Are Developing Country Firms Getting the Short End of the ESG Stick?

Jairaj Gupta
University of York, United Kingdom
jairaj.gupta@york.ac.uk

R. Shruti
Indian Institute of Management Kozhikode, India
shruti@iimk.ac.in

Xia Li
Aston University, United Kingdom
x.li40@aston.ac.uk

Amjad Ali
University of York, United Kingdom
amjad.ali@york.ac.uk

January 5, 2026

Abstract

Policymakers in emerging economies are increasingly concerned that global ESG-scoring firms based in developed countries are "unfairly punishing" their companies by assigning them lower ESG scores. Using Refinitiv—the most comprehensive global ESG database—this study shows that firms in developing countries have ESG scores that are about 16% lower (relative to the standard deviation) than those of firms in developed countries. Further analysis indicates that this "*score gap*" stems from features of the ESG scoring process itself rather than from any systematic "*rating bias*". Specifically, the "*score gap*" is driven by the implicit global industry-based benchmarking assumptions embedded in Refinitiv's ESG rating methodology, which fails to adequately account for domestic priorities in emerging markets and thus systematically disadvantages their firms. When this issue is addressed using our proposed Modified ESG scores—which benchmark firms jointly by country and industry, thereby isolating within-country, within-industry variation—the score differences become statistically insignificant.

Keywords: Institutional Theory; Theory of Human Needs; ESG; Developing Country; Rating Bias; Score Gap

JEL Codes: D82; G24; I31; O16; P48

1. Introduction

The introduction of the United Nations Sustainable Development Goals (SDGs) in 2015 has put sustainability on centre stage, gaining widespread prominence, particularly in response to the Global Financial Crisis of 2007–08, growing climate concerns, and the recent COVID-19 pandemic.[1] This renewed emphasis on sustainability has spurred the growth and demand of ESG rating providers such as Refinitiv, Sustainalytics, MSCI, and Bloomberg. These ESG rating firms quantify a company's environmental, social, and governance (ESG) performance and sell their ESG scores to various users. However, the lack of transparency and a standardised framework have cast doubt on the accuracy of these scores, leading investors, businesses, and regulators to question their credibility (Larcker, Tayan, and Watts 2022).[2] Adding to this controversy is a deepening apprehension that ESG raters may be intentionally assigning lower scores to companies in developing countries.[3]

If true, this development is concerning, as lower ESG scores can have several negative consequences, including: (i) deterring capital investments from sustainability-focused investors (Hartzmark and Sussman 2019); (ii) increasing financing costs for developing country firms, as lenders demand higher returns due to the perceived risks of poor ESG performance (Ng and Rezaee 2015; Apergis, Poufinas, and Antonopoulos 2022); (iii) heightened regulatory and compliance risks, with firms potentially facing penalties, trade restrictions, legal liabilities, or challenges in meeting evolving sustainability standards (Asante-Appiah and Lambert 2023); (iv) discouraging customers who prioritise ethical sourcing and sustainability (Parguel, Benoît-Moreau, and Larceneux 2011); (v) disrupting supply chains, as multinational corporations increasingly require ESG-compliant suppliers (Pagell, Wu, and Wasserman 2010; Schmidt, Foerstl, and Schaltenbrand 2017); (vi) making talent acquisition and retention more difficult, as skilled workers tend to prefer companies with strong ESG commitments (Bhattacharya, Sen, Edinger-Schons, and Neureiter 2023); and (vii) eroding competitive advantage, as firms with weak ESG ratings struggle to compete against more sustainable brands (Porter and Kramer 2006). Moreover, if sustainability efforts by developing country firms under challenging economic conditions aren't properly recognised, firms may be discouraged from adopting sustainable business practices (IMF 2021).

---

[1] See, ESG awareness is an enduring legacy of the global financial crisis. Accessed on July 1, 2024.
[2] See, Sustainability Institute, ERM (2023). "Rate the Raters 2023: ESG Ratings at a Crossroads." Accessed on January 19, 2024.
[3] See, ESG scoring 'unfairly punishing' emerging economies. Accessed on January 19, 2024.

Thus, in this study, we investigate the "*apprehension*", whether global ESG-scoring firms based in developed countries are "unfairly punishing" developing country firms by assigning them lower ESG scores compared to their developed counterparts.

One plausible explanation is that the global ESG scoring agencies, headquartered mainly in the United States (U.S.) and Europe, do not adequately consider the cultural and contextual distinctions between developed and developing nations.[4] For instance, employment generation, community development, and access to basic services that improve social equity could be more critical than environmental goals in developing countries. Thus, a failure to capture these nuances may adversely influence ESG scores of firms in developing countries (Pinto 2024; SEBI 2022, 2023). While these apprehensions have been examined in the context of credit ratings and corporate governance ratings (Almeida, Cunha, Ferreira, and Restrepo 2017; De Moor, Luitel, Sercu, and Vanpée 2018; Black, de Carvalho, Khanna, Kim, and Yurtoglu 2017), this study marks the first attempt to explain why firms in developing countries may receive lower ESG scores. Moreover, it is crucial to investigate whether firms in developing countries perform poorly on sustainability criteria due to weaker institutional settings and national priorities, or whether their lower scores are partly attributable to issues in the ESG rating methodologies or biases in ESG rating agencies towards developing country firms. This is important as investors, regulators, think tanks, industry associations, NGOs, media, and research bodies worldwide are increasingly relying on commercial ESG ratings to assess corporate sustainability (European Commission 2021).

To address this question, we first integrate Powell and DiMaggio's (1991) institutional theory with the theory of human needs (Maslow 1943, 1954; Deci and Ryan 1985; Ryan and Deci 2017) and the Environmental Kuznets Curve (EKC) hypothesis (Grossman and Krueger 1991; Shafik and Bandyopadhyay 1992; Panayotou 1993) to explain why ESG scores are lower in developing countries. The institutional theory highlights how differences in political systems, labour and education systems, national cultures, and legal origins shape corporate social performance (Ioannou and Serafeim 2012; Liang and Renneboog 2017). However, this theory alone cannot fully explain the variations in ESG scores between developed and developing countries. From the perspective of the theory of human needs, these variations may also arise because developing countries often prioritise fundamental needs, such as poverty alleviation and infrastructure development, over higher-level needs. According to the EKC hypothesis, focusing on improving basic infrastructure, such as transportation and energy, to

[4] See, UNCTAD (2023). Financing for sustainable development report 2023. Accessed on November 16, 2024.

increase industrial output inevitably leads to rampant pollution and rapid resource depletion. This reinforces the idea that countries prioritise economic stability over environmental sustainability (Grossman and Krueger 1995; Stern, Common, and Barbier 1996). Hence, such differences in SDG priorities result in varying ESG priorities in policymaking, which subsequently impact firm-level sustainability practices and disclosures.

Therefore, we contend that the ESG scores as reported by the rating agencies, which largely rely on the volume of voluntary disclosures (Raghunandan and Rajgopal 2022), are lower for firms in developing countries because these companies may fail to provide sufficient or quality information on key scoring criteria required by ESG raters—due to differing national priorities and/or poorer institutional settings.

To test the above prediction, we use a global sample of non-financial firms and classify them as either developing or developed based on the International Monetary Fund (IMF) designation of the country where the firms are incorporated. Accordingly, the *DVPG* dummy takes a value of one if the firm is from a developing country, else zero.[5] We obtain the ESG of global firms and the other financial information from the LSEG workspace (formerly Refinitiv Eikon) database. We consider the ESG scores from Refinitiv because, unlike its U.S./Western-centric competitors, it has the most exhaustive and longest historical coverage of global non-financial firms, offers methodological transparency and is widely used in the empirical literature (Drempetic, Klein, and Zwergel 2020; Basu, Vitanza, Wang, and Zhu 2022).

