link.

Accordingly, we develop four testable hypotheses structured into two conceptual relationships. First, regarding ESG performance, if national institutions complement firm responsibility by deterring misconduct and encouraging compliance, higher institutional quality should be associated with stronger ESG outcomes. Conversely, if institutions substitute for firm-level initiatives, higher governance may correspond to weaker measured performance. Empirically, we find that overall institutional quality is positively associated with ESG performance, consistent with the complementarity view. Second, regarding ESG controversies, if deterrence reduces the probability of misconduct, stronger institutional environments should reduce the incidence of controversies; however, if greater transparency increases the likelihood of detection, higher institutional quality may instead be associated with more reported controversies, even when true misconduct is lower. Our evidence supports the detection mechanism: transparency- and scrutiny-related dimensions of governance are associated with more reported controversies. Taken together, these findings suggest that institutional quality tends to complement firm-level sustainability practices while simultaneously heightening the visibility of misconduct.

This study makes three contributions. First, we integrate institutional theory with two distinct ESG outcomes—performance and controversies—measured consistently using Refinitiv data. Second, we articulate and test a dual-channel framework within the WGI architecture, explaining why countries with stronger institutions can exhibit higher ESG performance yet also record more controversies. Third, methodologically, we employ a Mundlak (within/between) decomposition that separates within-country changes in governance from between-country differences, clarifying whether the observed associations reflect temporal improvements or persistent cross-national gaps. This approach retains the advantages of fixed effects by controlling for time-invariant unobserved heterogeneity while also enabling informative cross-sectional inference, addressing concerns about “thin” institutional indicators in cross-national research (Jackson and Deeg 2008).

The rest of the paper is organised as follows. Section 2 reviews the literature and develops the hypotheses. Section 3 describes the data, variables and empirical strategy, including the Mundlak decomposition. Section 4 presents the main results and robustness analyses. Section 5 concludes and discusses limitations and policy implications.

2 | Literature Review and Hypotheses

2.1 | Literature Review

Research on ESG issues has expanded rapidly over the past two decades, reflecting growing pressures from stakeholders, regulators and investors (e.g., Kotsantonis et al. 2016; Huang

et al. 2025). A central insight from this literature is that ESG is not a unitary construct. ESG performance captures systematic policies, programmes and disclosures that firms adopt to demonstrate ongoing responsibility, while ESG controversies represent event-based breakdowns such as environmental accidents, corruption scandals, labour disputes or governance failures (Chatterji et al. 2016; Eccles and Strohle 2018). The two measures are only moderately correlated and capture complementary dimensions of corporate responsibility: performance reflects organisational investments in sustainability infrastructure, while controversies capture acute failures of compliance or risk management (Sullivan and Mackenzie 2006). A comprehensive understanding of corporate responsibility therefore requires analysing both performance and controversies.

Institutional theory has long provided the dominant framework for explaining why corporate responsibility varies across national settings. Campbell (2007) argued that firms are more likely to act responsibly when they operate in environments characterised by credible enforcement, effective regulation and active monitoring by civil society. Matten and Moon (2008) advanced this insight by distinguishing between ‘explicit’ and ‘implicit’ CSR. In liberal market economies such as the United States, CSR tends to be explicit, relying on voluntary firm-led initiatives. In coordinated market economies such as Germany, CSR is often implicit, embedded in regulation, collective agreements and long-standing institutional arrangements. Related perspectives, such as Whitley (1999) national business systems and Hall and Soskice’s (2001) varieties of capitalism, likewise emphasise how institutional complementarities across labour markets, legal regimes and financial systems shape firms’ social and environmental practices. Extending this view, CSR can be understood as institutionally embedded private governance that varies systematically with national systems (Brammer et al. 2012).

Empirical evidence supports the institutional account. Ioannou and Serafeim (2012) demonstrate that corporate social performance systematically varies with national institutions, including political systems, education and civil liberties. Jackson and Apostolakou (2010) show that CSR practices in Europe reflect differences in institutional configurations rather than convergence towards a global model. Liang and Renneboog (2017) link CSR to legal origin, showing that firms in civil-law systems with stronger stakeholder protections adopt broader responsibility practices, while firms in common-law systems emphasise shareholder value. This evidence demonstrates that institutions condition corporate responsibility beyond firm-level determinants such as ownership, governance and financial characteristics (Aguilera et al. 2007; Khan et al. 2016; Crifo and Forget 2015).

To study governance quality systematically, scholars increasingly use the World Bank’s WGI. The WGI measure six governance dimensions: CC, GE, PV, RL, RQ and VA (Kaufmann et al. 2011). These indicators provide a standardised, cross-national framework for capturing how governments are selected, how authority is exercised and how institutions constrain abuse of power. Prior research has shown that these dimensions matter for corporate outcomes. Strong RL and CC

are associated with higher-quality financial reporting and reduced earnings management (Leuz et al. 2003; Houqe and Monem 2016). GE and RQ are associated with more efficient capital allocation and a lower cost of capital (Easley and O’hara 2004; Djankov et al. 2008). Investor protection and enforcement likewise discipline firms and reduce expropriation (La Porta et al. 1998; Shleifer and Vishny 1997). In the ESG domain, stronger governance environments promote environmental disclosure (Meng et al. 2014), encourage social engagement (Wang et al. 2016) and mitigate corruption-related risks (Cuervo-Cazurra 2006). Together, this evidence supports a deterrence and compliance mechanism: high-quality governance increases the costs of misconduct and encourages firms to invest in responsibility.

However, institutional perspectives also highlight that governance may not always complement firm-level ESG practices. Matten and Moon’s (2008) distinction between explicit and implicit CSR implies a potential substitution effect: in contexts where social and environmental protections are mandated by law, firms may do less in terms of voluntary, explicit programmes. Jackson and Apostolakou (2010) show that firms in coordinated economies with stronger institutionalised protections report fewer discretionary CSR initiatives. For ESG performance scores that emphasise explicit programmes and disclosure, this means that firms in high-governance countries may paradoxically appear to perform less well than their counterparts in weaker institutional environments, even if their actual conduct is highly responsible. This substitution effect complicates the relationship between governance quality and ESG performance, suggesting that both positive and negative associations are theoretically plausible.

The distinction between ESG performance and ESG controversies is especially important for assessing how governance quality matters. While ESG performance captures organisational strategies, controversies capture discrete events that can destroy value. Kr ger (2015) shows that negative responsibility events trigger much stronger stock price declines than positive CSR announcements create gains. Flammer (2013) similarly finds that negative CSR news has larger effects on firm value than positive news. Studies of environmental accidents and product recalls confirm that controversies lead to significant abnormal losses and litigation risks (Karpoff et al. 2005; Amiram et al. 2018). These findings underscore that controversies are economically meaningful outcomes in their own right and cannot be inferred directly from ESG performance measures.

Country governance shapes controversies through both prevention and transparency. On the prevention side, the same four WGI dimensions that support ESG performance—RL, CC, GE and RQ—reduce the likelihood that misconduct escalates into controversies by increasing sanction certainty, improving regulatory consistency and limiting opportunities for evasion. On the transparency side, detection is not solely about civic freedoms: RL enables legal discovery and due process; CC supports audit integrity and credible whistleblowing; GE raises administrative capacity to investigate and share information; RQ mandates clearer reporting and incident notification; VA strengthens media/nongovernmental organisation (NGO) scrutiny and PV

sustains routine inspections and data collection. Empirical work shows that stronger monitoring and freer information flows raise reported wrongdoing even when underlying behaviour does not worsen (Dyck et al. 2010; Olken 2007; Reinikka and Svensson 2004; Besley and Prat 2006; Ferraz and Finan 2008). Applied to ESG, countries with stronger governance may appear to have more controversies simply because incidents are more readily detected, verified and disclosed, leading data providers to record them more systematically (Healy and Palepu 2001; Healy and Serafeim 2020).

A persistent methodological issue in ESG research is the divergence of ratings across providers. Berg et al. (2022) document substantial differences in scope, measurement and aggregation across vendors, producing only moderate correlations. For controversies, providers differ in how they define, classify and score incidents. To minimise this noise, researchers increasingly rely on a single provider when analysing both performance and controversies. Refinitiv is particularly suitable because it computes ESG performance scores from disclosed practices while separately tracking controversies based on event screening. Using a single source ensures consistency and reduces confounding differences in measurement scope. Nevertheless, even within a single provider, controversies remain partly shaped by institutional visibility. Countries with strong media and civil society actors will naturally generate more controversial records for comparable levels of misconduct than countries where information is tightly controlled. This underscores the importance of treating ESG controversies as the joint outcome of both misconduct incidence and detection capacity.

In summary, the literature establishes that country-level governance is central to ESG outcomes but leaves two tensions unresolved. First, while governance may enhance ESG performance through complementarity, measured performance may appear lower when responsibilities are institutionalised at the national level (Matten and Moon 2008; Jackson and Apostolakou 2010). Second, controversies may decline through prevention yet increase in recorded terms due to stronger detection and visibility. By relying on a single-provider dataset (Refinitiv), we minimise ratings divergence and consistently test two questions: whether governance complements or substitutes for firm-level ESG efforts, and whether it reduces the true incidence of ESG controversies or increases their measured incidence through enhanced detection and reporting.

2.2 | Hypotheses

Institutional theory emphasises that firms are embedded in national systems of rules and norms that shape both incentives and constraints (North 1990; Campbell 2007). Corporate sustainability is therefore not solely the outcome of managerial discretion or stakeholder pressure; it is also a product of the institutional environment. Country-level governance institutions determine whether firms face credible enforcement of regulations, whether corruption undermines compliance and whether civic actors can monitor and publicise misconduct. These mechanisms imply that governance quality has systematic effects on two distinct dimensions of corporate

responsibility: ESG performance, reflecting programmatic investments and disclosures, and ESG controversies, reflecting event-based failures and incidents. The World Bank's WGI—RL, CC, GE, RQ, VA and PV—provide a multidimensional framework for governance quality (Kaufmann et al. 2011).

Building on these dimensions, we link governance to corporate responsibility outcomes through theoretically grounded channels. The first channel is a complementarity mechanism. Governance complementarity arises when strong institutions amplify the returns to investing in credible ESG systems by reducing policy uncertainty, improving information quality and aligning private incentives with social expectations (Campbell 2007; Dhaliwal et al. 2014; Eliwa et al. 2021; Houque et al. 2020; Ioannou and Serafeim 2012). Four WGI dimensions—RL, CC, GE and RQ—are especially salient. RL provides predictable enforcement of contracts, protection of property rights and impartial sanctioning of violations, encouraging preventive systems (e.g., environmental management, compliance audits) because these investments are more likely to be recognised and rewarded. CC reduces the feasibility of evading scrutiny through bribery or favouritism; when illicit pathways are closed, firms must meet standards through genuine improvements rather than informal arrangements (Cuervo-Cazurra 2006). GE enhances administrative capacity—more consistent inspections, timely permitting, reliable oversight—lowering compliance frictions and favouring firms that institutionalise ESG processes. RQ improves the clarity and coherence of rules, reducing ambiguity about what 'good' looks like and facilitating the adoption of recognised practices and verifiable disclosures. Two additional WGI dimensions reinforce governance complementarity: VA, by safeguarding free media, association, and civic participation, strengthens external scrutiny and stakeholder engagement, thereby increasing the reputational and market rewards to substantive ESG performance and high-quality disclosure; and PV, by lowering disruption and policy risk, lengthens firms' planning horizons and supports irreversible, long-gestation ESG investments while maintaining continuity in public policy priorities (Daske et al. 2008; Garmaise and Liu 2005; Kaufmann et al. 2011).

Together, these features complement firm-level efforts: they raise the expected payoff to credible ESG investment (because enforcement and market signals work), reduce the risk of opportunism by rivals (making ESG less of a competitive disadvantage) and strengthen stakeholder channels that reward authentic performance over symbolic gestures. In short, better governance makes it easier, safer and more valuable for firms to implement substantive ESG programmes and to disclose them transparently (Dhaliwal et al. 2014; Cheng et al. 2014; North 1990; Campbell 2007; Kaufmann et al. 2011; Ioannou and Serafeim 2012).

H1a (Complementarity). *Higher country-level governance quality is positively associated with firms' ESG performance*

An alternative theoretical link between country governance and sustainable performance is substitution. When public institutions embed social and environmental responsibilities into laws, regulations and norms, the need for firms to develop additional

voluntary initiatives diminishes. In governance-rich environments, strong enforcement, clear statutory baselines and credible oversight already secure many of the outcomes that voluntary corporate sustainability programmes aim to deliver. Consequently, the marginal benefit of adopting high-visibility ESG initiatives or expansive disclosure is lower—even though underlying corporate conduct may remain fully responsible (Matten and Moon 2008; Jackson and Apostolou 2010). This is both a behavioural and a measurement effect: as responsibilities are institutionalised and audited by the state, firms prioritise compliance and ‘quiet’ process improvements, while ratings that weight the breadth of voluntary policies and public commitments register less incremental credit for going beyond already high baselines.

Examples of this institutional substitution include social protections (comprehensive labour law, occupational-safety regulation, mandatory parental leave), environmental policy frameworks (national carbon-pricing schemes, renewable-energy investment, emissions and efficiency standards) and governance safeguards (minority-shareholder protection, related-party transaction rules, disclosure requirements). In such settings, sustainable practices are largely mandated, monitored and enforced collectively rather than initiated voluntarily at the firm level. Conversely, where governance is weaker and institutional voids persist, firms often compensate with prominent self-regulatory programmes, certification schemes and extensive reporting to build legitimacy—activities many ratings methodologies reward (Matten and Moon 2008; Jackson and Apostolou 2010). In this framework, higher country-level governance quality should be negatively associated with firms’ disclosure-weighted ESG performance, as institutionalised responsibilities substitute for voluntary firm-level initiatives and disclosures.