Our sample for this study involves 7,969 listed firms from 2009 to 2023 across 46 countries. Since the firms in developed and developing countries differ substantially in firm-level characteristics, we employed an entropy-balanced sample for our regression analysis with industry and year-fixed effects (Hainmueller 2012).[6] In line with our predictions, panel regression analysis reveals that the ESG scores of firms in developing countries are lower than in developed countries. In terms of economic significance, the scores are lower by about 16% relative to the standard deviation, confirming the "*score gap*". This finding aligns with the predictions of the institutional theory. Notably, the environmental (*ENV*) and social (*SOC*) pillar scores exhibit a similar pattern as well. Furthermore, our empirical analysis suggests that endogeneity issues, particularly those arising from reverse causality, are unlikely. Entropy balancing addresses endogeneity arising from self-selection (see Hainmueller 2012, Shroff,

[5] Empirical analysis show that the *DVPG* dummy explains about 90% of the variations in the institutional differences between developed and developing countries. This confirms that the *DVPG* dummy is a robust proxy for institutional differences across countries. See section 4.1 for detailed analysis and discussion.

[6] We are unable to use firm-fixed effects or country-fixed effects since they are perfectly correlated with the *DVPG* dummy, which would result in the *DVPG* dummy being eliminated in the regressions due to multicollinearity.

Verdi, and Yost 2017), and we perform additional tests to tackle endogeneity concerns related to omitted variables (Oster's Coefficient Stability Test (Oster 2019)) and measurement errors. Test results confirm that our results are robust to endogeneity concerns.

While the differences outlined above are in line with the institutional theory, the theory of human needs and the EKC hypothesis, the possibility of the scores being lower due to institutional biases among ESG raters and measurement challenges or actual poor performance cannot be ruled out. Notably, the institutional bias observed in credit ratings (Borensztein, Cowan, and Valenzuela 2013; Almeida *et al.* 2017) and governance ratings (Doidge, Andrew Karolyi, and Stulz 2007; Black *et al.* 2017) for firms in developing countries may extend to ESG scores. Several factors may exacerbate the measurement problems when scoring firms in developing countries, including: (i) reliance on sustainability frameworks designed for developed markets (Berg, Kölbel, and Rigobon 2022); (ii) lack of institutional familiarity due to geographical distance between firms in developing countries and rating agencies (Cai, Pan, and Statman 2016); and (iii) lack of transparency in rating methodologies, data availability, and disclosure biases (Drempetic *et al.* 2020; Larcker *et al.* 2022; Raghunandan and Rajgopal 2022).[7] Therefore, we argue that these shortcomings could adversely impact the ESG scores of developing country firms even after accounting for firm-level, industry-level, and country-level differences.

To test this hypothesis, we first obtain data pertaining to the 186 indicators used by Refinitiv to construct their *ESG* scores.[8] Liang and Renneboog (2017) and Basu *et al.* (2022) highlight that the scores are constructed using a global benchmarking approach in which firms are compared against their industry peers. As such, this approach fails to capture the country-specific institutional differences and national priorities in the ESG scores (Basu *et al*. 2022). We refine the Refinitiv *ESG* scores by benchmarking against both country and industry peers when deriving category scores and weights. By doing so, the modified ESG (*MESG*) scores address the methodological limitation in Refinitiv's scoring process. Our regression results indicate that the *MESG* scores are not significantly different between developed and developing countries. This finding confirms that developing country firms are receiving lower ESG scores from the rating agencies, not due to an inherent "*rating bias*", but due to the implicit

---

[7] Data availability bias means the tendency of larger firms to generate larger volumes of ESG data to receive higher ESG scores regardless of industry (Drempetic *et al.* 2020). Data disclosure bias means the tendency of firms disclosing more ESG-related information to receive higher ESG scores regardless of whether their activities are impactful or not (Raghunandan and Rajgopal 2022).

[8] Upon request and subject to a non-disclosure agreement, the London Stock Exchange Group (LSEG) provided access to the list of indicators, together with their polarity and industry relevance.

benchmarking assumptions, disadvantaging the developing country firms. We observe a similar pattern in the modified component scores, particularly *MENV* and *MSOC*.

While this analysis confirms the existence of an ESG *score gap*, it is possible that the lower scores could be due to poor disclosures rather than methodological issues. This is because *ESG* scores have been criticised for capturing disclosure volume irrespective of the substantiveness of the disclosed information (Raghunandan and Rajgopal 2022). To rule out this possibility, we perform a staggered Difference-in-Differences analysis around the passage of mandatory ESG disclosures around the world (Krueger, Sautner, Tang, and Zhong 2024). Our findings confirm that the measurement issues in the construction of *ESG* scores, rather than poor disclosures in developing countries, are the source of the lower *ESG* scores in developing country firms.

To ensure the robustness of our findings, we conduct several additional tests. First, incorporating additional fixed effects like *Industry*×*Year* does not change our findings. Second, using the United Nations and World Bank classification for developed and developing countries, we confirm the consistency of our results. Third, a sample excluding dominant countries also provides similar results, confirming that our inferences are based on a group dynamic arising from the developed-developing classification rather than on individual country-specific attributes.

Overall, we find compelling empirical evidence that the ESG scores of firms in developing countries are lower than those of firms in developed countries, not only due to a poorer institutional environment and varying national priorities but also due to measurement-driven limitations among ESG raters. Further, by developing the *MESG* score that accounts for the limitations stemming from Refinitiv's global benchmarking process, we suggest that rating agencies may consider our approach to correct such issues and provide ESG scores that are reflective of varying institutional settings and national priorities. We also encourage regulators and investors to consider using *MESG* scores in cross-country assessments of firms' sustainability performance and in informing portfolio allocation decisions.

Overall, our findings contribute to two strands of literature. First, we extend research on how the institutional environment shapes sustainable business practices. While prior studies (Ioannou and Serafeim 2012; Cai et al. 2016; Liang and Renneboog 2017), predominantly draw on institutional theory (Powell and DiMaggio 1991) to explain cross-country ESG performance, we incorporate insights from the theory of human needs (Maslow 1943, 1954; Deci and Ryan 1985; Ryan and Deci 2017) and the Environmental Kuznets Curve (Grossman and Krueger 1991; Shafik and Bandyopadhyay 1992; Panayotou 1993). In doing so, we emphasise that cross-country ESG disparities reflect group dynamics between developed and developing economies, rather than single institutional factors (Leuz, Nanda, and Wysocki 2003).

Second, we contribute to the literature on ESG measurement challenges. Prior studies highlight issues such as size bias (Drempetic *et al*. 2020), disclosure bias (Raghunandan and Rajgopal 2022), score divergence across agencies (Chatterji, Durand, Levine, and Touboul 2016; Berg *et al*. 2022), and lack of standardised disclosures (Christensen, Serafeim, and Sikochi 2022). In response to these issues and similar concerns, the European Union has adopted the regulation (EU) 2024/3005 to improve the "transparency and integrity" of ESG ratings and data providers, applicable from July 2026, placing providers under the supervision of the European Securities and Markets Authority (ESMA) (OJEU 2024). Building on this discussion, we demonstrate how applying a global rating framework systematically disadvantages firms in developing countries.