H1b (Substitution). *Higher country-level governance quality is negatively associated with firms’ ESG performance.*

Another mechanism linking national governance to ESG outcomes operates through controversies—event-based, publicly observable incidents such as environmental accidents, labour disputes, corruption scandals or governance failures. Whereas ESG performance reflects firms’ proactive initiatives and disclosure choices, controversy metrics are shaped by the information and enforcement environment. In practice, the number of incidents recorded by data providers (e.g., Refinitiv) depends on two components: (i) the latent incidence of misconduct and (ii) the probability that misconduct is detected, verified and publicised (Healy and Palepu 2001; Dyck et al. 2010). Country-level governance quality therefore matters in two conceptually distinct ways: as a prevention/deterrence force that lowers true misconduct and as a transparency/detection force that raises the likelihood that any misconduct is observed and recorded (Kaufmann et al. 2011).

Each of the six WGI dimensions—RL, CC, GE, RQ, VA and PV—shapes enforcement and the information environment through distinct channels (Kaufmann et al. 2011). RL strengthens contract enforcement and impartial sanctioning; CC curbs rent-seeking and bribery that would otherwise enable evasion; GE raises administrative capacity for monitoring and

inspections and RQ clarifies and stabilises rules across environmental, labour and governance domains. VA protects media freedom and civic oversight, heightening scrutiny and reputational consequences. PV reduces disruption and security risk, sustaining routine inspections and continuity of oversight. On the deterrence margin, these features reduce negligence and opportunism *ex ante*, shrinking the true frequency of ESG-relevant failures. Consistent with classic deterrence theory, higher expected costs of misconduct shift firms towards compliance and precautionary investment (Becker 1968).

H2a (Prevention/Deterrence). *Higher country-level governance quality is negatively associated with firms’ ESG controversies (i.e., fewer event-based failures in better-governed environments).*

Governance also shapes how much of the remaining misconduct becomes visible. Crucially, this is not solely about civic freedoms or media scrutiny; it is a system-wide effect that draws on multiple governance dimensions. RL enables due process and legal discovery; CC supports audit integrity and whistleblower credibility; GE equips agencies to collect, validate and share information; RQ mandates clearer reporting, incident notification and safety protocols with traceable documentation and VA empowers media, NGOs and citizens to scrutinise and escalate concerns (Kaufmann et al. 2011). In such settings, incidents are more likely to be detected (through monitoring, audits, inspections), verified (via credible institutions) and publicised (through functioning information channels). Empirical work shows that stronger monitoring and freer information flows raise reported wrongdoing even when behaviour does not worsen—for instance, media and whistle-blowers are pivotal in surfacing corporate fraud (Dyck et al. 2010), and improvements in monitoring increase recorded corruption without necessarily increasing true corruption (Olken 2007; Reinikka and Svensson 2004; Besley and Prat 2006; Ferraz and Finan 2008). Applied to ESG, better-governed countries may appear to have more controversies because incidents are more readily detected, documented and coded by data providers, whereas weak-governance contexts can under-record similar misconduct.

H2b (Detection/Transparency). *Higher country-level governance quality is positively associated with measured ESG controversies.*

Together, this framework explains why high-governance countries can exhibit strong ESG performance yet more recorded controversies: the same institutions that deter misconduct also expose what remains. Conversely, weak-governance settings may show fewer recorded controversies, but poorer ESG performance if limited monitoring suppresses detection while enforcement is lax (Besley and Prat 2006). Because both pairs of mechanisms pull in opposite directions—H1a (complementarity) versus H1b (substitution) for performance, and H2a (prevention/deterrence) versus H2b (detection/transparency) for controversies—the net effects of governance quality on measured outcomes are indeterminate *a priori* and therefore empirical. Importantly, controversy counts blend behaviour and visibility, treating them as pure misconduct risks, conflating under-detection with prevention and transparency

with over-reporting. Modelling governance through these channels—alongside the performance-side complementarity–substitution logic—provides a unified institutional lens: governance can complement or substitute firm ESG inputs, and it can both deter and reveal outcomes. The empirical tests below evaluate these mechanisms across WGI dimensions.

3 | Data and Methodology

3.1 | Data

This study's empirical analysis draws on an international firm-year panel covering publicly listed companies from 86 countries between 2002 and 2023, obtained from Refinitiv Eikon (LSEG Data & Analytics). The dataset provides the overall ESG score, scaled from 0 to 100, together with the three pillar scores—ESG—that assess firms' performance on each sustainability dimension. ESG scores are widely recognised as a reliable proxy for evaluating corporate sustainability (e.g., Moussa et al. 2026), where higher values indicate stronger performance. The panel is unbalanced due to heterogeneous data coverage across firms, years and jurisdictions.

In addition, the analysis incorporates the ESG Controversies Score, also sourced from Refinitiv Eikon. This score ranges from 0 to 100, where a value of 100 indicates that the firm has no recorded ESG-related controversies in the given year. When controversies occur—such as environmental accidents, labour rights violations, consumer safety issues, corruption or fraud investigations or governance misconduct—the score decreases below 100, with the magnitude of the reduction reflecting both the frequency and the severity of the reported incidents. The measure is event-based and dynamically updated, offering an outcome-oriented complement to the ESG pillar scores, which are predominantly disclosure- and policy-based. Thus, the Controversies Score captures reputational and behavioural risks that may not be evident from firms' stated sustainability practices alone.

Our key explanatory variables capture the notion of country-level governance quality. Data were obtained from the World Bank's WGI, which provide six measures of governance quality. These include VA, PV, GE, RQ, RL and CC. The World Bank provides the following descriptions: CC reflects perceptions regarding the extent to which public authority is misused for private benefit, whether through minor or large-scale corruption, as well as state capture by elites or private actors. GE evaluates the quality of public services, the competence and independence of the civil service, the effectiveness of policy design and implementation and the government's credibility in adhering to its policies. PV refers to the perceived probability of political unrest, instability or violence driven by political motives, including terrorism. RQ assesses the extent to which governments are seen as able to design and enforce sound policies and regulations that enable private sector development. RL measures perceptions of confidence in and compliance with societal rules, particularly regarding contract enforcement, property rights, the judicial system, the police and exposure to crime and violence. Finally, VA evaluates the degree to which citizens can participate in government selection and enjoy freedoms of expression, association and media. Each of these six

indicators represents an estimated score for a country on an aggregate index, normalised to follow a standard normal distribution, typically ranging between -2.5 and $+2.5$.

Control variables follow the standard ESG–finance literature and include company age, firm size (logarithm of market capitalisation), organisational scale (logarithm of employees), profitability (return on assets), valuation (price-to-book ratio), innovation intensity (R&D expenditures over total assets), investment (capital expenditures over assets), liquidity (quick ratio), leverage (debt-to-equity ratio) and cash holdings (cash over assets).

Table 1 shows the summary statistics of the variables used in the study across ~146k firm–years: overall ESG averages 46.1, with Governance highest (51.2), Social mid-range (46.8) and Environmental lowest (39.4) and most dispersed (SD: 29.6). The ESG Controversies Score averages 89.2 with a median of 100, and 79% of observations are at the maximum (100), indicating a strong ceiling effect. World Governance Indicators (WGI) means are positive on the -2.5 to $+2.5$ scale—CC (1.16), GE (1.27), PV (0.42), RL (1.21), RQ (1.21) and VA (0.84)—suggesting generally solid institutional quality but comparatively weaker stability and voice.

In Table 2, we present the correlation coefficients. The composite ESG score is very highly correlated with the Social (0.906) and Environmental (0.864) pillars, and less so with Governance (0.713). That pattern is expected because the overall ESG score is an aggregation of the E, S and G pillars, and Refinitiv/LSEG's methodology also overlays the Controversies signal—hence the observed negative correlations with Controversies (e.g., ESG = -0.328). Turning to the WGI block, the indicators are strongly intercorrelated (e.g., CC with RL, GE and RQ are all ≥ 0.90), with VA showing comparatively lower correlations (e.g., with GE and PV), so VA behaves a bit differently in your data. This suggests we should be cautious about including them all together in multivariate models (risk of multicollinearity). The WGI authors (Kaufmann and Kraay 2008, 2024) argue that—even though the six governance dimensions are highly intercorrelated—each indicator is conceptually distinct, capturing separate aspects of governance quality. They caution that aggregating these into a single index would obscure these conceptual differences and introduce statistical complications, such as correlated measurement errors across indicators. This position aligns with OECD guidance, which warns that undifferentiated aggregation of governance measures involves arbitrary weighting, sensitivity to scaling and the potential to mislead by masking the specific institutional drivers of governance outcomes (Arndt and Oman 2006).

3.2 | Empirical Strategy

We study how country-level governance quality shapes firms' sustainability outcomes using a multi-country firm–year panel in which the governance indicators vary at the country–year level while outcomes and controls vary at the firm–year level. This configuration calls for an empirical strategy that cleanly separates (i) within-country, over-time movements in governance from (ii) between-country

TABLE 1 | Summary statistics.

	Variable	N	Mean	Standard deviation	Min	Median	Max
(1)	ESG score	146,439	46.085	21.563	0.390	45.340	95.74
(2)	Social pillar score	146,429	46.806	24.412	0.053	45.464	98.402
(3)	Environmental pillar score	146,429	39.358	29.629	0.000	37.437	99.097
(4)	Governance pillar score	146,439	51.164	23.000	0.037	52.151	99.413
(5)	ESG controversies score	146,429	89.193	24.888	0.472	100.0	100.0
(6)	CC	146,439	1.162	0.794	-1.585	1.341	2.459
(7)	GE	146,439	1.269	0.580	-1.859	1.472	2.470
(8)	PV	146,439	0.423	0.610	-2.566	0.491	1.753
(9)	RL	146,439	1.212	0.706	-1.883	1.496	2.125
(10)	RQ	146,439	1.209	0.672	-2.302	1.439	2.309
(11)	VA	146,033	0.836	0.812	-1.907	1.077	1.801
(12)	Company age (log)	146,439	3.061	1.038	0.000	3.429	4.812
(13)	Market cap (log)	143,040	22.094	1.777	5.037	22.174	25.085
(14)	Employees (log)	146,439	7.786	3.503	0.000	8.757	14.648
(15)	ROA	146,439	3.945	6.698	-12.357	3.469	17.461
(16)	M/B ratio	146,439	2.907	3.853	0.000	1.708	26.064
(17)	R&D (% total assets)	145,330	24.063	9152.7	0.000	0.000	3489200
(18)	CAPEX (% total assets)	146,439	0.046	0.069	0.000	0.026	0.439
(19)	Quick ratio	146,439	2.198	1.828	0.000	1.319	5.218
(20)	Debt-to-equity	146,439	95.263	110.093	-0.579	56.058	415.608
(21)	Cash (% assets)	145,330	0.068	0.115	-0.034	0.024	100.0

Note: This table reports the number of observations (N), mean, standard deviation, minimum (Min), median (p50) and maximum (Max) for each variable used in this study. ESG Score, Social Pillar Score, Environmental Pillar Score, Governance Pillar Score and ESG Controversies Pillar Score are Refinitiv Eikon measures on a 0–100 scale (higher = better; for controversies, 100 indicates no controversies). The six country-level governance indicators—CC, control of corruption, GE, government effectiveness, PV, political stability and absence of violence/terrorism, RL, rule of law, RQ, regulatory quality, and VA, voice and accountability—are from the World Bank’s worldwide governance indicators (WGI) database and are reported on the standard WGI scale (approximately -2.5 to +2.5; higher = better). Company age (log), Market Cap (log) and Employees (log) are natural logarithms of the firm’s age, market capitalisation and number of employees, respectively. ROA is return on assets; the M/B ratio is the market value of equity divided by the book value of equity; R&D (% total assets) and CAPEX (% total assets) are each scaled by total assets; Quick Ratio is (current assets—inventories) divided by current liabilities; Debt-to-equity measures leverage as total debt over shareholders’ equity and Cash (% Assets) is cash and short-term investments over total assets. Country-level variables (CC–VA) come from WGI, while ESG and firm-level variables come from Refinitiv Eikon.

persistent differences, while addressing unobserved, time-invariant firm heterogeneity. We therefore estimate a Mundlak (1978) correlated-random-effects (CRE) model that augments a random-intercept specification with the time means of the time-varying regressors, allowing the firm effect to correlate with persistent levels of both governance and firm characteristics. This delivers fixed-effects-equivalent identification for the ‘within’ effects of governance, retains information on ‘between’ differences and avoids the loss of observations that a strict fixed-effects design would entail—an important consideration given our dependent variable’s limited within-firm variation.

In this framework, consider the following general baseline specification:

$$y_{it} = \alpha + W_{c(i)t}^T \beta + X_{it}^T \gamma + c_i + \lambda_t + u_{it}, \quad (1)$$

where y_{it} is the outcome for firm i in year t , $W_{c(i)t}$ is a vector of country-level governance measures observed at the country-year of firm i , X_{it} are firm-level controls, c_i is an unobserved time-invariant firm effect, λ_t are year effects and u_{it} is an idiosyncratic error.

Define time averages over the sample window:

$$\bar{W}_{c(i)} = \frac{1}{T_{c(i)}} \sum_t W_{c(i)t}, \quad \bar{X}_i = \frac{1}{T_i} \sum_t X_{it}. \quad (2)$$

Following Mundlak (1978), allow the firm effect to correlate with the regressors via their means:

$$c_i = \bar{W}_{c(i)}^T \theta_w + \bar{X}_i^T \theta_x + a_i, E[a_i | \{W_{c(i)t}, X_{it}\}] = 0. \quad (3)$$

Using the identity $W_{c(i)t} = (W_{c(i)t} - \bar{W}_{c(i)}) + \bar{W}_{c(i)}$ and $X_{it} = (X_{it} - \bar{X}_i) + \bar{X}_i$, we have $W_{c(i)t}^T \beta = (W_{c(i)t} - \bar{W}_{c(i)})^T \beta + \bar{W}_{c(i)}^T \beta$ and $X_{it}^T \gamma = (X_{it} - \bar{X}_i)^T \gamma + \bar{X}_i^T \gamma$, which yields equation (4):

$$y_{it} = \alpha + (W_{c(i)t} - \bar{W}_{c(i)})^T \beta + \bar{W}_{c(i)}^T (\beta + \theta_w) + (X_{it} - \bar{X}_i)^T \gamma + \bar{X}_i^T (\gamma + \theta_x) + \lambda_t + a_i + u_{it}. \quad (4)$$

For estimation, it is convenient to reparametrise and let the mean terms have their own free coefficient vector ϕ_w and ϕ_x , rather than imposing $\phi_w = \beta + \theta_w$, $\phi_x = \gamma + \theta_x$. So, we redefine the model as:

TABLE 2 | Correlation coefficients.

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(1) ESG score	1.000									
(2) Social pillar score	0.906*	1.000								
(3) Environmental pillar score	0.864*	0.743*	1.000							
(4) Governance pillar score	0.713*	0.459*	0.445*	1.000						
(5) ESG controversies	-0.328*	-0.312*	-0.312*	-0.204*	1.000					
(6) CC	0.011*	0.013*	0.004	0.018*	-0.064*	1.000				
(7) GE	-0.009*	-0.012*	-0.011*	0.007*	-0.069*	0.938*	1.000			
(8) PV	0.015*	-0.012*	0.044*	0.022*	-0.016*	0.819*	0.783*	1.000		
(9) RL	0.020*	0.036*	-0.001	0.015*	-0.093*	0.955*	0.947*	0.777*	1.000	
(10) RQ	0.015*	0.042*	-0.022*	0.016*	-0.072*	0.929*	0.902*	0.749*	0.952*	1.000
(11) VA	0.084*	0.117*	0.053*	0.033*	-0.108*	0.796*	0.697*	0.677*	0.821*	0.823*
(12) Company age (log)	0.107*	0.077*	0.124*	0.059*	-0.034*	-0.090*	-0.066*	-0.067*	-0.068*	-0.088*
(13) Market cap (log)	0.502*	0.449*	0.507*	0.319*	-0.373*	0.000	0.053*	-0.032*	0.017*	-0.027*
(14) Employees (log)	0.367*	0.321*	0.368*	0.234*	-0.271*	0.002	0.047*	-0.060*	0.012*	-0.042*
(15) ROA	0.085*	0.066*	0.077*	0.064*	0.000	-0.035*	-0.028*	-0.034*	-0.046*	-0.050*
(16) M/B ratio	-0.018*	0.015*	-0.069*	-0.031*	0.007*	0.023*	0.036*	-0.052*	0.047*	0.042*
(17) R&D (% total assets)	-0.005	-0.005	-0.002	-0.003	0.001	-0.002	-0.001	0.003	-0.002	-0.002
(18) CAPEX (% total assets)	-0.095*	-0.091*	-0.067*	-0.057*	0.044*	0.012*	0.001	0.026*	-0.003	0.001
(19) Quick ratio	-0.076*	-0.044*	-0.100*	-0.029*	0.005	0.004	0.021*	-0.031*	0.038*	0.037*
(20) Debt-to-equity	0.131*	0.122*	0.159*	0.080*	-0.128*	-0.019*	-0.021*	-0.038*	-0.011*	-0.017*
(21) Cash (% assets)	-0.134*	-0.101*	-0.163*	-0.110*	0.064*	0.038*	0.065*	0.022*	0.069*	0.057*

Variables	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)
(11) VA	1.000										
(12) Company age (log)	-0.053*	1.000									
(13) Market cap (log)	-0.025*	0.057*	1.000								
(14) Employees (log)	-0.018*	0.088*	0.566*	1.000							
(15) ROA	-0.046*	0.056*	0.281*	0.136*	1.000						
(16) M/B ratio	0.019*	-0.028*	0.151*	0.037*	0.209*	1.000					
(17) R&D (% total assets)	-0.001	0.001	-0.048*	-0.006	-0.002	-0.002	1.000				
(18) CAPEX (% total assets)	0.004	-0.027*	-0.093*	-0.235*	0.046*	-0.025*	-0.002	1.000			
(19) Quick Ratio	0.021*	-0.037*	0.035*	-0.080*	-0.172*	-0.077*	0.004	-0.185*	1.000		
(20) Debt-to-equity	0.019*	-0.027*	0.152*	0.137*	-0.182*	0.102*	-0.002	-0.102*	0.086*	1.000	
(21) Cash (% Assets)	0.043*	-0.013*	-0.132*	-0.080*	-0.053*	0.200*	0.021*	-0.038*	0.093*	-0.188*	1.000

Note: Refer to Table 1 notes for variable definitions.

*Indicates significance at the 10% level.

$$y_{it} = \alpha + (W_{c(i)t} - \bar{W}_{c(i)})^\top \beta + \bar{W}_{c(i)}^\top \phi_w + (X_{it} - \bar{X}_i)^\top \gamma + \bar{X}_i^\top \phi_x + \lambda_t + a_i + u_{it}. \quad (5)$$

In this CRE formulation, coefficients on deviations β capture within-country (over-time) associations, γ within-firm (over-time) associations and ϕ_w , ϕ_x the between-country/firm components.

After deriving the Mundlak (CRE) specification, we tailor the outcome model to the distribution of each dependent variable. To study sustainability performance, we use the Refinitiv ESG scores, which in our data are approximately symmetric and well behaved; we therefore estimate a linear CRE (Mundlak-adjusted RE) model,

retaining the within/between decomposition and allowing the firm effect to correlate with the time means of the covariates. To study ethical behaviour, we use the Refinitiv ESG Controversies Score, which presents substantial distributional challenges: a heavy mass at the upper bound (many firm-years with no controversies) and limited within-firm movement over time. In this setting, a linear model is not suited—it can extrapolate outside [0,1] and is inefficient under such skewness—and a strict FE design would discard many observations with no within variation. Accordingly, we implement the Mundlak (CRE) structure within a fractional logit framework (Papke and Wooldridge 1996), which ensures fitted values remain in [0,1], accommodates the skew and boundary mass, and preserves both within-country (over-time) and between-country information. To further enhance the

reliability of our results, all specifications include year and sector fixed effects, and inference is based on robust standard errors clustered at the firm level.

4 | Results Discussion

4.1 | Governance Indicators and ESG

The baseline results of our study concerning the relationship between WGI and ESG performance are presented in Table 3. The results show that higher country-level governance quality is strongly and consistently associated with stronger ESG performance, providing compelling cross-sectional evidence for H1a (complementarity). The between-country ('Bar') coefficients are uniformly positive and estimated with a high degree of statistical precision for all six WGI dimensions—VA, RQ, PV, RL, GE and CC—with especially large associations for VA (+4.66), RQ (+3.88) and PV (+3.99), followed by RL (+3.15), GE (+2.22) and CC (+0.97).

These magnitudes indicate that firms in countries with durably stronger institutions systematically report higher ESG performance because governance makes credible ESG cheaper to implement and more valuable to disclose. Four WGI dimensions are central to this complementarity. RL provides predictable enforcement of contracts, protection of property and impartial sanctioning of violations, so investments in prevention (e.g., environmental management systems, compliance audits) are more likely to be recognised and rewarded. CC removes illicit pathways (bribery, favouritism) that would otherwise allow firms to bypass standards; when these channels close, firms must improve practices rather than rely on informal arrangements (Cuervo-Cazurra 2006). GE raises administrative capacity—more consistent inspections, timely permitting, reliable oversight—lowering compliance frictions and favouring firms that institutionalise ESG processes. RQ clarifies and harmonises rules, reducing ambiguity about what 'good' looks like and easing adoption of recognised practices and verifiable disclosures.

Two additional WGI dimensions reinforce these channels. VA safeguards free media, association and civic participation, strengthening external scrutiny and stakeholder engagement and increasing reputational and market rewards for substantive ESG and high-quality reporting. PV lowers disruption and policy risk, lengthening planning horizons and supporting irreversible, long-gestation ESG investments while maintaining continuity in public priorities (Campbell 2007; Dhaliwal et al. 2014; Eliwa et al. 2021; Houqe et al. 2020; Ioannou and Serafeim 2012; Daske et al. 2008; Garmaise and Liu 2005; Kaufmann et al. 2011).

Together, these features complement firm-level efforts: they raise the expected payoff to genuine ESG investment (because enforcement and market signals work), reduce the risk of opportunism by rivals (making ESG less of a competitive disadvantage) and strengthen stakeholder channels that reward substance over symbolism. In short, better governance makes it easier, safer and more valuable for firms to implement substantive ESG programmes and to disclose them transparently (Dhaliwal et al. 2014; Cheng et al. 2014; North 1990;

Campbell 2007; Kaufmann et al. 2011; Ioannou and Serafeim 2012). Taken together, these results strongly support H1a—that higher country-level governance quality is positively associated with firms' ESG performance.

Turning to the within-country dynamics, the results provide strong support for H1a for four of the six WGI dimensions: when a country improves over time on RL (+4.68), CC (+3.64), GE (+2.26) or PV (+1.03), firms' ESG scores rise materially and with high precision. The two exceptions are RQ, whose short-run within effect is small and statistically indistinguishable from zero, and VA, whose within coefficient is negative and significant (−2.26). A plausible interpretation of the negative within-country effect of VA—despite its strong positive cross-sectional association—is a 'greenwashing clean-up' dynamic. When VA rises, the expected intensity of media scrutiny, NGO monitoring and stakeholder participation increases. Anticipating closer probing—and fearing accusations of greenwashing or unethical conduct—firms tighten their reporting: they withdraw or narrow under-substantiated initiatives, restate emissions baselines and target pathways to meet stricter standards and upgrade assurance (e.g., from limited to reasonable). In the short run, stronger civic accountability induces a reporting clean-up: firms drop or narrow claims that cannot be substantiated, audits reveal documentation gaps that must be corrected and restatements or corrective action plans are recorded by rating agencies. As a result, disclosure-weighted ESG scores temporarily decline—not because underlying practices worsen, but because reporting becomes more conservative and verifiable. The temporary dip is thus less a deterioration in real conduct than an adjustment from 'promotional' ESG to verifiable ESG under a higher bar of civic accountability.

This distinction also explains why VA differs from the enforcement-oriented WGI dimensions. Improvements in RL, CC, GE and PV primarily reduce compliance frictions and increase sanction certainty. Firms can therefore translate existing systems and disclosures into higher measured ESG quickly, yielding positive within-country effects. By contrast, VA upgrades first raise expectations and reveal deficiencies: regulatory processes become more participatory, information flows more freely and stakeholders gain stronger means to question or verify firms' claims. These mechanisms expose documentation gaps and compel firms to validate existing disclosures before expanding or announcing new ESG initiatives. In short, the immediate effect of higher VA is diagnostic—identifying and correcting deficiencies—rather than additive—expanding and publicising programmes. This produces a short-run dip in scores that reverses as remediation is completed and disclosures align with the higher standard, consistent with VA's strong positive association in the cross-section.

The results for RQ highlight a complementary temporal nuance. While the cross-sectional association between RQ and ESG performance is strong and positive, short-run within-country changes in RQ are small and statistically insignificant, suggesting that policy design and regulatory coherence primarily influence ESG through gradual, structural channels rather than immediate shifts. Taken together, the evidence shows heterogeneous temporal profiles across governance dimensions: enforcement and administrative capacity (RL, CC,

TABLE 3 | Worldwide governance indicators (WGI) and ESG score.

Variables	(1) ESG score	(2) ESG score	(3) ESG score	(4) ESG score	(5) ESG score	(6) ESG score
CC Dev	3.6393*** (10.0036)					
CC Bar	0.9732*** (3.8887)					
GE Dev		2.2644*** (5.3157)				
GE Bar		2.2236*** (9.5099)				
PV Dev			1.0257*** (3.3808)			
PV Bar			3.9930*** (18.1421)			
RL Dev				4.6776*** (8.0544)		
RL Bar				3.1540*** (16.6692)		
RQ Dev					0.6920 (1.6327)	
RQ Bar					3.8819*** (20.7801)	
VA Dev						-2.2576*** (-3.3963)
VA Bar						4.6641*** (33.3355)
Company age (log) Dev	2.1775*** (8.6878)	2.2087*** (8.8517)	2.1653*** (8.6915)	2.2117*** (8.8819)	2.1123*** (8.4785)	2.1641*** (8.6774)
Company age (log) Bar	0.7175*** (5.4415)	0.6506*** (4.8640)	0.7708*** (5.8487)	0.7170*** (5.3480)	0.8477*** (6.3421)	0.8557*** (6.5447)
Market cap (log) Dev	2.6258*** (24.0895)	2.6451*** (24.2154)	2.6592*** (24.2614)	2.6381*** (24.1267)	2.6309*** (24.0588)	2.6336*** (23.9708)
Market cap (log) Bar	5.8575*** (43.9400)	5.8184*** (43.6611)	5.7218*** (43.0821)	5.8106*** (43.5105)	5.8119*** (43.6044)	5.9596*** (44.9767)
Employees (log) Dev	0.2760*** (9.5951)	0.2641*** (9.1181)	0.2564*** (8.8054)	0.2720*** (9.4137)	0.2582*** (8.9021)	0.2482*** (8.5053)
Employees (log) Bar	0.9709*** (15.5581)	0.9240*** (14.7753)	1.0584*** (16.9460)	0.9338*** (14.8975)	0.9744*** (15.5679)	0.9393*** (15.2873)
ROA Dev	-0.0482*** (-5.3004)	-0.0490*** (-5.4051)	-0.0483*** (-5.3088)	-0.0484*** (-5.3374)	-0.0480*** (-5.2913)	-0.0485*** (-5.3310)
ROA Bar	0.1931*** (8.0724)	0.2023*** (8.3780)	0.2126*** (8.9421)	0.2198*** (9.1398)	0.2316*** (9.6762)	0.2376*** (10.0539)
M/B ratio Dev	-0.2584*** (-12.9727)	-0.2640*** (-13.2914)	-0.2639*** (-13.3055)	-0.2596*** (-13.0788)	-0.2646*** (-13.3152)	-0.2679*** (-13.4864)
M/B ratio Bar	-0.9617*** (-24.1234)	-0.9615*** (-23.8990)	-0.9147*** (-22.9005)	-0.9696*** (-24.1879)	-0.9653*** (-24.1302)	-0.9850*** (-25.0110)

(Continues)

TABLE 3 | (Continued)