2. Relevant Literature, Theoretical Background and Hypothesis Development

2.1. Evolution and Significance of ESG Scores

With the growing importance of non-financial information in investors' decision-making process, ESG factors began capturing investors' attention in the 1990s, leading to the rise of socially responsible investing (SRI). However, early ESG scores, often based on self-reported data, lacked standardisation and transparency (Michelson, Wailes, Van Der Laan, and Frost 2004). Therefore, SRI in the 1990s involved negative screening to avoid investing in companies deemed unethical, for example, companies engaged in alcohol, tobacco, or gambling businesses (Sparkes and Cowton 2004). Since this approach did not encourage non-ethical companies to adopt responsible business practices (Heinkel, Kraus, and Zechner 2001), there was a significant push in the 2000s towards standardisation of ESG metrics through the Global Reporting Initiative (GRI) and the United Nations Principles for Responsible Investment (PRI) (Renneboog, Ter Horst, and Zhang 2008). This led to positive screening or "best-in-class" practices and the integration of ESG factors into mainstream investment strategies of asset managers and institutional investors (Statman and Glushkov 2009). Moreover, the global financial crisis further increased investor focus on sustainability, encouraging the development of sophisticated ESG scoring methodologies using extensive data sets and analytics by major scoring agencies such as MSCI, Sustainalytics, and Refinitiv.[9]

[9] See, ESG awareness is an enduring legacy of the global financial crisis. Accessed on July 1, 2024.

Today, ESG scoring agencies process voluminous amounts of data from various sources on firms' environmental, social, and governance performance and make sustainability measurable. Further, by benchmarking business practices on sustainability, ESG scores guide companies towards continuous improvement and help in mitigating negative incidents (Eccles, Ioannou, and Serafeim 2014). Thus, institutional investors, particularly signatories to the PRI initiative, sovereign wealth funds, and pension funds, use commercially available corporate ESG scores to incorporate sustainability into their investment strategy.[10] Further, government agencies are increasingly relying on these ESG scores to guide policymaking for promoting responsible corporate business conduct.[11] However, Larcker *et al.* (2022) argue that the notion of ESG scores measuring ESG performance is a myth due to issues with the construction of such metrics. Therefore, the reliance of SRI funds on ESG scores as a proxy for sustainability performance is being questioned (Nofsinger and Varma 2014). Despite these concerns, ESG scores remain influential in shaping global trade and investment, meriting further empirical investigation.

2.2. ESG Scores Through the Lens of Institutional Theory

The institutional theory by Powell and DiMaggio (1991) posits that organisations conform to societal norms to gain legitimacy and support. Therefore, this theory emphasises that external institutions are instrumental in shaping organisational behaviour and practices. Accordingly, in stakeholder-oriented economies, like France and Germany, socially responsible businesses are held in high regard, whereas in shareholder-oriented economies, like the U.S., economically responsible businesses are held in higher regard (Maignan 2001). This difference in stakeholder pressure influences managerial incentives to act in a socially responsible manner (Maignan and Ralston 2002). Therefore, the extent to which firms engage in socially responsible business practices is largely shaped by country-level institutions (Jackson and Apostolakou 2010) and by the institutionalised norms that govern corporate behaviour (Campbell 2007).

Subsequent studies have provided empirical evidence on the role of institutional differences in explaining the variations in the firms' ESG or CSR practices. In a cross-country study spanning 2,787 firms across 42 countries from 2002 to 2008, Ioannou and Serafeim (2012) report that political systems, labour and education systems, and national cultures drive corporate social performance (CSP). They observe that CSP is negatively affected by

---

[10] See, UNCTAD (2020). How public pension and sovereign wealth funds mainstream sustainability. Accessed on July 1, 2024.

[11] See, OECD (2020). Responsible business conduct and the OECD Guidelines for Multinational Enterprises. Accessed on November 16, 2024.

increasing national corruption and power distance, leftist ideology, laws promoting competition, and higher levels of shareholder protection, but is positively influenced by strong labour unions. Liang and Renneboog (2017), studying 23,000 companies in 114 countries, find that legal origin significantly explains CSP differences, with firms in civil law countries with stronger stakeholder protection scoring higher on ESG metrics than those in common law countries with stronger shareholder protection. Notably, Cai *et al.* (2016) demonstrate that country-level characteristics matter more than firm-level characteristics in explaining sustainability as measured by commercial ESG scores. Using 2,632 unique firms across 36 countries from 2006 to 2011, they find that country-level characteristics such as economic development, culture, and institutions explain 13.4% of the variation in ESG scores, while firm-level characteristics explain only 6.7% of the variation in ESG scores.

2.3. Why Are Corporate ESG Scores Lower in Developing Countries?

While institutional theory explains the importance of cross-country differences in shaping corporate ESG outcomes, it fails to explain whether the group dynamics of these institutional differences influence ESG scores. Empirically, examining institutional factors in isolation suffers from the problem of co-dependence among country attributes (Leuz *et al.* 2003). Practically, it is the interplay of multiple institutional factors that ultimately drives corporate behaviour, making it challenging to attribute variations in ESG scores to any single factor in isolation.

To understand the group dynamics of institutional differences in corporate ESG scores, we categorise countries as either developed or developing due to their practical relevance. First, institutions (e.g., the World Bank, IMF and UN) providing funding assistance to countries set the covenants in the assistance programmes according to the development status, as to whether a country is developed, developing, or in transition (Kentikelenis, Stubbs, and King 2016). Second, investors (i.e., foreign institutional investors and pension funds) managing global portfolios diversify their portfolios across geographical regions by following a similar classification of countries (Didier, Rigobon, and Schmukler 2013). Third, the different set of challenges facing developing countries leads to differences in policymaking between developed and developing countries.

To explain why ESG scores are expected to be different between developed and developing countries, we integrate the existing insights from institutional theory with the theory of human needs and the Environmental Kuznets Curve (EKC) hypothesis. By focusing on psychological and societal prioritisation, the theory of human needs suggests that in the context of developing countries, economic growth, poverty alleviation, and infrastructure development take precedence over environmental concerns because they directly impact people's survival and well-being. Accordingly, the first two goals of the UN SDGs, released in 2015—the eradication of poverty and hunger—are the prime focus in many developing countries. This, in turn, affects their ESG priorities. For example, in India, the emphasis on employment generation, gender diversity, and promoting inclusive development highlights the priority of addressing the fundamental socio-economic issues before shifting focus to the more advanced sustainability issues related to climate, environment, and biodiversity relevant to developed markets (SEBI 2022, 2023). Similar priorities drive policymaking in the countries of the African Union (Pinto 2024).

While the human needs theory provides a psychological foundation, the EKC hypothesis provides the appropriate microeconomic framework driving this behaviour. This hypothesis suggests that during the early stages of development, in the pursuit of economic growth, society's focus is on industrialisation, which increases pollution and environmental degradation. As the infrastructure expands to improve the standard of living, society reaches its peak level of environmental damage. Once the income levels have improved, society recognises the harmful effects of environmental degradation and begins to prioritise environmental sustainability over economic growth. Grossman and Krueger (1995) and Stern et al. (1996) empirically validate this hypothesis by showing that during the early stages of development, most countries prioritise investments in basic infrastructure like transportation and energy, with growing industrial output, which is accompanied by rampant pollution and rapid resource depletion. This reinforces the idea that countries focus on economic stability before considering environmental sustainability. The recent discussions at COP29 emphasise the ongoing dilemma for developing countries—economic growth or environmental sustainability—and stress the urgent need for financial support from developed countries to assist them in their climate adaptation efforts (Goar 2024). Identifying that the lack of fiscal strength prevents developing countries from balancing economic growth with climate adaptation. IMF has advocated that the poorest and most climate-vulnerable nations require substantial international financial assistance to achieve this goal (IMF 2023).

Integrating these perspectives with the institutional theory suggests that corporate ESG priorities and abilities also differ between developed and developing countries. Since the sustainability framework underpinning international ESG scores is based on institutional norms established in developed economies, firms in developing countries that are primarily focused on addressing local priorities and challenges may not prioritise such norms due to the limited financial support. As a result, these firms may fail to provide the quantity and quality of disclosures required by the ESG raters to score them comprehensively. Accordingly, the first hypothesis is stated as follows:

H1: *ESG Scores of firms in developing countries are lower than their developed counterparts.*

2.4. Systematic Rating Bias or Score Gap

In the next hypotheses, we explore whether the lower ESG scores of firms in developing countries also reflect a systematic "*rating bias*" or simply a "*score gap*" arising from features of the ESG scoring process itself.

2.4.1. Why is Score Gap a possibility?

The ESG rating measurement and methodology are not without their flaws (Larcker *et al.,* 2022). One significant measurement issue is the size bias, which results in disproportionately higher ESG scores for larger firms irrespective of industry (Drempetic *et al.* 2020). This issue arises from the ability of larger firms to generate larger volumes of ESG data than smaller firms due to their financial strength. Drempetic *et al.* (2020) term this as the data availability bias. Closely related is the disclosure bias, Raghunandan and Rajgopal (2022) explain this as the propensity of firms with extensive ESG disclosures to receive higher ESG scores, regardless of the real impact of such initiatives.

Another major concern is the lack of transparency in ESG methodologies. While some agencies disclose their scoring frameworks, crucial details necessary for meaningful interpretation and accurate comparison remain undisclosed. Additionally, the absence of a standardised ESG framework leads to differences in data collection methods, inconsistencies in data formats, and differences in quality control—all of which make comparability across firms and countries difficult (Berg *et al.* 2022; Christensen *et al.* 2022; Semenova and Hassel 2015).

The issues highlighted above affect both developed and developing country firms. However, a particular feature of the ESG rating methodology can translate these concerns into systematically lower ESG scores for developing country firms than developed country firms. Liang and Renneboog (2017) reflect that ESG rating agencies assess companies relative to their industry peers across international markets, making ESG scores less reflective of the domestic institutional environment. For instance, Refinitiv benchmarks environmental and social scores against global industry averages, but governance scores are assessed against national averages (LSEG 2024). This global benchmarking approach results in systematic variations in ESG scores across countries (Ioannou and Serafeim 2012; Cai *et al.* 2016; Liang and Renneboog 2017). Such a global benchmarking approach disproportionately penalises firms in developing countries that may lack the financial resources and regulatory requirements to produce extensive ESG disclosures, even if they engage in meaningful sustainability efforts.

To examine whether firms in developing countries are systematically disadvantaged by Refinitiv's global benchmarking approach, we need an alternative, diagnostic version of ESG scores that benchmarks firms jointly by country and industry, thereby isolating within-country, within-industry variation. We call these Modified ESG scores (*MESG*). A statistically insignificant difference between the *MESG* scores of developed and developing country firms would suggest that the lower ESG scores of developing country firms stem from Refinitiv's implicit global benchmarking assumptions. We test this assertion using hypothesis H2a, stated below:

H2a: *MESG Scores of firms in developing countries are not different from their developed counterparts.*

One should note that the "*score gap*" is not a "*rating bias*" because it does not arise from differential treatment or intentional (or unintentional) discrimination by the ESG rater against firms in developing countries. Instead, it results from structural features of the scoring methodology itself. Specifically, Refinitiv applies implicit global, industry-based benchmarking assumptions that compare firms across countries without adequately accounting for differences in domestic priorities, institutional environments, and developmental contexts. Under such a framework, firms in developing countries may systematically score lower even when they perform well relative to their local peers.

2.4.2. Why is Rating Bias a possibility?

While lower ESG scores of firms in developing countries may partly reflect developmental and domestic priorities, concerns raised by developing countries that ESG raters may be systematically assigning lower scores to their firms represent a serious concern that merits careful investigation (Matthews 2022).

The susceptibility of ratings to biases adversely affecting developing country firms is not new. Sovereign credit ratings have been criticised for reflecting home country optimism rather than economic fundamentals (De Moor *et al.* 2018). For instance, after the 2008 financial crisis, sovereign credit ratings were criticised for being overly optimistic about regions with stronger economic, geopolitical, and cultural ties to the West (Fuchs and Gehring 2017). Developing nations, despite their significant economic growth, often received unfairly low procyclical credit ratings disconnected from their economic fundamentals (Cornaggia, Cornaggia and Israelsen 2002). This introduced a rating ceiling wherein corporate credit ratings never surpassed the sovereign credit rating, thereby introducing a bias (Borensztein *et al.* 2013; Almeida *et al.* 2017).[12,13] Recent evidence shows that such sovereign-driven rating constraints can also reduce firms' ESG engagement (Boumparis, Florackis, Guedhami, and Sainani 2025).

A similar problem exists with commercial corporate governance ratings, where governance norms in developed countries are applied as a universal benchmark. As a result, such ratings fail to recognise and appreciate alternative governance mechanisms existing in developing countries. Consequently, firms in developing economies often receive lower ratings, which may not truly reflect their governance quality (Doidge *et al.* 2007; Black *et al.* 2017).

We contend that ESG scores are also vulnerable to this problem. Rating agencies often use a sustainability framework designed for developed countries, prioritising environmental issues while underweighting concerns more important in developing countries, such as socio-economic upliftment of the society through employment generation in smaller towns, gender diversity among employees, and inclusive development (Grossman and Krueger 1995; Stern *et al.* 1996). This imbalance is also evident in how funds implement ESG, with environmental factors often dominating at the expense of social and governance issues (Agoraki, Kouretas, and Zhao 2025). Regulatory authorities in developing countries have acknowledged these issues. For example, the consultation paper released by the Securities Exchange Board of India (SEBI) in February 2023 highlights that existing ESG rating agencies overlook the domestic context in their corporate sustainability assessments. It argues that ESG issues plaguing emerging markets differ fundamentally from developed countries (SEBI 2023). Similar concerns have been raised by the economies of the African Union (Pinto 2024). As scarcity of resources necessitates developing economies to align their sustainability goals with their economic needs, implementing stringent climate policies and pollution norms at the cost of economic growth remains a challenge (Cai *et al.* 2016).[14]

This institutional bias is further amplified by geographic distance. Most commercial ESG scoring agencies are headquartered in Europe and the U.S., making developing country firms geographically distant from their raters (Cai *et al.* 2016). As geographic distance increases information asymmetry and reduces institutional familiarity (Ayers,

---

[12] See, India's ratings don't reflect economy's fundamentals: CEA. Accessed on July 1, 2024.
[13] See, Moody's politically biased credit outlook cut won't affect China's long-term upward growth trend: experts. Accessed on July 1, 2024.
[14] See, UNCTAD (2021). World Investment Report 2021: Investing in Sustainable Recovery. Accessed on July 1, 2024.

Ramalingegowda, and Eric Yeung 2011; Kim, Miller, Wan, and Wang 2016), developing country firms are prone to receiving lower ESG scores than their developed counterparts (Cai *et al.* 2016). Consequently, we test whether ESG raters systematically assign lower scores to developing country firms using hypothesis H2b, stated below:

H2b: *MESG Scores of firms in developing countries are lower than their developed counterparts.*

Since *MESG* benchmarks firms jointly by country and industry, thereby isolating within-country, within-industry variation—unlike Refinitiv's implicit global benchmarking approach—lower *MESG* scores for firms in developing countries would indicate the presence of intentional and/or unintentional bias.