Variables	(1) ESG score	(2) ESG score	(3) ESG score	(4) ESG score	(5) ESG score	(6) ESG score
R&D (% assets) Dev	0.0753*** (7.3213)	0.0761*** (7.1386)	0.0756*** (6.7099)	0.0755*** (7.1953)	0.0780*** (7.4219)	0.0833*** (7.3397)
R&D (% assets) Bar	0.0753*** (7.3228)	0.0761*** (7.1397)	0.0756*** (6.7104)	0.0755*** (7.1967)	0.0780*** (7.4233)	0.0833*** (7.3411)
CAPEX (% assets) Dev	-4.7657*** (-3.2139)	-4.6265*** (-3.1266)	-4.6154*** (-3.1230)	-4.6200*** (-3.1273)	-4.6639*** (-3.1550)	-4.5873*** (-3.0841)
CAPEX (% assets) Bar	-12.8292*** (-5.5280)	-13.0876*** (-5.6108)	-12.5995*** (-5.4192)	-13.0257*** (-5.5630)	-12.6148*** (-5.3860)	-13.1289*** (-5.5703)
Quick ratio Dev	-0.1645** (-2.1538)	-0.1577** (-2.0634)	-0.1557** (-2.0339)	-0.1663** (-2.1835)	-0.1479* (-1.9313)	-0.1326* (-1.7279)
Quick ratio Bar	-1.8075*** (-16.8868)	-1.8807*** (-17.4981)	-1.7717*** (-16.5443)	-1.8861*** (-17.5911)	-1.8852*** (-17.6135)	-1.8611*** (-17.4663)
Debt-to-equity Dev	0.0104*** (11.3179)	0.0105*** (11.5204)	0.0104*** (11.4349)	0.0105*** (11.4491)	0.0104*** (11.3709)	0.0104*** (11.3333)
Debt-to-equity Bar	0.0191*** (12.1826)	0.0195*** (12.2401)	0.0191*** (12.1524)	0.0195*** (12.2767)	0.0194*** (12.3303)	0.0180*** (11.6113)
Cash (% assets) Dev	-3.3208*** (-5.2400)	-3.4041*** (-5.3675)	-3.5596*** (-5.6120)	-3.3947*** (-5.3546)	-3.5164*** (-5.5426)	-3.5532*** (-5.6062)
Cash (% Assets) Bar	0.1865 (0.1701)	-0.0619 (-0.0560)	-0.6435 (-0.5893)	-0.8873 (-0.8035)	-1.5390 (-1.3984)	-3.0817*** (-2.7811)
Constant	-124.64*** (-42.5934)	-124.27*** (-42.2374)	-122.25*** (-42.1334)	-125.16*** (-42.7505)	-126.73*** (-43.4847)	-127.93*** (-44.4640)
Sector fixed effects	Y	Y	Y	Y	Y	Y
Year fixed effects	Y	Y	Y	Y	Y	Y
Observations	143,038	143,038	143,038	143,038	143,038	142,641
Number of firms	15,674	15,674	15,674	15,674	15,674	15,646
R-squared within	0.516	0.516	0.516	0.517	0.516	0.516
R-squared between	0.275	0.263	0.275	0.270	0.278	0.304
R-squared overall	0.406	0.395	0.402	0.402	0.407	0.430

Note: This table reports results from correlated random-effects (Mundlak) panel regressions of ESG Score on the worldwide governance indicators (WGI: CC, GE, PV, RL, RQ and VA). Each specification includes both the within-country/firm (demeaned) and between-country/firm (mean) components of the governance indicators and control variables. 'Dev' denotes the within-country/firm (time-demeaned) component, computed as the deviation from the country/firm-specific mean (e.g., $CC_{c(t)t} - \overline{CC}_{c(t)}$). 'Bar' denotes the between-country/firm (mean) component, that is, the country/firm-specific average over time (e.g., $\overline{CC}_{c(t)}$). All models include sector and year fixed effects. Please see Table 1 notes for variable definitions. Models are estimated by random-effects GLS with robust standard errors clustered at the firm level. Robust z-statistics are in parentheses. *, ** and *** indicate significance at the 10%, 5% and 1% levels, respectively.

GE) map to prompt within-country gains in ESG, whereas VA—by institutionalising stakeholder rights and intensifying scrutiny—can initially depress measured, programme-weighted scores even as it strengthens the broader accountability environment.

Control variables behave in line with expectations for disclosure-weighted ESG measures. Larger, older and more employee-intensive firms score higher on ESG both within and across firms, consistent with scale advantages in reporting and programmatic capacity. Profitability exhibits a split pattern—negative within but positive between—implying that transitory earnings pressure may compress ESG activity even as structurally profitable firms sustain higher baseline programmes. Market valuation (M/B) is negatively related to ESG, while R&D intensity is positive, Capital Expenditures _

(CAPEX) is negative, liquidity (quick ratio) is modestly negative, leverage is positive and cash holdings turn negative in the between dimension in the VA specification. These stable and interpretable covariate patterns provide a useful validity check on the main results.

Model fit is strong for panel ESG regressions of this scope. The within R-squared centres around 0.516–0.517 across specifications, and the overall R-squared increases to 0.430 when VA is included, reflecting the substantial cross-sectional explanatory power of voice and civic freedom for measured ESG performance. The Mundlak (CRE) structure ensures that these estimates separately capture within-country movements and between-country differences, mitigating the confounding that would arise in a pure pooled or strict fixed-effects design, given the limited within variation of ESG.

Overall, Table 3 shows that country governance quality predominantly complements firm-level ESG (H1a): within-country improvements in CC, GE, PV and RL are associated with higher ESG scores, and cross-sectional associations are positive for all six WGI dimensions. By contrast, the negative within coefficient on VA appears to reflect a scrutiny-driven reporting clean-up (tighter verification, restatements, trimmed unverifiable claims) that temporarily lowers disclosure-weighted scores.

Table 4 reports the relationship between the WGI and firms' Environmental Pillar Scores. The estimates reveal that stronger national governance is consistently associated with better environmental performance, extending the complementarity pattern observed for overall ESG. Both the cross-sectional ('Bar') and within-country ('Dev') coefficients are mostly positive and economically large, confirming that improvements in institutional quality support environmental investment, disclosure and monitoring (H1a).

Among the six WGI dimensions, RL exhibits the strongest association: a one-unit within-country improvement in RL raises the Environmental Pillar Score by +8.76, the largest coefficient in the table. These findings point to predictable enforcement and strong legal protection as key drivers of environmental compliance and long-term abatement investment. When contracts and property rights are reliably enforced, firms face stronger liability for environmental harm and greater incentives to internalise externalities through cleaner technologies and process upgrades.

Similarly, GE and CC show large positive within effects (+4.39 and +4.47, respectively), suggesting that efficient, impartial administrations and reduced rent-seeking enhance the credibility and implementation of environmental regulation. Where inspections, licensing and monitoring are timely and fair, firms can plan environmental projects with less uncertainty and are less tempted to circumvent compliance through informal payments. PV also supports environmental progress: both its within (+1.10) and between (+5.67) effects are positive and significant. Stable political environments reduce policy reversals and the risk of disruption, extending firms' planning horizons and enabling long-gestation environmental investments such as emissions-control systems or renewable-energy projects.

The cross-sectional ('Bar') coefficients for all six governance dimensions are positive and highly significant, confirming that firms located in countries with durably stronger institutions exhibit higher baseline environmental scores. This pattern aligns with the idea that credible governance frameworks provide the legal certainty, regulatory coherence and enforcement capacity necessary for sustained environmental performance.

Two governance dimensions display a different within-country profile: VA and RQ. For VA, the pattern is consistent with the 'greenwashing clean-up' dynamic documented for the overall ESG Score: a negative within coefficient (−5.00) alongside a large positive between effect (+5.09). As media freedom and civic participation expand, environmental claims face tighter external scrutiny; firms withdraw or narrow unverifiable assertions (e.g., net-zero trajectories, Scope 3 estimates, the use and quality of offsets), restate baselines and target pathways to

meet stricter evidentiary standards and upgrade assurance (from limited to reasonable). In the short run, this accountability-driven clean-up reduces reported breadth and pushes down disclosure-weighted Environmental scores—not because underlying practices worsen, but because reporting becomes more conservative and verifiable. Over time, greater transparency and third-party validation sustain the strong positive cross-sectional relationship.

For RQ, a similar transitional clean-up mechanism helps explain the negative within coefficient (−1.63) despite a strong positive cross-sectional effect (+3.96). When RQ improves—via clearer, auditable standards and new reporting protocols (e.g., mandatory Scope 1–3 accounting rules, taxonomy alignment, MRV systems, tighter definitions of 'renewable' or 'low-carbon')—firms are pushed to re-document claims, reclassify activities, cap or drop overstated initiatives and correct data gaps uncovered by early audits and compliance reviews. This raises short-run compliance costs and compresses disclosure-weighted Environmental scores as reported breadth shrinks, and documentation is re-verified. As regulatory coherence persists and firms' systems catch up, measured performance recovers, consistent with the positive between-country association.

Overall, Table 4 shows that improvements in enforcement-oriented governance dimensions translate into stronger environmental outcomes. These results reinforce H1a—that good governance complements corporate environmental responsibility—primarily by lowering enforcement uncertainty and enhancing credibility in environmental reporting and investment.

Table 5 presents the results for the relationship between the WGI and firms' Social Pillar Scores. The estimates reveal a strong and systematic association between national governance quality and firms' social performance, extending the evidence for complementarity observed in the overall ESG and Environmental dimensions. Both the within-country ('Dev') and cross-sectional ('Bar') coefficients are positive and highly significant across all six governance dimensions, indicating that stronger institutional environments foster more consistent and credible social practices and disclosures (H1a).

Among the WGI dimensions, RL again exhibits the largest within-country effect (+8.28), confirming that predictable enforcement and legal protection play a central role in shaping corporate social conduct. Reliable judicial systems and contract enforcement strengthen labour rights, workplace safety and nondiscrimination standards, while also giving credibility to grievance mechanisms and compliance frameworks. CC shows a similarly large within coefficient (+6.17), suggesting that the suppression of bribery and favouritism directly supports fair employment practices and transparent stakeholder relations. When informal payments and favouritism decline, firms must compete on compliance and merit, improving working conditions and social accountability.

GE also exhibits a substantial positive within effect (+3.38) and a strong between effect (+2.71), consistent with the notion that capable public administration enhances the implementation and oversight of labour, safety and social-welfare regulation.

TABLE 4 | Worldwide governance indicators (WGI) and environmental pillar score.

Variables	(1) Environmental pillar score	(2) Environmental pillar score	(3) Environmental pillar score	(4) Environmental pillar score	(5) Environmental pillar score	(6) Environmental pillar score
CC Dev	4.4659*** (8.5318)					
CC Bar	1.0714*** (3.3279)					
GE Dev		4.3929*** (6.7473)				
GE Bar		2.6870*** (9.2626)				
PV Dev			1.0993** (2.3667)			
PV Bar			5.6711*** (20.6374)			
RL Dev				8.7631*** (10.0065)		
RL Bar				3.4309*** (14.6403)		
RQ Dev					-1.6320*** (-2.6476)	
RQ Bar					3.9629*** (17.2849)	
VA Dev						-5.0032*** (-5.0851)
VA Bar						5.0897*** (29.0649)
Constant	-194.04*** (-50.3128)	-193.78*** (-49.9549)	-191.24*** (-50.0380)	-194.24*** (-50.1292)	-195.63*** (-50.6655)	-196.96*** (-51.4589)
Controls	Y	Y	Y	Y	Y	Y
Sector fixed effects	Y	Y	Y	Y	Y	Y
Year fixed effects	Y	Y	Y	Y	Y	Y
Observations	143,028	143,028	143,028	143,028	143,028	142,631
Number of firms	15,674	15,674	15,674	15,674	15,674	15,646
R-squared within	0.429	0.430	0.429	0.432	0.429	0.431
R-squared between	0.374	0.361	0.377	0.365	0.371	0.388
R-squared overall	0.438	0.429	0.440	0.434	0.435	0.453

Note: This table reports results from correlated random-effects (Mundlak) panel regressions of ESG Environmental Pillar Score on the Worldwide Governance Indicators (WGI: CC, GE, PV, RL, RQ and VA). Each specification includes both the within-country/firm (demeaned) and between-country/firm (mean) components of the governance indicators and control variables. 'Dev' denotes the within-country/firm (time-demeaned) component, computed as the deviation from the country/firm-specific mean (e.g., $CC_{c(i)t} - \overline{CC}_{c(i)}$). 'Bar' denotes the between-country/firm (mean) component, that is, the country/firm-specific average over time (e.g., $\overline{CC}_{c(i)}$). All models include sector and year fixed effects. Please see Table 1 notes for variable definitions. Models are estimated by random-effects GLS with robust standard errors clustered at the firm level. Robust z-statistics are in parentheses. ** and *** indicate significance at the 5% and 1% levels, respectively.

TABLE 5 | Worldwide governance indicators (WGI) and social pillar score.

Variables	(1) Social pillar score	(2) Social pillar score	(3) Social pillar score	(4) Social pillar score	(5) Social pillar score	(6) Social pillar score
CC Dev	6.1721*** (14.0036)					
CC Bar	0.6210** (2.0353)					
GE Dev		3.3750*** (6.5243)				
GE Bar		2.7144*** (9.5425)				
PV Dev			2.2922*** (6.1374)			
PV Bar			4.2123*** (15.7243)			
RL Dev				8.2799*** (11.9727)		
RL Bar				4.3554*** (18.7533)		
RQ Dev					2.4404*** (4.6294)	
RQ Bar					5.3172*** (23.3031)	
VA Dev						2.4316*** (2.8194)
VA Bar						6.5452*** (38.4980)
Constant	-136.47*** (-39.8212)	-135.75*** (-39.3826)	-133.52*** (-39.2797)	-137.22*** (-40.0790)	-139.47*** (-40.9756)	-141.99*** (-42.6891)
Controls	Y	Y	Y	Y	Y	Y
Sector fixed effects	Y	Y	Y	Y	Y	Y
Year fixed effects	Y	Y	Y	Y	Y	Y
Observations	143,028	143,028	143,028	143,028	143,028	142,631
Number of firms	15,674	15,674	15,674	15,674	15,674	15,646
R-squared within	0.475	0.475	0.474	0.477	0.474	0.474
R-squared between	0.194	0.178	0.188	0.188	0.202	0.246
R-squared overall	0.347	0.330	0.334	0.339	0.346	0.379

Note: This table reports results from correlated random-effects (Mundlak) panel regressions of ESG Social Pillar Score on the worldwide governance indicators (WGI: CC, GE, PV, RL, RQ and VA). Each specification includes both the within-country/firm (demeaned) and between-country/firm (mean) components of the governance indicators and control variables. 'Dev' denotes the within-country/firm (time-demeaned) component, computed as the deviation from the country/firm-specific mean (e.g., $CC_{c(t)t} - \overline{CC}_{c(t)}$). 'Bar' denotes the between-country/firm (mean) component, that is, the country/firm-specific average over time (e.g., $\overline{CC}_{c(t)}$). All models include sector and year fixed effects. Please see Table 1 notes for variable definitions. Models are estimated by random-effects GLS with robust standard errors clustered at the firm level. Robust z-statistics are in parentheses. ** and *** indicate significance at the 5% and 1% levels, respectively.