3. Data, Covariates and Descriptive Analysis

3.1. Data and Sample Selection

We source ESG scores from the London Stock Exchange Group (LSEG) Workspace database, formerly ASSET4 by Thomson Reuters or Refinitiv Eikon, which offers global coverage starting from 2002. The Refinitiv ESG score reflects a firm's commitment to environmental, social, and governance dimensions, constructed using 186 unique metrics. There are three subcategories within the environmental pillar: innovation, emissions, and resource use, containing 20, 28, and 20 indicators, respectively. Workforce, human rights, community, and product responsibility are the four subcategories of the social pillar, each of which has 30, 8, 14, and 10 indicators, respectively. The three subcategories comprising the governance pillar—management, shareholders, and CSR strategy—are measured using 35, 12, and 9 indicators, respectively. Along with the ESG scores and the pillar scores, we also extract the data for 186 metrics that form ESG scores, and the other financial data used in our analysis from the LSEG Workspace. Based on the country where the firms are incorporated, firms are classified as 'developed' or 'developing' following the IMF (2023) classification.

We begin with a global sample of 13,323 firms with annual data from 2002 to 2023, comprising 102,079 firm-year observations. After removing private firms and firms belonging to academic and educational services, financial services, government activity, real estate and utilities sector, we arrive at 8,996 unique firms representing 72,508 firm-year observations. Next, we remove firms belonging to countries with fewer than 15 observations per year and firms belonging to countries that are not classified as developed or developing by the IMF. Finally, we remove firm-year observations with missing data for the computation of control variables in our study. This leaves us with 54,534 firm-year observations of 7,970 firms across

46 countries. After this sample selection process, we note that observations from developing countries before 2009 are dropped. Hence, to maintain comparability, we consider firm-year observations from 2009 for our analysis. The final sample covers 7,969 firms across 46 countries, comprising 50,327 firm-year observations. The sample selection process is depicted in Table A1 in the Appendix. Table A2 depicts the sample distribution across developed and developing countries. It reveals that approximately 21% of the final sample comprises observations from developing countries, while 79% comprises observations from developed countries. To mitigate the impact of outliers, we winsorized all continuous variables (except the *ESG* and *MESG*) at their 1st and 99th percentiles. Further, Table A3 of the Appendix, which depicts the country-wise distribution of the sample, indicates that more than 50% of the representation in the developing countries is from India and China, while the United States alone accounts for about 40% of the representation amongst developed countries.

3.2. Why Refinitiv ESG Scores?

We use Refinitiv ESG scores for our analysis for four main reasons: coverage, depth, usage, and methodological transparency.

Regarding coverage, Refinitiv provides substantially better coverage of firms from developing countries than others, such as MSCI and Sustainalytics. The more limited coverage of developing country firms by these providers constitutes a major impediment to meaningful cross-country analysis (Cai *et al.* 2016). As of 2024, Refinitiv (Asset4/LSEG/Thomson Reuters) offers the broadest global coverage, with data on more than 16,500 firms (LSEG 2024). Sustainalytics (Morningstar) covers over 16,000 firms, while MSCI (KLD) covers more than 10,000 firms.[15] Although all three agencies employ global industry benchmarking, their underlying methodologies differ substantially, leading to significant divergence in ESG scores (Berg *et al.* 2022).

In terms of depth, Berg *et al.* (2022) compare ESG databases across 65 material ESG issues and document the number of indicators used under each ESG category. Refinitiv provides the most extensive coverage of these issues, with 282 indicators, compared to 163 indicators for Sustainalytics and 68 for MSCI. As of 2024, Refinitiv reports more than 800 ESG indicators in total, of which 186 are used to construct its composite ESG score and pillar-level scores (LSEG 2024). By contrast, MSCI covers 35 ESG issues (MSCI 2024), and

[15] MSCI 2024 *ESG Ratings*. Available at: https://www.msci.com/data-and-analytics/sustainability-solutions/esg-ratings-climate-search-tool. Sustainalytics 2024 *ESG Risk Ratings*. Available at: https://www.sustainalytics.com/esg-data (Accessed: 31 December 2025).

Sustainalytics reports more than 250 indicators (Sustainalytics 2024).[16] Importantly, unlike Refinitiv, neither MSCI nor Sustainalytics discloses the specific indicators used to construct their aggregate ESG ratings, which is essential for estimating the Modified ESG (*MESG*) scores.

With respect to usage, having been cited in more than 1,500 academic articles (Berg, Fabisik, and Sautner 2021), we assessed adoption in academic research by conducting a Scopus search of article titles, abstracts, and keywords for publications published or in press as of 31 December 2025.[17] The search yielded 463 academic articles (66%) referring to Refinitiv ESG data, compared to 160 articles (23%) referring to MSCI and 77 articles (11%) referring to Sustainalytics. In terms of practical relevance, Refinitiv data are integrated into widely used financial data and trading terminals, providing access to a broad user base of practitioners and investors.

Finally, regarding methodological transparency, Refinitiv offers comparatively detailed documentation of its ESG scoring framework, including the mapping of individual indicators to category-level, pillar-level, and aggregate ESG scores. This level of disclosure enhances the accessibility and replicability of Refinitiv's ESG measures for researchers and practitioners, in contrast to MSCI and Sustainalytics, which do not provide indicator-level transparency.

Considering Refinitiv's superior coverage, depth, widespread usage, and methodological transparency, we rely on Refinitiv ESG scores as the primary ESG measure in this study.

3.3. Covariates

3.3.1. Refinitiv ESG Scores and Modified ESG Scores

The dependent variables of interest in our study are the ESG scores (*ESG*) and the Modified ESG scores (*MESG*). *ESG* represents the ESG scores from Refinitiv for the company *i* in year *t*. The *ESG* scores are composed of 10 category scores with respect to environmental (E), social (S) and governance (G) pillars, capturing over 870 firm-level ESG measures (indicators hereafter), of which 186 indicators are considered for calculating the score. The list of indicators, their polarity and industry relevance are disclosed by the LSEG upon request,

[16] https://www.sustainalytics.com/material-esg-issues-resource-center Accessed on 31 December 2025.
[17] Refinitiv Query: ("Refinitiv" OR "LSEG" OR "Asset4" OR "Thomson*" OR "Eikon") AND ("ESG"); MSCI Query: ("MSCI" OR "KLD") AND ("ESG"); Sustainalytics Query: ("Sustainalytics" OR "Morningstar") AND ("ESG").

subject to a non-disclosure agreement.[18] Table 1 shows the categories under each ESG pillar and the number of corresponding indicators for each category. To empirically test whether the ESG scores may be leading to a *"score gap"* or a *"rating bias",* we first briefly explain the scoring procedure. *ESG* score calculation is a three-step process: a) category score calculations; b) category weight calculations; and c) ESG score calculations.[19] The three steps are applied separately for each year (LSEG 2024).

<Insert Table 1 Here>

Step 1: Category Scores

To calculate the category scores, first, a percentile scoring methodology is applied to each indicator using the formula in Eq. (1). Then, the percentile scores of the individual indicators are summed at the category level for each company-year. Finally, the percentile score formula is then reapplied to this aggregate to produce the category score.

$$Score = \frac{\begin{pmatrix} Number\ \ of\ companies\ with \\ a\ worse\ value \end{pmatrix} + \begin{pmatrix} Number\ of\ companies\ with\ the\ same\ value \\ included\ in\ the\ current\ one \end{pmatrix} / 2}{Number\ \ of\ companies\ with\ a\ value} \quad (1)$$

To calculate the percentile scores for environmental and social indicators and categories, LSEG uses the TRBC (The Refinitiv Business Classification) industry group as a benchmark; for governance indicators and categories, the country of incorporation is used as a benchmark.

Step 2: Category Weights

The category weights are calculated using a data-driven approach based on the relative importance of the material indicators within each category.[20] For environmental and social pillars, category weights are derived using a combination of two methods: a) the quant industry median method, and b) the transparency weights method. Under the quant industry median method, the median of each material indicator is calculated for each industry. These median values are then ranked into deciles across industries. Under the transparency weight method, the disclosure percentage is calculated for each industry, and then the resulting values are ranked into deciles across industries. These decile ranks reflect the magnitude for each

[18] Polarity indicates whether a higher value is positive or negative. For example, having environmental reduction policy is positive (LSEG 2024).

[19] The data for the indicator are either in Boolean ("Yes", "No" or "Null") or numeric format. The Boolean indicators are converted to numeric based on their polarity. For Boolean indicators with positive polarity, a data point containing "Yes" are assigned a value of 1, "No" or "Null" is assigned a value of 0. Conversely, for a Boolean indicator with negative polarity, a data point containing "No" is assigned a value of 1, whereas "Yes" or "Null" is assigned a value of 0.

[20] See, LSEG (2024), page 12, for the list of adjacent themes and the material indicators for each category.

category, which varies across industries. To obtain the category magnitude, the average of ranks for material indicators under each category is calculated.

Refinitiv considers all indicators equally important for the calculation of magnitude weights of governance categories for all TRBC industry groups. The default category weights are assigned at five points, with the points' distribution ranging from 1 to 10. Since governance comprises three categories, the total points under governance would be 15. For the governance pillar, the magnitude for each category is calculated by dividing the count of indicators in each governance category by all indicators in the governance pillar and multiplying the resultant value by 15. The magnitude remains the same across all industry groups across all countries for categories under the governance pillar. The final category weights are obtained by dividing the magnitude of a category by the sum of the magnitudes of all categories.

Step 3: ESG Score

The aggregate *ESG* score for each firm-year is calculated as the weighted average of all category scores, where category scores are multiplied by corresponding category weights and then added. For the pillar scores (i.e., environmental, social and governance), the category weights within each pillar are first normalised, so that the normalised total weight is equal to one. The pillar scores are then calculated as the weighted average of the category scores within that pillar, using these normalised weights.

3.3.2. Modified ESG scores (MESG)

As highlighted in steps 1 and 2 of the *ESG* score calculation process, the benchmark for environmental and social pillar scores is the TRBC industry groups. Hence, this method fails to capture the country-specific institutional differences and national priorities in the ESG scores (Basu *et al.* 2022). We modify the Refinitiv *E, S, G* and *ESG* scores by making two key improvements. First, when calculating the percentile scores for the indicators and categories, we use both country and industry as benchmarks. And second, when deriving category magnitudes through the quant industry median and transparency weights methods, we again apply the country and industry benchmark for environment and social categories. Consistent with the LSEG methodology, the country and industry classifications are based on TRBC industry groups and the country of incorporation. By doing so, the *MESG* scores address the methodological shortcomings in the scoring process.

3.3.3. Classification of Developed and Developing Countries

Following the IMF (2023) classification, we classify firms as 'developed' or 'developing' based on their country of incorporation. While similar classifications are provided by other agencies, including the United Nations and World Bank, we rely on the IMF classification for our firm-level analysis because it is based on capital market integration, financial openness, and investor perceptions (Carrieri, Errunza, and Hogan 2007). In contrast, the World Bank's income-based categorisation emphasises developmental stages (World Bank 2022), while the United Nations classification focuses on socio-political dimensions (United Nations 2021). As a result, these classifications may not be suitable for financial analysis. By adopting the IMF classification, we ensure that firm-level comparisons are grounded in categories that capture differences most relevant for financial analysis. Accordingly, our independent variable of interest is *DVPG*, an indicator variable that takes the value of 1 if a firm is incorporated in a developing country as per the IMF 2023 classification, else zero. See Table 2 for variable description. For a discussion on whether the *DVPG* dummy is a robust proxy for institutional differences across countries, please refer to Section 4.1.

3.3.4. Control Variables

We include several firm-specific control variables that cause variations in corporate ESG scores (Ioannou and Serafeim 2012; Cai *et al.* 2016; Liang and Renneboog 2017; Drempetic *et al.* 2020). First, we control for size (*SIZE*), profitability (*ROA*), and working capital (*LIQ*) because larger companies that are more profitable and have better liquidity exhibit better sustainability performance. Next, since firms with a greater degree of product differentiation exhibit better ESG performance, we use R&D intensity (*RDEXP*) as an indicator of product differentiation. Further, we include leverage (*LEV*) to control for the effect of credit constraints. We include capital expenditure (*CAPEX*) as an indicator of resource allocation since higher *CAPEX* would strain the company's resources, thereby limiting the funds available for ESG initiatives. Next, we control for *GROWTH* (change in sales scaled by lagged sales), since an aggressive growth strategy could be an indicator of compromise on ESG initiatives. Finally, we control for financial health through a dividend payment dummy (*DIV*).

A detailed description of all the variables in the study is presented in Table 2. Our univariate regression analysis, presented in Table A4 in the Appendix, confirms that these control variables are significant predictors of both *ESG* and *MESG* and their respective individual components.

<Insert Table 2 Here>

3.4. Descriptive Analysis

Table 3 presents the mean differences in corporate ESG scores between developed and developing countries. The *ESG* for firms in developing countries (mean of 0.452) are significantly lower than that in developed countries (mean of 0.475). This provides an indication that the corporate ESG scores could be lower for developing country firms due to differing priorities. Though the *MESG* also differs between developed (0.524) and developing countries (0.516) by a much smaller margin, it appears to be statistically different between the two groups. Further, a univariate comparison of the means may not account for the influence of other firm-specific factors in explaining the difference. Hence, further investigation is required to conclude whether the lower ESG scores are due to poor ESG performance, measurement issues or due to ratings biases.

<Insert Table 3 Here>

Further, Table A5 in the Appendix presents descriptive statistics of the full list of variables, highlighting the differences between developed and developing countries. Firms in developing countries exhibit higher allocation towards capital investments (*CAPEX*), better profitability (*ROA*) and pay higher dividends (*DIV*) than firms in developed countries. However, these firms are smaller in size (*SIZE*), have lower leverage levels (*LEV*), exhibit lesser product differentiation (*RDEXP*) and lower liquidity (*LIQ*) than those in developed countries. The growth prospects (*GROWTH*) are not much different between developing and developed country firms. The correlation matrix in Table A6 of the Appendix confirms no severe multicollinearity issues.

4. Empirical Results and Discussion

4.1. Does *DVPG* Capture Differences Between Developed and Developing Countries?

Our key independent variable is the *DVPG* dummy, which takes the value of 1 if a firm is incorporated in a developing country and 0 otherwise. A key issue raised by the international literature is that institutional differences should be accounted for in cross-country studies (Ball, Robin, and Wu 2003; Leuz *et al.* 2003; Bhattacharya, Daouk, and Welker 2003; Hail and Leuz 2006; Isidro, Nanda, and Wysocki 2020). In hypothesis H1, we argue that these differences, along with national priorities, drive variations in ESG scores between firms in developed and developing countries. Hence, before testing our hypotheses, we test whether the *DVPG* dummy adequately captures these institutional differences.