Where inspections and enforcement are timely and impartial, firms are more likely to adhere to formal standards on worker protection, community engagement and health and safety compliance. PV likewise contributes positively to firms' social performance, both within countries over time (+2.29) and across countries (+4.21). By reducing policy uncertainty,

conflict risk and institutional disruption, stability allows firms to sustain long-term human-capital investments, worker training and community programmes.

Two governance dimensions with strong informational and participatory components—RQ and VA—display particularly

large between-country effects (+5.32 and +6.55, respectively) and statistically significant within-country improvements (+2.44 and +2.43). These results suggest that enhanced regulatory coherence and civic participation not only improve compliance but also broaden the social accountability landscape. As regulatory clarity increases, firms face stronger incentives to align with international labour and social standards (e.g., ILO conventions, diversity and inclusion frameworks). Meanwhile, greater VA empower workers, media and civil society to monitor firm behaviour, promote transparency in supply chains and pressure companies to adopt inclusive and equitable practices.

Taken together, the social pillar results reinforce the governance complementarity hypothesis (H1a). Improvements in enforcement-oriented dimensions (RL, CC, GE) strengthen the formal underpinnings of labour and stakeholder protection, while information- and participation-based dimensions (RQ, VA) expand external oversight and reputational incentives. The consistent and positive associations across all six WGI indicators demonstrate that stronger governance not only deters social misconduct but also enables sustained, verifiable progress in firms' social performance.

Table 6 presents the results for the relationship between national governance indicators (WGI) and firms' Governance Pillar Scores. Overall, the evidence shows that country-level institutional quality remains positively associated with firms' internal governance structures in the cross-section but displays more heterogeneous, and in some cases negative, within-country dynamics.

Across all specifications, the between-country ('Bar') coefficients are positive and statistically significant for every WGI dimension. Firms located in countries with persistently stronger governance institutions—particularly those with higher RQ (+1.76), PV (+2.05), RL (+1.15) and VA (+1.70)—exhibit systematically higher Governance Pillar Scores. This cross-sectional pattern aligns with the complementarity hypothesis (H1a): durable institutional quality supports better firm-level governance practices through clearer rules, stronger investor protections and more credible oversight frameworks. In these environments, formal disclosure standards, board independence requirements and audit integrity tend to be embedded in both public and private governance systems, allowing firms to sustain higher-quality internal controls and transparency mechanisms.

By contrast, the within-country ('Dev') coefficients suggest that short-run improvements in institutional governance do not translate immediately into higher firm-level governance scores—and in some cases, the effects are negative. Notably, RL (−3.55), GE (−1.81) and VA (−4.74) all show significant negative within coefficients. These results likely reflect a short-term adjustment process similar to the 'accountability clean-up' dynamic observed for the overall ESG and environmental pillar scores.

When institutional enforcement, administrative efficiency or civic scrutiny improves, existing firm-level governance systems come under closer inspection. Firms may be compelled to restate governance structures, revise disclosure of board

composition or audit practices and correct previously overstated or incomplete governance claims. As stronger legal and civic oversight exposes deficiencies—such as inconsistencies in ownership reporting, weak board diversity or inadequate audit assurance—measured Governance Pillar scores can temporarily decline. This decline does not signal deterioration in governance quality per se but rather a transitional tightening of reporting and assurance standards. Over time, as firms align internal governance frameworks with higher institutional expectations, scores are likely to recover, consistent with the positive long-run (cross-sectional) relationship.

For CC and RQ, the within-country effects are statistically insignificant, indicating that no clear short-term pattern can be extracted from these indicators. Similarly, PV shows no measurable within-country impact on firms' governance scores, although its cross-sectional association remains positive and significant.

Taken together, Table 6 highlights that improvements in national governance quality are associated with higher corporate governance standards across countries, but that within-country transitions often reveal an adjustment phase where enhanced scrutiny and institutional tightening temporarily compress governance scores. These results reinforce the broader finding that good governance environments foster more credible and transparent corporate governance systems, even if the short-run effects of reform are initially disruptive to measured performance.

Table 7 synthesises the evidence across the ESG composite score and its three pillars, providing a coherent view of how national governance quality relates to firms' sustainability performance. The results reveal a clear and consistent pattern of governance complementarity in the cross-section: all six WGI dimensions exhibit strong and statistically significant positive associations with ESG outcomes at the between-country level. Firms headquartered in countries with durable institutional quality—characterised by effective governments, low corruption, PV, predictable legal enforcement, regulatory clarity and civic accountability—systematically display higher ESG scores.

Within-country dynamics, however, are more heterogeneous. Four enforcement-oriented dimensions—CC, GE, PV and RL—show robust positive within-country effects for the overall ESG score and for the Environmental and Social pillars. These results indicate that as national governance improves over time, firms' sustainability practices strengthen correspondingly, consistent with the complementarity hypothesis (H1a): stronger institutions lower compliance frictions, raise sanction certainty and make credible ESG investment more valuable.

Two dimensions—VA and RQ—show more complex short-run behaviour. VA is negative within for ESG, Environmental and Governance, but positive for Social, consistent with a greenwashing clean-up: greater transparency and civic scrutiny prompt firms to withdraw/verify claims and restate baselines, temporarily compressing disclosure-weighted scores. RQ displays the same transitional dip in the Environmental pillar as clearer, auditable rules tighten eligibility and reporting. In the cross-section, both VA and RQ remain strongly positive.

TABLE 6 | Worldwide governance indicators (WGI) and governance pillar score.

Variables	(1) Governance pillar score	(2) Governance pillar score	(3) Governance pillar score	(4) Governance pillar score	(5) Governance pillar score	(6) Governance pillar score
CC Dev	0.4586 (1.0050)					
CC Bar	0.8593*** (2.8724)					
GE Dev		-1.8106*** (-3.3531)				
GE Bar		0.9096*** (3.4276)				
PV Dev			-0.6348 (-1.6097)			
PV Bar			2.0534*** (8.1895)			
RL Dev				-3.5490*** (-4.7850)		
RL Bar				1.1521*** (5.4622)		
RQ Dev					0.6412 (1.1550)	
RQ Bar					1.7572*** (8.1493)	
VA Dev						-4.7400*** (-5.8548)
VA Bar						1.7024*** (10.2708)
Constant	-45.3847*** (-15.6352)	-44.9872*** (-15.4947)	-44.2582*** (-15.3273)	-45.7052*** (-15.7801)	-46.6221*** (-16.0868)	-46.2403*** (-15.9491)
Controls	Y	Y	Y	Y	Y	Y
Sector fixed effects	Y	Y	Y	Y	Y	Y
Year fixed effects	Y	Y	Y	Y	Y	Y
Observations	143,038	143,038	143,038	143,038	143,038	142,641
Number of firms	15,674	15,674	15,674	15,674	15,674	15,646
R-squared within	0.119	0.119	0.119	0.120	0.119	0.120
R-squared between	0.148	0.147	0.150	0.148	0.150	0.150
R-squared overall	0.161	0.160	0.163	0.161	0.163	0.166

Note: This table reports results from correlated random-effects (Mundlak) panel regressions of ESG governance pillar score on the worldwide governance indicators (WGI: CC, GE, PV, RL, RQ and VA). Each specification includes both the within-country/firm (demeaned) and between-country/firm (mean) components of the governance indicators and control variables. 'Dev' denotes the within-country/firm (time-demeaned) component, computed as the deviation from the country/firm-specific mean (e.g., $CC_{c(t)t} - \overline{CC}_{c(t)}$). 'Bar' denotes the between-country/firm (mean) component, that is, the country/firm-specific average over time (e.g., $\overline{CC}_{c(t)}$). All models include sector and year fixed effects. Please see Table 1 notes for variable definitions. Models are estimated by random-effects GLS with robust standard errors clustered at the firm level. Robust z-statistics are in parentheses. *** indicate significance at the 1% levels, respectively.

TABLE 7 | Summary of the relationship between WGI dimensions and ESG scores and pillars.

	ESG score	Social pillar score	Environmental pillar score	Governance pillar score
Control of corruption (CC)				
Within	+ (***)	+ (***)	+ (***)	0 (n.s.)
Between	+ (***)	+ (**)	+ (***)	+ (***)
Government effectiveness (GE)				
Within	+ (***)	+ (***)	+ (***)	− (***)
Between	+ (***)	+ (***)	+ (***)	+ (***)
Political Stab. and Abs. Violence (PV)				
Within	+ (***)	+ (***)	+ (**)	0 (n.s.)
Between	+ (***)	+ (***)	+ (***)	+ (***)
Rule of law (RL)				
Within	+ (***)	+ (***)	+ (***)	− (***)
Between	+ (***)	+ (***)	+ (***)	+ (***)
Regulatory quality (RQ)				
Within	0 (n.s.)	+ (***)	− (***)	0 (n.s.)
Between	+ (***)	+ (***)	+ (***)	+ (***)
Voice and accountability (VA)				
Within	− (***)	+ (***)	− (***)	− (***)
Between	+ (***)	+ (***)	+ (***)	+ (***)

Note: This table summarises the direction and statistical significance of coefficients from the correlated random-effects (Mundlak) regressions reported in Tables 3–6. ‘Within’ refers to the effect of the time-demeaned (within-country/firm) component (the deviation from the country/firm’s time mean). ‘Between’ refers to the effect of the country/firm-level time mean (cross-section). Entries report the sign followed by significance in parentheses: + (***) , − (**) or 0 (n.s.). A value of 0 (n.s.) indicates that the coefficient is not significant at the 10% level. All effects are conditional on the same controls, sector fixed effects and year fixed effects as in the base models; standard errors are clustered at the firm level. Significance levels: *** $p < 0.01$, ** $p < 0.05$.

For the Governance Pillar, within-country improvements in enforcement capacity—notably RL and GE—are associated with short-run declines, reflecting reassessment and stricter verification of internal controls rather than deteriorating practices. Cross-sectionally, however, all WGI dimensions are positive, indicating that strong public governance ultimately supports higher firm-level governance standards.

Overall, the results show that good country governance complements corporate sustainability across all dimensions, but that transitions towards stronger institutions often induce temporary dips in disclosure-based metrics as firms adapt to heightened expectations and more rigorous verification.

4.2 | Governance Indicators and Firm Controversies

Table 8 examines how national governance relates to firms’ ESG controversies (with higher values of the controversy fraction indicating fewer incidents). Cross-sectionally, all six WGI dimensions display negative and highly significant ‘Between’ coefficients, implying that governance-rich countries register more recorded controversies. This pattern is consistent with H2b (detection/transparency): stronger courts, cleaner administrations, clearer reporting mandates and freer civic oversight make incidents more likely to be detected, verified and publicised, so measured controversy counts rise even if underlying conduct is not worse. Within countries, over time, the signs are more varied but still dominated by detection pressures. VA and CC show significant negative

‘Within’ effects, indicating that expansions in civic scrutiny and anti-corruption enforcement increase the effective reach of audits, inspections and investigations, thereby surfacing incidents that might otherwise remain hidden. PV is similarly negative within, consistent with stability, enabling uninterrupted monitoring. GE and RL do not yield discernible short-run within effects, so no immediate inference can be drawn there.

RQ is the notable exception on the within margin: its positive ‘Within’ coefficient indicates fewer controversies immediately following improvements in RQ. This pattern is credible and fits a two-step mechanism. When RQ rises—through clearer, auditable incident-reporting rules, standardised definitions, mandatory notification thresholds and enforceable measurement, reporting and verification systems—firms must rationalise their processes, close procedural gaps and invest in preventive controls (e.g., incident escalation protocols, supplier compliance attestations, internal audits). These upgrades produce quick wins that reduce the incidence of recordable events in the near term, especially where earlier controversies stemmed from ambiguous standards, poor documentation or inconsistent thresholds that invited procedural failures. At the same time, the negative ‘Between’ association for RQ shows that in countries where high RQ is entrenched, the long-run information infrastructure ultimately raises visibility: mature taxonomies, audit trails and mandatory disclosure regimes allow authorities, investors and data providers to capture more of what happens. In short, RQ reduces controversies in the short run by removing ambiguity and forcing better internal controls, but in the long run it supports an information-rich ecosystem that detects and

records more of the residual incidents. This reconciliation is fully consistent with our prevention/detection framework: near-term compliance gains dominate within, while persistent transparency capacity dominates in levels.

4.3 | Robustness Analyses

Our results clearly show an overall positive relation between country-level governance quality (WGI) and firms' ESG performance across pillars. This pattern is consistent in all 'between' specifications and in many of the 'within' specifications. Moreover, when we pool the within and between components, the between effect largely dominates. In pooled OLS regressions that abstract from the Mundlak decomposition (not reported), WGI variables are jointly and individually positive and highly statistically significant predictors of ESG scores. When we estimate a fractional logit model for the ESG controversy ratio, the WGI coefficients are uniformly negative and highly significant, indicating fewer controversies where governance quality is higher. Taken together, these findings confirm that the between variation—persistent cross-country differences in institutions—dominates the within-country time-series response. This is not surprising given our panel's time dimension (average firm panel length ≈ 9.35 years) and broad coverage (86 countries), which limits short-run within-country institutional movement relative to large cross-sectional dispersion.

While our baseline designs (with firm and year fixed effects) mitigate endogeneity from time-invariant firm heterogeneity and common shocks, we also probe remaining concerns—most notably reverse causality and omitted time-varying country factors—through two additional strategies: (i) instrumental-variables regressions and (ii) a quasi-natural experiment around France's 2017 Sapin II Anti-Corruption Law. Across these complementary approaches, the objective is to verify that our main inferences are not an artefact of model choice or residual endogeneity.

First, we estimate instrumental-variables two stage least squares (2SLS) regressions that use lagged WGI pillars as instruments for contemporaneous WGI. Consistent with the dominance of between variation and the pooled evidence above, these IV models are estimated without the Mundlak adjustment while retaining the full control set and sector fixed effects. The identifying assumption is that past institutional quality predicts current institutional quality (relevance) and that lagged WGI are orthogonal to contemporaneous firm-level ESG shocks after conditioning on controls and sector fixed effects.

Table 9 reports the IV estimates. Across all specifications, coefficients are positive and highly significant, confirming that stronger country governance is associated with higher firm-level ESG outcomes. In Panel A (overall ESG Score), point estimates range from about 1.46 (RL) to 2.29 (PV) with large z -statistics (≈ 6), and model fit is stable ($R^2 \approx 0.29$ – 0.32). Panels B–D show the same pattern at the pillar level: effects are especially pronounced for the Social and Environmental pillars (≈ 2.60 – 5.02 and ≈ 2.71 – 4.23 , respectively), while the Governance pillar effects are positive but smaller (≈ 1.17 – 1.62).

Overidentification tests do not reject instrument validity (Sargan–Hansen p -values spanning ≈ 0.19 – 0.99).

Table 10 turns to ESG controversies using an IV fractional logit model. All six WGI pillars load negatively and highly statistically significantly on the Controversy Fraction, indicating fewer controversies where country governance is stronger. The magnitudes are economically meaningful, with the largest reductions associated with GE (-0.082) and sizable effects for VA and PV (≈ -0.047 and -0.046). RL, RQ and CC are also negative and precise (≈ -0.030 to -0.022). Wald χ^2 statistics are large, and Sargan–Hansen tests again fail to reject the instruments (p -values ≈ 0.14 – 0.64).

Taken together, the IV evidence corroborates our main message from the baseline analyses: persistent cross-country differences in governance quality (the 'between' variation) are strongly associated with higher ESG performance and fewer controversies, with especially strong transmission to the Environmental and Social dimensions and uniformly negative links to controversy incidence.

In a different approach, we study a quasi-experiment exploiting the Sapin II Law, which tightened France's anti-corruption regime from 2017 onward. This reform—strengthening anti-bribery rules, enhancing transparency and whistleblower protection, and creating the French Anti-Corruption Agency—constitutes a plausibly exogenous improvement in national governance quality. We treat this institutional change as a positive shock to country-level governance and evaluate its implications for corporate behaviour using a global difference-in-differences (DiD) design that compares French firms (treated) to firms headquartered in other countries (control) before and after 2017. The specification includes the standard firm-level controls and sector fixed effects. Our DiD focus is the overall average effect of improved governance on firms' ESG performance and controversies, providing a straightforward validation of the main findings under an alternative identification strategy.

Table 11 reports the DiD estimates. The interaction term (Treat \times Post) is positive and statistically significant for the composite ESG score ($+3.665$; $z = 4.59$) and for each pillar—Social ($+5.344$; $z = 5.14$), Governance ($+4.950$; $z = 4.71$) and Environmental ($+2.468$; $z = 2.17$)—indicating that French firms improved their ESG outcomes relative to the control group after Sapin II. For the controversy fraction, the interaction is negative (-0.208 ; $z = -2.46$). Because higher values of this index indicate fewer controversies, the negative sign implies a relative uptick in reported controversies among French firms' postreform. This pattern is consistent with heightened detection and disclosure following stronger enforcement and whistleblower protections, even as underlying ESG practices improve. These findings are consistent with Chen and Li (2024), who use a quasi-natural experiment based on China's environmental regulatory reform and show that the reform leads to improvements in firms' ESG outcomes, primarily driven by increased environmental investment.

To probe sensitivity, we ran a battery of robustness checks: (i) re-estimating all models on a large-cap subsample, (ii) on a

TABLE 8 | Worldwide governance indicators (WGI) and controversies score.

Variables	(1) Controversy fraction	(2) Controversy fraction	(3) Controversy fraction	(4) Controversy fraction	(5) Controversy fraction	(6) Controversy fraction
CC Dev	-0.2947*** (-5.2506)					
CC Bar	-0.2034*** (-5.1457)					
GE Dev		-0.0606 (-0.9269)				
GE Bar		-0.3071*** (-9.0421)				
PV Dev			-0.2965*** (-5.6525)			
PV Bar			-0.0781*** (-2.7957)			
RL Dev				-0.1106 (-1.3455)		
RL Bar				-0.4227*** (-15.8330)		
RQ Dev					0.2693*** (3.6914)	
RQ Bar					-0.4614*** (-15.7271)	
VA Dev						-0.6573*** (-6.3632)
VA Bar						-0.5255*** (-26.4032)
Company age (log) Dev	-0.2708*** (-7.6613)	-0.2845*** (-7.9791)	-0.3411*** (-9.3519)	-0.2271*** (-6.5284)	-0.2272*** (-6.5355)	-0.1659*** (-4.8716)
Company age (log) Bar	-0.0257* (-1.6518)	-0.0209 (-1.3310)	-0.0165 (-1.0563)	-0.0250 (-1.6110)	-0.0333** (-2.1754)	-0.0301** (-1.9904)
Market cap (log) Dev	-0.1326*** (-4.2380)	-0.1370*** (-4.3813)	-0.1536*** (-4.9624)	-0.1235*** (-3.8851)	-0.1267*** (-3.9724)	-0.1099*** (-3.4080)
Market cap (log) Bar	-0.7145*** (-39.1728)	-0.7078*** (-38.8318)	-0.7201*** (-39.9872)	-0.6985*** (-38.1416)	-0.6988*** (-38.1633)	-0.7088*** (-38.6183)
Employees (log) Dev	-0.0057 (-0.9395)	-0.0040 (-0.6653)	-0.0031 (-0.5248)	-0.0054 (-0.8594)	-0.0048 (-0.7691)	-0.0088 (-1.4016)
Employees (log) Bar	-0.1297*** (-11.3490)	-0.1285*** (-11.2778)	-0.1261*** (-11.5863)	-0.1353*** (-11.4637)	-0.1412*** (-11.9052)	-0.1443*** (-12.0427)
ROA Dev	0.0258*** (10.9303)	0.0259*** (10.9735)	0.0266*** (11.2120)	0.0252*** (10.7086)	0.0249*** (10.5854)	0.0248*** (10.5169)
ROA Bar	0.0445*** (11.4795)	0.0438*** (11.2905)	0.0470*** (12.0571)	0.0420*** (10.9400)	0.0428*** (11.1154)	0.0434*** (11.3514)
M/B ratio Dev	0.0269*** (5.8954)	0.0276*** (6.0725)	0.0273*** (5.9819)	0.0272*** (5.9952)	0.0276*** (6.0974)	0.0265*** (5.7810)
M/B ratio Bar	0.0366*** (6.1326)	0.0366*** (6.2180)	0.0350*** (5.8883)	0.0390*** (6.6402)	0.0399*** (6.7328)	0.0368*** (6.1918)

(Continues)

TABLE 8 | (Continued)

Variables	(1) Controversy fraction	(2) Controversy fraction	(3) Controversy fraction	(4) Controversy fraction	(5) Controversy fraction	(6) Controversy fraction
R&D (% assets) Dev	-0.0333*** (-4.5810)	-0.0331*** (-4.5589)	-0.0340*** (-4.6768)	-0.0330*** (-4.5448)	-0.0334*** (-4.5950)	-0.0331*** (-4.5490)
R&D (% assets) Bar	-0.0333*** (-4.5801)	-0.0331*** (-4.5579)	-0.0340*** (-4.6757)	-0.0330*** (-4.5438)	-0.0334*** (-4.5940)	-0.0330*** (-4.5480)
CAPEX (% assets) Dev	2.6328*** (6.1608)	2.7000*** (6.2918)	2.7408*** (6.3425)	2.5943*** (6.1056)	2.6375*** (6.2231)	2.4715*** (5.8088)
CAPEX (% assets) Bar	1.5290*** (2.9400)	1.6331*** (3.0985)	1.7095*** (3.2258)	1.3967*** (2.7017)	1.3094** (2.5422)	1.1416** (2.2248)
Quick ratio Dev	0.0008 (0.0387)	-0.0016 (-0.0763)	0.0042 (0.1940)	-0.0040 (-0.1855)	-0.0061 (-0.2860)	-0.0028 (-0.1306)
Quick ratio Bar	0.0523*** (3.4720)	0.0548*** (3.6380)	0.0545*** (3.6019)	0.0555*** (3.7251)	0.0562*** (3.7697)	0.0539*** (3.5925)
Debt-to-equity Dev	-0.0009*** (-5.4857)	-0.0009*** (-5.6367)	-0.0009*** (-5.8012)	-0.0009*** (-5.5894)	-0.0009*** (-5.5873)	-0.0009*** (-5.5175)
Debt-to-equity Bar	-0.0018*** (-9.7885)	-0.0018*** (-10.0474)	-0.0018*** (-10.0229)	-0.0017*** (-9.6870)	-0.0017*** (-9.6145)	-0.0015*** (-8.2870)
Cash (% assets) Dev	0.9761*** (4.5605)	0.9998*** (4.6838)	0.9994*** (4.7120)	1.0377*** (4.8241)	1.0550*** (4.9280)	1.0771*** (4.9053)
Cash (% assets) Bar	-0.3786* (-1.7117)	-0.3000 (-1.3526)	-0.3795* (-1.7119)	-0.2856 (-1.2926)	-0.3836* (-1.7452)	-0.3308 (-1.4819)
Constant	19.9476*** (49.0289)	19.8399*** (48.5645)	19.6561*** (48.1278)	19.9315*** (49.1098)	20.0681*** (49.5015)	20.3753*** (50.6129)
Sector fixed effects	Y	Y	Y	Y	Y	Y
Year fixed effects	Y	Y	Y	Y	Y	Y
Observations	143,028	143,028	143,028	143,028	143,028	142,631
Pseudo R-squared	0.227	0.226	0.224	0.231	0.231	0.237

Note: This table reports fractional logit coefficients from correlated random-effects (Mundlak) panel regressions of the Controversy Fraction on the worldwide governance indicators (WGI: CC, GE, PV, RL, RQ, VA) and firm covariates. The dependent variable is the Controversy Fraction (higher values indicate fewer controversies); thus, positive coefficients imply fewer controversies, and negative coefficients imply more controversies. Each specification includes both the within (demeaned) and between (mean) components of all regressors. 'Dev' denotes the within-country/firm (time-demeaned) component, computed as the deviation from the country/firm-specific mean (e.g., $CC_{c(i)t} - \overline{CC}_{c(i)}$). 'Bar' denotes the between-country/firm (mean) component, i.e., the country/firm-specific average over time (e.g., $\overline{CC}_{c(i)}$). All models include sector and year fixed effects. Please see Table 1 notes for variable definitions. Models are estimated with robust standard errors clustered at the firm level. Robust z-statistics are in parentheses. *, ** and *** indicate significance at the 10%, 5% and 1% levels, respectively.

small-cap subsample, and (iii) excluding, in turn, firms headquartered in the United States and then China. Across these variations, the qualitative conclusions remain intact—the WGI–ESG relations stay positive for scores and negative for controversies—although statistical significance attenuates in some cases, consistent with reduced sample size and changing composition. For brevity, these results are not tabulated.

5 | Conclusions, Limitations and Policy Implications

This paper documents a robust and economically meaningful association between country-level governance quality and firms' sustainability outcomes. Using a large global panel of listed firms spanning 86 countries and an average firm-level time dimension of roughly 9.35 years, we show that higher values on

the WGI are positively related to overall ESG scores and to each pillar and are negatively related to controversy incidence. Interpreting ESG performance as a form of institutionally embedded private governance provides a useful lens for understanding why these effects vary across national contexts (Brammer et al. 2010; 2012).

Decompositions that separate within- versus between-country variation reveal that the between effects—persistent cross-country institutional differences—dominate the within effects, consistent with the slow evolution of national institutions over the sample horizon. At the between-country level, the evidence is uniform: all six WGI dimensions are positively associated with ESG outcomes and negatively associated with the Controversy Fraction—by construction, an inverse metric in which lower values indicate more recorded controversies—consistent with a detection/transparency channel under

TABLE 9 | IV regression of ESG scores on worldwide governance indicators (WGI).