For this purpose, we first compile data on 72 country-level attributes from Isidro *et al.* (2020), encompassing both static (time-invariant) and dynamic (time-variant) measures across economic, regulatory, social, and political dimensions. Following Isidro *et al.* (2020), we first take the average of the time-varying variables from 2009 and 2023, excluding variables with insufficient observations (less than 30), leaving us with 64 variables (listed in Appendix Table A8). To avoid adjustment bias, we replace missing values for the retained variables with their sample mean across developing and developed countries. For comparability, all variables are standardised by their mean and standard deviation across the sample of 46 countries.

Given the high correlation, Principal Component Analysis (PCA) results presented in Table A7 of the Appendix show that 14 latent factors sufficiently encompass the information across 64 country-level attributes. Collectively, these fourteen mutually orthogonal factors explain 84.95% of the total variance in the data. Factor 1 alone explains 35.67% of the variance (eigenvalue = 22.830), followed by Factors 2 to 4 explaining 13.36% (8.551), 4.98% (3.189), 4.69% (3.006), respectively. Table A8 of the Appendix shows that the item-wise communalities (extractions) range between 0.743 and 0.970, exceeding the 0.50 threshold.[21] Variables are grouped by their highest factor loadings, with weights determined by their corresponding factor loadings and eigenvalues.[22] Factor 1 captures most of the variation in 30 variables, largely reflecting economic, regulatory and governance-specific attributes.

The predicted factors and their scores are used in univariate regressions of the *DVPG* dummy (Table 4). Regressing the *DVPG* dummy on these 14 factors shows that Factor 1 has highest explanatory power, with a significant (1%) negative coefficient across all three classifications—IMF, World Bank (WB), and United Nations (UN), with the best fit for IMF classification (Pseudo R^2 = 88.6%), followed by UN (62.9%) and WB (57.0%) classifications. The first factor provides a highly parsimonious representation, capturing most of the information embedded in thirty of the sixty-four variables, including measures of institutional strength and governance, political stability, economic and market development, and social capital.[23] This indicates that the countries with stronger economic conditions, better regulatory and governance mechanisms, political stability and social capital have a lower probability of

[21] A value closer to 1 suggests that extracted factors explain more of the variance of an individual item (Hair, Black, Babin and Anderson 2018).

[22] PCA predicts factor scores as a linear combination of variables where factor loadings and eigen value of a factor determine the weight assigned to each variable.

[23] Table A8 in the Appendix provides details on the variables that make up factor 1. The institutional strength and strong governance are reflected by the measures of the rule of law, control of corruption, regulatory quality, law and order, judicial independence, judicial efficiency, property rights, creditor rights, enforcement of audit/accounting standards, and the assessment of tax evasion. The political stability is reflected by low political risk, low repudiation of contracts by govt., political stability, and political score. GDP per capita, Economic freedom, Information and knowledge, Bank money in private sector to GDP, IPOs to GDP, Big4 market share are the measures of economic and market development, and media freedom, trust, individualism, power distance (negative), Secrecy (negative), ownership concentration (negative), Muslim (negative), latitude show social capital and cultural traits.

being classified as developing by all three classifications. The remaining 13 factors fail to exhibit statistically significant explanatory power. Together, these findings suggest that the incremental value of additional factors is limited and that the IMF's developing–developed classification most effectively captures cross-country institutional differences, consistent with the dominant explanatory role of Factor 1 in the variance structure.

<Insert Table 4 here>

4.2. Empirical Model

We consider the following baseline model to test our hypotheses:

$$ESG_{i,t} = \alpha_0 + \beta_1 DVPG_{i,t} + Controls_{i,t-1} + IndFE + YearFE + \varepsilon_{it} \quad (2)$$

In Eq. (2), *DVPG* represents the developing country dummy. *Controls* represent the control variables as described in Section 3.3.4 and are lagged by one year. *IndFE* and *YearFE* represent industry and year-fixed effects, respectively. ε_{it} represents the error term. To reduce heteroskedasticity in the error term, we clustered the standard errors at the firm level.

Since firms in developing countries differ from firms in developed countries on several firm-specific attributes, we consider an entropy-balanced sample of firms from developed and developing countries to address sample selection issues and eliminate self-selection bias. This further ensures that varying characteristics between firms in developed and developing countries do not influence our inferences. Following the procedure as in Hainmueller (2012) and Shroff *et al.* (2017), we match firms on all the firm-specific attributes considered in our study (i.e., *Controls*) and arrive at an entropy-balanced sample. The results of the entropy balancing procedure across three moments (mean, variance, and skewness) are presented in Table A9 of the Appendix. Panel regression analysis is carried out with this entropy balanced sample.

Consistent with prior research (Cai *et al.* 2016; Liang and Renneboog 2017), we do not control for firm-fixed effects since it is perfectly collinear with the *DVPG* dummy. We have not controlled for country-fixed effects in our regression models due to two important reasons: First, from a conceptual standpoint, as explained in Section 4.1, the *DVPG* dummy captures about 90% of the institutional differences across countries using factors identified by Isidro *et al.* (2020). This makes the *DVPG* dummy a comprehensive metric that consolidates the variations across institutional differences. This further justifies excluding institutional controls, as their inclusion would amount to introducing the same information twice and would lead to multicollinearity. Second, from an empirical standpoint, the *DVPG* dummy is highly correlated with the country-fixed effects, and multicollinearity would eliminate the *DVPG* dummy from the regression models. We acknowledge that some may still raise concerns about the reliability of our results due to omitted variable bias, as we have not directly controlled for various factors that differentiate developed and developing countries. However, Section 4.3.1.1 presents and discusses the results of Oster's Coefficient Stability analysis (Oster, 2019), confirming that our statistical inferences are robust to omitted variable bias. *ESG* is the dependent variable and denotes *ESG* and its components—environmental (*ENV*), social (*SOC*) and governance (*GOV*) when testing H1. As the firm-level differences impacting ESG scores (e.g., performance and size) are taken care of by the entropy balancing procedure, if β_1 is negative in the *ESG* regressions, it confirms our first hypothesis that corporate ESG scores are lower in developing countries than in developed countries due to institutional differences and national priorities. To test H2a and H2b, we replace *ESG* in Eq. (1) with the modified ESG score (*MESG*) and the modified component scores (*MENV*, *MSOC*, and *MGOV*). If β_1 still remains negative in the *MESG* regressions, it means that the lower corporate ESG scores for developing country firms are also due to "*rating bias*", thereby confirming H2b. However, if β_1 becomes insignificant, it would mean that the lower scores are primarily due to features of the ESG scoring process itself—implicit global benchmarking assumptions embedded in Refinitiv's ESG rating methodology—rather than from any systematic "*rating bias*". Thereby confirming the existence of a "*score gap*" in the ESG scores in line with our hypothesis, H2a.

4.3. Test of H1

Applying the institutional theory and the theory of human needs, H1 predicts that the *ESG* are lower in developing countries compared to developed countries. In line with this prediction, our panel regression results show that the *DVPG* dummy is negative and significant in the *ESG* regression, suggesting that corporate ESG scores are lower in developing countries (see Column (2) of Table 5). In terms of economic significance, *ESG* scores in developing countries are lower by 16.34% than in developed countries, when measured with respect to the standard deviation (see Mitton 2024).[24]

<Insert Table 5 Here>

Upon examining the component scores, the results in Table 6 indicate that the *DVPG* dummy is negative and significant only in the *ENV* and social *SOC* regressions but not in the

[24] The economic significance is calculated following Mitton (2024). With respect to standard deviation: for *ESG*, −0.034/0.208 = −0.1634; for *MESG*, −0.002/0.175 = −0.0114. The standard deviation of the ESG scores for the entire sample are provided in Table A5 in the Appendix.