Variables	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Dep. Var. ESG Score						
CC	1.4882*** (6.2295)					
GE		1.8109*** (6.2352)				
PV			2.2911*** (6.2021)			
RL				1.4597*** (6.2714)		
RQ					1.6064*** (6.2084)	
VA						1.8542*** (6.0783)
Other control variables and sector fixed effects	Y	Y	Y	Y	Y	Y
Observations	78,184	78,184	78,184	78,184	78,184	77,926
R-squared	0.298	0.293	0.300	0.297	0.299	0.316
Sargan–Hansen χ^2	2.984	2.664	1.896	2.485	2.847	2.207
p-value	0.2249	0.2640	0.3875	0.2887	0.2408	0.3318
Panel B: Dep. Var. social pillar score						
CC	2.9462*** (14.1657)					
GE		4.1682*** (14.7669)				
PV			5.0161*** (14.5248)			
RL				2.6022*** (10.0129)		
RQ					2.8365*** (9.8738)	
VA						3.1949*** (9.5640)
Other control variables and sector fixed effects	Y	Y	Y	Y	Y	Y
Observations	100,255	100,190	100,255	88,605	88,605	88,320
R-squared	0.252	0.241	0.251	0.252	0.256	0.283
Sargan–Hansen χ^2	0.153	2.754	0.000	1.780	1.780	2.039
p-value	0.6958	0.2523	0.9878	0.4107	0.4106	0.3607
Panel C: Dep. Var. environmental pillar score						
CC	2.7106*** (11.0367)					
GE		3.0035*** (11.0516)				
PV			4.1036*** (10.8108)			
RL				2.9473*** (11.1077)		

(Continues)

TABLE 9 | (Continued)

Variables	(1)	(2)	(3)	(4)	(5)	(6)
RQ					3.0988*** (10.6477)	
VA						4.2259*** (9.1113)
Other control variables and sector fixed effects	Y	Y	Y	Y	Y	Y
Observations	113,247	113,247	113,247	113,247	113,247	88,343
R-squared	0.319	0.312	0.327	0.315	0.318	0.319
Sargan–Hansen χ^2	0.026	1.226	1.180	0.238	0.708	1.204
<i>p</i> -value	0.8715	0.2683	0.2773	0.6256	0.4002	0.2726
Panel D: Dep. Var. governance pillar score						
CC	1.2454*** (4.1693)					
GE		1.6192*** (4.1573)				
PV			1.1745*** (3.4869)			
RL				1.3926*** (4.1624)		
RQ					1.6147*** (4.2718)	
VA						1.3863*** (4.2148)
Other control variables and sector fixed effects	Y	Y	Y	Y	Y	Y
Observations	100,259	100,259	88,630	100,259	100,259	99,945
R-squared	0.109	0.106	0.106	0.108	0.110	0.112
Sargan–Hansen χ^2	1.354	1.211	0.548	1.243	0.432	1.680
<i>p</i> -value	0.2446	0.2712	0.4592	0.2650	0.5112	0.1949

Note: This table reports the coefficients from instrumental-variables regressions of firm-level ESG outcomes on the six worldwide governance indicators (WGI): CC, GE, PV, RL, RQ and VA. Each column treats the indicated WGI pillar as the endogenous regressor and is estimated in a separate specification; the set of instruments is the lagged values of WGI. All specifications include the same firm-level controls and year fixed effects as in the baseline models; sample sizes vary with data availability. For each panel, the reported Sargan–Hansen χ^2 statistic tests the overidentifying restrictions; the accompanying *p*-value indicates that the instruments are not rejected at conventional levels in any specification shown. Robust *z*-statistics are shown in parentheses. *** indicate significance at the 1% levels, respectively.

stronger governance. The within-country estimates, while smaller and less consistent, are broadly aligned in sign: Environmental and Social scores tend to rise with improvements in RL, CC, GE (and often PV), whereas VA frequently exhibits small or negative within effects (a scrutiny/diagnostic dynamic), RQ is often weak or insignificant and within-pillar results for Governance are muted or occasionally negative. Overall, the between effects dominate in magnitude—reflecting slow institutional change—but the within patterns largely point in the same direction.

This positive WGI–ESG relation accords with the hypothesis that stronger public institutions reduce policy uncertainty, bolster enforcement credibility and improve information quality, thereby increasing the expected payoff to substantive ESG investments and verifiable disclosure (H1a). At the same time, as governance strengthens (clearer rules, whistleblower protection, freer media, greater administrative capacity), the

probability that residual misconduct is detected and recorded rises. Because our controversies metric is inverse—the controversy fraction is lower when more controversies are recorded—governance upgrades can lead to a decline in the measured fraction even as underlying practices improve (H2b). These findings dovetail with comparative CSR research showing that corporate responsibility tends to mirror institutional strengths (Jackson and Apostolou 2010).

To address endogeneity concerns, we implement instrumental-variables models using lagged WGI as instruments and a quasi-natural experiment exploiting France’s 2017 Sapin II anti-corruption law. The IV results confirm positive effects on ESG and negative effects on controversies; the DiD evidence shows that, relative to comparable firms elsewhere, French firms increased ESG scores postreform while recorded controversies rose modestly—consistent with improved detection and whistleblower protections operating alongside

TABLE 10 | IV regression of controversies score on worldwide governance indicators (WGI).

Variables	(1) Controversy fraction	(2) Controversy fraction	(3) Controversy fraction	(4) Controversy fraction	(5) Controversy fraction	(6) Controversy fraction
CC	−0.0222*** (−8.9100)					
GE		−0.0823*** (−19.9357)				
PV			−0.0462*** (−9.6299)			
RL				−0.0298*** (−8.8053)		
RQ					−0.0313*** (−8.8521)	
VA						−0.0471*** (−14.6945)
Other control variables and year fixed effects	Y	Y	Y	Y	Y	Y
Observations	88,628	99,993	60,652	60,652	60,652	53,135
Wald chi-squared	4586	5262	4085	3654	3673	3670
Sargan–Hansen χ^2	1.623	0.214	2.145	1.790	1.676	1.854
<i>p</i> -value	0.2027	0.6439	0.1431	0.1810	0.1955	0.1734

Notes: This table reports the coefficients from instrumental-variables regressions of Controversy Fraction on the six worldwide governance indicators (WGI): CC, GE, PV, RL, RQ and VA. Each column treats the indicated WGI pillar as the endogenous regressor and is estimated in a separate specification; the set of instruments is the lagged values of WGI. All specifications include the same firm-level controls and year fixed effects as in the baseline models; sample sizes vary with data availability. For each panel, the reported Sargan–Hansen χ^2 statistic tests the overidentifying restrictions; the accompanying *p*-value indicates that the instruments are not rejected at conventional levels in any specification shown. Robust *z*-statistics are shown in parentheses. *** indicate significance at the 1% levels, respectively.

genuine upgrades in practices. This pattern aligns with cross-country findings that ESG regulatory reforms spur ESG growth (Huang et al. 2025).

Several limitations suggest caution and avenues for future work. First, measurement remains a central challenge. ESG scores embed both performance and disclosure dimensions and may respond to changes in transparency and assurance practices. Likewise, our Controversy Fraction aggregates incidents across heterogeneous domains and, crucially, does not disaggregate by the E, S and G pillars; as a result, we cannot analyse which components—environmental, social or governance—drive the controversy dynamics most strongly. This limits our ability to map mechanisms with precision and to reconcile pillar-specific improvements with incident reporting in the same domain.

Second, while we mitigate omitted variable bias through rich controls and sector fixed effects, and we probe identification with IV and DiD designs, residual endogeneity cannot be completely ruled out. The IV strategy relies on lagged WGI for relevance and exclusion; although standard diagnostics support instrument validity, lagged institutions may still correlate with slow-moving, unobserved country trends that influence firms' sustainability. The DiD analysis around a single national reform delivers intuitive evidence, but inference with one treated

country is sensitive to cross-country shock heterogeneity and potential spillovers; moreover, the observed increase in recorded controversies postreform highlights that enforcement intensity and reporting incentives can shift the mapping from true misconduct to observed incidents.

Third, our data features limited within-country institutional variation over time, which moderates the precision of within-effect estimates and may mask dynamic adjustment paths. The dominance of between effects, therefore, should not be interpreted as evidence that institutional reforms have a limited short-run impact; rather, it reflects slow institutional drift relative to substantial cross-sectional dispersion. Relatedly, our coverage is necessarily limited to listed firms with available ESG and controversy data; selection on disclosure and index membership may affect external validity, especially for small private enterprises and firms in markets with weaker reporting regimes.

Fourth, construct heterogeneity across ESG providers and across WGI subindices may introduce measurement error. While our findings are consistent across multiple specifications and robustness checks (large-cap and small-cap subsamples; samples excluding the United States and, separately, China), attenuated significance in some restricted samples underscores sensitivity to composition and power. Future work should

TABLE 11 | Difference-in-differences around France's 2017 Sapin II Anti-Corruption Law: Effects on ESG scores and controversies.

Variables	(1) ESG score	(2) Social pillar score	(3) Governance pillar score	(4) Environmental pillar score	(5) Controversy fraction
Treat	4.8846*** (4.6028)	6.3151*** (4.7633)	−5.5306*** (−4.1716)	12.9240*** (9.8270)	0.1823*** (3.3264)
Post	11.5866*** (86.3527)	13.9595*** (86.4593)	5.8193*** (34.1641)	11.2370*** (56.8271)	0.1770*** (12.0337)
Treat*post	3.6648*** (4.5920)	5.3441*** (5.1422)	4.9502*** (4.7110)	2.4684** (2.1742)	−0.2082** (−2.4552)
Company age (log)	5.7934*** (39.0173)	5.7273*** (32.8223)	3.4789*** (19.7611)	7.0819*** (35.1010)	−0.0311*** (−4.3622)
Market cap (log)	4.9151*** (56.3181)	5.1381*** (51.1020)	3.1964*** (35.2758)	6.7715*** (54.8678)	−0.6734*** (−84.1802)
Employees (log)	0.3603*** (12.3292)	0.3400*** (9.1190)	0.4029*** (12.9224)	0.4334*** (11.2648)	−0.0849*** (−19.0425)
ROA	−0.1448*** (−15.2693)	−0.1567*** (−14.0205)	−0.0582*** (−5.0979)	−0.2196*** (−15.6070)	0.0401*** (26.1165)
M/B ratio	−0.4977*** (−23.1512)	−0.4869*** (−20.2947)	−0.3426*** (−13.1302)	−0.7282*** (−23.7398)	0.0494*** (18.3119)
R&D (% total assets)	0.0744*** (3.4005)	0.1145*** (3.3545)	−0.0072 (−0.2866)	0.0822*** (9.8507)	−0.0134*** (−3.7773)
CAPEX (% total assets)	−18.7806*** (−13.6055)	−16.9828*** (−10.6495)	−11.8172*** (−6.6195)	−31.9125*** (−17.6408)	2.3802*** (9.9718)
Quick ratio	−0.6809*** (−9.6245)	−0.6310*** (−7.6633)	−0.3432*** (−3.8549)	−1.2059*** (−12.6009)	0.0623*** (8.0463)
Debt-to-equity	0.0157*** (16.9897)	0.0147*** (14.0038)	0.0108*** (9.3967)	0.0259*** (20.7896)	−0.0016*** (−22.2541)
Cash (% assets)	−3.6973*** (−5.9487)	−3.6020*** (−4.7942)	−4.5247*** (−5.5723)	−3.9979*** (−4.3960)	−0.0117 (−0.1107)
Constant	−91.9031*** (−37.8582)	−98.5601*** (−34.5949)	−37.3828*** (−16.0645)	−144.3061*** (−42.5238)	17.5288*** (99.6113)
Sector fixed effects	Y	Y	Y	Y	Y
Observations	144,397	144,387	144,397	144,387	144,387
Number of firms	15,910	15,910	15,910	15,910	15,910
R-squared	0.301	0.249	0.131	0.347	0.206

Note: This table reports results from a global difference-in-differences (DiD) design that exploits the 2017 implementation of France's Sapin II anti-corruption law as an exogenous, country-level governance shock. Treat equals 1 for French firms and 0 for firms headquartered in all other countries; Post equals 1 for years 2017 and after and 0 otherwise. The coefficient on Treat×Post is the DiD estimator—the average treatment effect on treated French firms relative to the control group after the reform. Columns report separate regressions for the ESG composite score, the three ESG pillar scores, and the Controversy Fraction (higher values indicate fewer controversies; thus, positive coefficients imply fewer controversies, and negative coefficients imply more controversies). All specifications include the firm-level controls described in Table 1 and sector fixed effects. Robust z-statistics are shown in parentheses. Coefficients and z-statistics in bold correspond to the interaction term Treat × Post, which captures the difference-in-differences treatment effect. *, ** and *** indicate significance at the 10%, 5% and 1% levels, respectively.

triangulate with alternative sustainability measures (e.g., emissions, verified safety incidents, litigation-based governance infractions), text-based disclosure audits and assurance quality indicators.

The results carry clear implications for policymakers, regulators and market participants. At a high level, stronger national governance systems are associated with higher corporate ESG

performance and, in the short run, more recorded controversies. The patterns are most pronounced for dimensions linked to enforcement and administrative capacity. Strengthening the RL, improving GE, tightening anti-corruption regimes and stabilising the policy environment reduce compliance frictions, sharpen incentives and provide credible discipline for firms. Evidence from France's Sapin II reform illustrates the mechanism: combining clear prohibitions, active oversight (a dedicated anti-corruption

agency) and whistleblower protection can raise ESG scores while initially increasing reported incidents as detection improves. Policymakers should therefore anticipate and communicate that an early rise in reported cases may signal better governance rather than deteriorating behaviour, and design monitoring systems that distinguish detection dynamics from underlying conduct. This policy implication accords with a central result in institutional economics: durable improvements in the institutional framework reshape private incentives and, consequently, economic outcomes (North 1990; Rodrik et al. 2004).

For regulators, the findings support continued investment in the institutional infrastructure that underpins credible ESG reporting and enforcement. This includes resourcing supervisory agencies, harmonising disclosure standards and incentivising independent assurance. Given that controversies respond not only to conduct but also to visibility, coherent whistleblower frameworks and safe reporting channels can improve information quality and accelerate remediation. For standard setters and data providers, the absence of controversy disaggregation by E, S and G is a material limitation; developing pillar-specific controversy taxonomies and severity metrics would significantly enhance research and policy evaluation.

For firms, the results suggest that improvements in the institutional environment amplify the returns to genuine sustainability investments. Managers should expect heightened scrutiny following governance upgrades and proactively strengthen compliance systems, internal controls, supply-chain due diligence and third-party assurance to convert institutional improvements into durable ESG gains. For global investors, country-level governance indicators should be treated as systematic ESG risk factors: portfolio construction, engagement strategies and cost-of-capital models can be improved by explicitly integrating institutional quality and its trajectory. Importantly, investors should interpret short-run increases in reported controversies around enforcement reforms with nuance, distinguishing detection effects from changes in underlying risk.

In sum, national governance quality is a powerful correlate of corporate sustainability outcomes. Convergent evidence from a Mundlak (within-between) model, an instrumental-variables strategy, and a DiD design around a policy shock points in the same direction. Addressing remaining limitations—especially disaggregating controversy data and adding hard outcome measures—will sharpen mechanism tests and strengthen policy design. Even so, the current results already support a practical message: investments in public governance complement private sustainability efforts and yield measurable improvements in firms' ESG profiles. More broadly, this aligns with the view that institutional context is first-order in shaping corporate behaviour (North 1990).

Acknowledgements

The author gratefully acknowledges financial support from FCT—Portuguese Foundation for Science and Technology under Grant UIDB/04007/2020. Open access publication funding provided by FCT (b-on).

Conflicts of Interest

The author declares no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are available from Refinitiv/LSEG. Restrictions apply to the availability of these data, which were used under license for this study. Data are available from the author(s) with the permission of Refinitiv/LSEG.

References

- Aguilera, R. V., and G. Jackson. 2010. "Comparative and International Corporate Governance." *Academy of Management Annals* 4, no. 1: 485–556. <https://doi.org/10.5465/19416520.2010.495525>.
- Aguilera, R. V., D. E. Rupp, C. A. Williams, and J. Ganapathi. 2007. "Putting the S Back in Corporate Social Responsibility: A Multilevel Theory of Social Change in Organizations." *Academy of Management Review* 32, no. 3: 836–863. <https://doi.org/10.5465/amr.2007.25275678>.
- Amiram, D., Z. Bozanic, J. D. Cox, Q. Dupont, J. M. Karpoff, and R. Sloan. 2018. "Financial Reporting Fraud and Other Forms of Misconduct: A Multidisciplinary Review of the Literature." *Review of Accounting Studies* 23, no. 2: 732–783. <https://doi.org/10.1007/s11142-017-9435-x>.
- Arndt, C., and C. Oman. 2006. *Uses and Abuses of Governance Indicators, Development Centre Studies*. OECD Publishing. <https://doi.org/10.1787/9789264026865-en>.
- Becker, G. S. 1968. "Crime and Punishment: An Economic Approach." *Journal of Political Economy* 76, no. 2: 169–217. <https://www.jstor.org/stable/1830482>.
- 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>.
- Besley, T., and A. Prat. 2006. "Handcuffs for the Grabbing Hand? Media Capture and Government Accountability." *American Economic Review* 96, no. 3: 720–736. <https://doi.org/10.1257/aer.96.3.720>.
- Brammer, S., G. Jackson, and D. Matten. 2012. "Corporate Social Responsibility and Institutional Theory: New Perspectives on Private Governance." *Socio-Economic Review* 10, no. 1: 3–28. <https://doi.org/10.1093/ser/mwr030>.
- Campbell, J. L. 2007. "Why Would Corporations Behave in Socially Responsible Ways? An Institutional Theory of Corporate Social Responsibility." *Academy of Management Review* 32, no. 3: 946–967. <https://doi.org/10.5465/amr.2007.25275684>.
- Chatterji, A. K., R. Durand, D. I. Levine, and S. Touboul. 2016. "Do Ratings of Firms Converge? Implications for Managers, Investors and Strategy Researchers." *Strategic Management Journal* 37, no. 8: 1597–1614. <https://doi.org/10.1002/smj.2407>.
- Chen, Q., and M. Li. 2024. "Environmental Regulatory System Reform and Corporate ESG Ratings: Evidence From China." *Economic Modelling* 135: 106710. <https://doi.org/10.1016/j.econmod.2024.106710>.
- Cheng, B., I. Ioannou, and G. Serafeim. 2014. "Corporate Social Responsibility and Access to Finance." *Strategic Management Journal* 35, no. 1: 1–23. <https://doi.org/10.1002/smj.2131>.
- Crifo, P., and V. D. Forget. 2015. "The Economics of Corporate Social Responsibility: A Firm-Level Perspective Survey." *Journal of Economic Surveys* 29, no. 1: 112–130. <https://doi.org/10.1111/joes.12055>.
- Cuervo-Cazurra, A. 2006. "Who Cares about Corruption?" *Journal of International Business Studies* 37, no. 6: 807–822. <https://doi.org/10.1057/palgrave.jibs.8400223>.
- Daske, H., L. Hail, C. Leuz, and R. Verdi. 2008. "Mandatory IFRS Reporting Around the World: Early Evidence on the Economic Consequences." *Journal of Accounting Research* 46, no. 5: 1085–1142. <https://doi.org/10.1111/j.1475-679X.2008.00306.x>.
- Dhaliwal, D., O. Z. Li, A. Tsang, and Y. G. Yang. 2014. "Corporate Social Responsibility Disclosure and the Cost of Equity Capital: The Roles of

- Stakeholder Orientation and Financial Transparency.” *Journal of Accounting and Public Policy* 33, no. 4: 328–355. <https://doi.org/10.1016/j.jaccpubpol.2014.04.006>.
- Djankov, S., R. La Porta, F. Lopez-de-Silanes, and A. Shleifer. 2008. “The Law and Economics of Self-Dealing.” *Journal of Financial Economics* 88, no. 3: 430–465. <https://doi.org/10.1016/j.jfineco.2007.02.007>.
- Dyck, A., A. Morse, and L. Zingales. 2010. “Who Blows the Whistle on Corporate Fraud?” *Journal of Finance* 65, no. 6: 2213–2253. <https://doi.org/10.1111/j.1540-6261.2010.01614.x>.
- Easley, D., and M. O’hara. 2004. “Information and the Cost of Capital.” *Journal of Finance* 59, no. 4: 1553–1583. <https://doi.org/10.1111/j.1540-6261.2004.00672.x>.
- Eccles, R. G., and J. Strohle, Exploring Social Origins in the Construction of ESG Measures (July 12, 2018). Available at SSRN: <http://dx.doi.org/10.2139/ssrn.3212685>.
- Eliwa, Y., A. Aboud, and A. Saleh. 2021. “ESG Practices and the Cost of Debt: Evidence From EU Countries.” *Critical Perspectives on Accounting* 79: 102097. <https://doi.org/10.1016/j.cpa.2019.102097>.
- Ferraz, C., and F. Finan. 2008. “Exposing Corrupt Politicians: The Effects of Brazil’s Publicly Released Audits on Electoral Outcomes.” *Quarterly Journal of Economics* 123, no. 2: 703–745. <https://doi.org/10.1162/qjec.2008.123.2.703>.
- Flammer, C. 2013. “Corporate Social Responsibility and Shareholder Reaction: The Environmental Awareness of Investors.” *Academy of Management Journal* 56, no. 3: 758–781. <https://doi.org/10.5465/amj.2011.0744>.
- Garmaise, M. J., and J. Liu. 2005. Corruption, Firm Governance, and the Cost of Capital. <https://doi.org/10.2139/ssrn.644017>.
- Hall, P. A., and D. Soskice. 2001. *Varieties of Capitalism: The Institutional Foundations of Comparative Advantage*. Oxford Academic. <https://doi.org/10.1093/0199247757.001.0001>.
- Healy, P. M., and K. G. Palepu. 2001. “Information Asymmetry, Corporate Disclosure, and the Capital Markets: A Review of the Empirical Disclosure Literature.” *Journal of Accounting and Economics* 31, no. 1–3: 405–440. [https://doi.org/10.1016/S0165-4101\(01\)00018-0](https://doi.org/10.1016/S0165-4101(01)00018-0).
- Healy, P. M., and G. Serafeim. 2020. “Voluntary, Self-Regulatory, and Mandatory Disclosure of Oil and Gas Company Payments to Foreign Governments.” *Accounting Horizons* 34, no. 1: 111–129. <https://doi.org/10.2308/horizons-17-133>.
- Houqe, M. N., K. Ahmed, and G. Richardson. 2020. “The Effect of Environmental, Social, and Governance Performance Factors on Firms’ Cost of Debt: International Evidence.” *International Journal of Accounting* 55, no. 3: 2050014. <https://doi.org/10.1142/S1094406020500146>.
- Houqe, M. N., and R. M. Monem. 2016. “IFRS Adoption, Extent of Disclosure, and Perceived Corruption: A Cross-Country Study.” *International Journal of Accounting* 51, no. 3: 363–378. <https://doi.org/10.1016/j.intacc.2016.07.002>.
- Huang, R., H. Cui, and X. Sun. 2025. “Legal Determinants of ESG Growth: Cross-National Evidence From Regulatory Reforms.” *Finance Research Letters* 86: 108396. <https://doi.org/10.1016/j.frl.2025.108396>.
- Ioannou, I., and G. Serafeim. 2012. “What Drives Corporate Social Performance? The Role of Nation-Level Institutions.” *Journal of International Business Studies* 43, no. 9: 834–864. <https://doi.org/10.1057/jibs.2012.26>.
- Jackson, G., and A. Apostolou. 2010. “Corporate Social Responsibility in Western Europe: An Institutional Mirror or Substitute?” *Journal of Business Ethics* 94, no. 3: 371–394. <https://doi.org/10.1007/s10551-009-0269-8>.
- Jackson, G., and R. Deeg. 2008. “Comparing Capitalisms: Understanding Institutional Diversity and Its Implications for International Business.” *Journal of International Business Studies* 39, no. 4: 540–561. <https://doi.org/10.1057/palgrave.jibs.8400375>.
- Karpoff, J. M., J. R. Lott, Jr., and E. W. Wehrly. 2005. “The Reputational Penalties for Environmental Violations: Empirical Evidence.” *Journal of Law and Economics* 48, no. 2: 653–675. <https://doi.org/10.1086/430806>.
- Kaufmann, D., and A. Kraay. 2007. “Governance Indicators: Where Are We, Where Should We Be Going?” *World Bank Research Observer* 23, no. 1: 1–30. <http://documents.worldbank.org/curated/en/260351468179971208>.
- Kaufmann, D., and A. Kraay. 2024. *The Worldwide Governance Indicators (WGI) Project: 2024 Update*. The World Bank. <https://documents.worldbank.org/en/publication/documents-reports/documentdetail/099005210162424110>.
- Kaufmann, D., A. Kraay, and M. Mastruzzi. 2011. “The Worldwide Governance Indicators: Methodology and Analytical Issues.” *Hague Journal on the Rule of Law* 3, no. 2: 220–246. <https://doi.org/10.1017/S1876404511200046>.
- Khan, M., G. Serafeim, and A. Yoon. 2016. “Corporate Sustainability: First Evidence on Materiality.” *Accounting Review* 91, no. 6: 1697–1724. <https://doi.org/10.2308/accr-51383>.
- Kotsantonis, S., C. Pinney, and G. Serafeim. 2016. “ESG Integration in Investment Management: Myths and Realities.” *Journal of Applied Corporate Finance* 28, no. 2: 10–16. <https://doi.org/10.1111/jacf.12169>.
- Krüger, P. 2015. “Corporate Goodness and Shareholder Wealth.” *Journal of Financial Economics* 115, no. 2: 304–329. <https://doi.org/10.1016/j.jfineco.2014.09.008>.
- Leuz, C., D. Nanda, and P. D. Wysocki. 2003. “Earnings Management and Investor Protection: An International Comparison.” *Journal of Financial Economics* 69, no. 3: 505–527. [https://doi.org/10.1016/S0304-405X\(03\)00121-1](https://doi.org/10.1016/S0304-405X(03)00121-1).
- Liang, H., and L. Renneboog. 2017. “On the Foundations of Corporate Social Responsibility.” *Journal of Finance* 72, no. 2: 853–910. <https://doi.org/10.1111/jofi.12487>.
- Matten, D., and J. Moon. 2008. “Implicit” and ‘Explicit’ CSR: A Conceptual Framework for a Comparative Understanding of Corporate Social Responsibility.” *Academy of Management Review* 33, no. 2: 404–424. <https://doi.org/10.5465/amr.2008.31193458>.
- Meng, X. H., S. X. Zeng, C. M. Tam, and X. D. Xu. 2014. “Whether Top Executives’ Turnover Influences Environmental Responsibility: From the Perspective of Institutional Theory.” *Journal of Business Ethics* 122, no. 4: 535–549. <https://doi.org/10.1007/s10551-012-1351-1>.
- Moussa, T., N. Krasnikova, M. Postoyeva, A. Hawas, and A. Helfaya. 2026. “Sustainability Performance, Cost of Debt, and Institutional Environment Quality: Global Evidence From the Hospitality and Tourism Industry.” *Business Strategy and the Environment* 35, no. 2: 1712–1734. <https://doi.org/10.1002/bse.70230>.
- Mundlak, Y. 1978. “On the Pooling of Time Series and Cross Section Data.” *Econometrica* 46, no. 1: 69–85. <https://doi.org/10.2307/1913646>.
- North, D. C. 1990. *Institutions, Institutional Change, and Economic Performance*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511808678>.
- Olken, B. A. 2007. “Monitoring Corruption: Evidence From a Field Experiment in Indonesia.” *Journal of Political Economy* 115, no. 2: 200–249. <https://doi.org/10.1086/517935>.
- Papke, L. E., and J. M. Wooldridge. 1996. “Econometric Methods for Fractional Response Variables With an Application to 401(k) Plan Participation Rates.” *Journal of Applied Econometrics* 11, no. 6: 619–632. [https://doi.org/10.1002/\(SICI\)1099-1255\(199611\)11:6<619::AID-JAE418>3.0.CO;2-1](https://doi.org/10.1002/(SICI)1099-1255(199611)11:6<619::AID-JAE418>3.0.CO;2-1).
- La Porta, R., F. Lopez-de-Silanes, A. Shleifer, and R. W. Vishny. 1998. “Law and Finance.” *Journal of Political Economy* 106, no. 6: 1113–1155. <https://doi.org/10.1086/250042>.

- Reinikka, R., and J. Svensson. 2004. "Local Capture: Evidence From a Central Government Transfer Program in Uganda." *Quarterly Journal of Economics* 119, no. 2: 679–705. <https://doi.org/10.1162/0033553041382120>.
- Rodrik, D., A. Subramanian, and F. Trebbi. 2004. "Institutions Rule: The Primacy of Institutions over Geography and Integration in Economic Development." *Journal of Economic Growth* 9, no. 2: 131–165. <https://doi.org/10.1023/B:JOEG.0000031425.72248.85>.
- Shleifer, A., and R. W. Vishny. 1997. "A Survey of Corporate Governance." *Journal of Finance* 52, no. 2: 737–783. <https://doi.org/10.1111/j.1540-6261.1997.tb04820.x>.
- Sullivan, R. and Mackenzie, C., eds. 2006. *Responsible Investment*, 1st ed. Routledge. <https://doi.org/10.4324/9781351283441>.
- Wang, H., L. Tong, R. Takeuchi, and G. George. 2016. "Corporate Social Responsibility: An Overview and New Research Directions: Thematic Issue on Corporate Social Responsibility." *Academy of Management Journal* 59, no. 2: 534–544. <https://doi.org/10.5465/amj.2016.5001>.
- Whitley, R. 1999. *Divergent Capitalisms: The Social Structuring and Change of Business Systems*. Oxford University Press.