GOV regressions (Columns (2) to (4)). Unlike *ENV* and *SOC* score calculations, Refinitiv employs country benchmarking when calculating *GOV* scores, arguing that the best governance practices are more consistent within countries (LSEG 2024). This implies that institutional factors dominate the industry norms, which is inconsistent with their assumptions for *ENV* and *SOC* scores. Consequently, we find an insignificant difference in *GOV* scores between the firms in developing and developed countries. However, studies argue that corporate governance practices vary across both industries and regulatory environments—see Adams, Hermalin and Weisbach (2010) survey. We find support for this line of argument; our results show that *MGOV* scores are significantly higher for firms in developed countries when benchmarked with respect to country and industry. Overall, empirical tests provide strong support for hypothesis H1, suggesting that environmental and social aspects of sustainability take a back seat in developing countries, as the focus here is on achieving economic stability. However, this finding does not rule out the possibility that either genuinely lower performance, measurement issues or bias in ESG scores may explain the lower ESG scores for developing country firms.

<Insert Table 6 Here>

4.3.1 Endogeneity Checks

Endogeneity in regression-based empirical research may arise due to four issues: self-selection, reverse causality, omitted variables, and measurement error (Roberts and Whited 2013; Hill, Johnson, Greco, O'Boyle, and Walter 2021). We have addressed the self-selection issue by employing an entropy-balanced sample (Hainmueller, 2012) while estimating our regression models. Reverse causality is not an issue in our context, as ESG scores do not influence the *DVPG* variable. However, measurement error and omitted variable bias remain a concern even after the inclusion of numerous firm-specific controls. Therefore, we perform tests addressing endogeneity concerns arising from omitted variables and measurement errors.

4.3.1.1 Endogeneity Due to Omitted Variables

To address the potential issue of unobserved confounding variables on our inferences, we employ Oster's Coefficient Stability Test (Oster 2019), which is a bounding technique wherein the robustness of the OLS inferences to omitted variable bias is assessed based on a proportional selection relationship arising from the correlation between observable control variables and unobservable controls. For this purpose, Oster (2019) defines two parameters: the ratio of the degree of selection on unobservables to observables (δ) and R^2 from a hypothetical regression of the outcome on the treatment and both observable and unobservable controls (R_{max}). If the coefficient of interest remains relatively stable (i.e., if the magnitude of

δ is substantially greater than one) when the R^2 from the baseline regression is increased to R_{max}, then we conclude that the omitted variable bias is not strong enough to invalidate our findings.

<Insert Table 7 Here>

The results of this analysis are presented in Table 7. Rows (1) and (2) present the coefficient of the *DVPG* variable and R^2 from the multivariate regression results in Table 5.[25] Row (3) presents the assumed value of δ, which is set to one, following Oster (2019). Row (4) specifies the R_{max} value for the regression model. Following Oster (2019), we set R_{max} (i.e., the upper bound on R^2) to 1.3 times the R^2 from our baseline multivariate regression results. Next, we compute the bounds of treatment effects $[\beta_{\text{baseline}}, \beta \times (\min\{1.3 \times R^2_{\text{baseline}}, 1\},1)]$ and check whether the interval excludes zero. Rows (5) and (6) highlight that there is little change in the *DVPG* coefficients, and the values exclude zero. This suggests that the inclusion of unobserved control variables will not lead to a different conclusion from our baseline results. Next, we present the estimated δ, that would be required to change the *DVPG* coefficient equal to zero when R^2 is increased to R_{max}. Rows (7) and (8) highlight that the magnitude of δ is substantially greater than one. Comparing these values with the threshold of one based on the recommendations of Oster (2019), we conclude that the degree of selection on unobservables must be 2.079; it is highly unlikely that unobservable covariates would be 2.079 times more important than the observable covariates. These findings confirm that our results are robust to omitted variable bias concerns.

4.3.1.2 Endogeneity Due to Measurement Error

Measurement error refers to the possibility that the dependent and/or independent variable in our study could be capturing an unobserved phenomenon rather than the observed phenomenon, thereby invalidating our findings (Hill *et al.* 2021). This concern is also addressed by the DiD analysis performed using mandatory ESG disclosures as an exogenous shock (see Section 4.4.1 for details), as mandatory ESG disclosures have a first-order effect on ESG disclosures through the information channel (Christensen *et al.* 2022). Overall, our results are robust to endogeneity concerns due to measurement error.

4.4. Test of H2a and H2b

[25] We use Stata code regress instead of reghdfe to use psacalc. The coefficients are slightly different from Table 5, but results are qualitatively unchanged.

In H2a, we argue that the lower ESG scores of developing country firms may also reflect measurement errors among ESG raters, while H2b suggests that they may also be due to rating bias. Regressing *MESG* scores on *DVPG* and a set of firm-level controls reveals that the coefficient of *DVPG* becomes insignificant (see Column (3) of Table 5). Regarding the component scores, the *DVPG* dummy becomes insignificant in regression estimates employing *MENV* and *MSOC* as dependent variables (see Columns (5) and (6) of Table 6). Further, the economic significance of the *DVPG* dummy in the *MESG* regressions is only 1.14%. Since the difference disappears once the global benchmarking issue is accounted for, this confirms that the "*score gap*" arises from features of the ESG scoring process itself rather than from any systematic rating bias. This methodological "*score gap*" is driven mainly by implicit global benchmarking assumptions embedded in Refinitiv's ESG rating methodology. Refinitiv's global industry-based benchmarking framework does not adequately account for the domestic priorities of emerging-market firms, thereby systematically disadvantaging them. While this supports H2a, the insignificant coefficients also confirm the absence of any systematic "*rating bias*", thereby refuting H2b.

Overall, the results collectively indicate that the observed disparity in ESG scores between firms in developed and developing countries reflects a methodological "*score gap*" embedded in the scoring process, rather than a systematic "*rating bias*" on the part of the rater.

4.4.1 Is Poor Disclosure the Source of Lower ESG Scores?

A second reason why the ESG scores of developing-country firms may be lower is their relatively lower levels of disclosure, which ESG raters often interpret as poor performance. This occurs because raters typically evaluate firms based on the volume of ESG disclosures rather than on actual sustainability outcomes (Raghunandan and Rajgopal 2022). To test this possibility, we construct a staggered Difference-in-Differences (DiD) model around the implementation of mandatory ESG disclosure requirements in different countries (Krueger, Sautner, Tang, and Zhong 2024). Since mandates were introduced at different points in time across countries, we follow Krueger *et al.* (2024) in adopting a staggered design and modify our baseline model in Eq. (2) as follows:

$$\begin{aligned} ESG_{i,t} = {} & \alpha_0 + \beta_1 DVPG_{i,t} + \beta_2\, MAN + \beta_3 DVPG_{i,t} \times MAN + Controls_{i,t-1} \\ & + IndFE + YearFE + \varepsilon_{it} \end{aligned} \tag{3}$$

Here, *MAN* takes the value 1 from the year a country introduces mandates and zero otherwise. In the *ESG* regressions, if β_3 is positive and significant in Eq. (3), it would indicate that lower ESG scores in developing countries were primarily due to weaker disclosures, and that scores improved once disclosure requirements were introduced—consistent with

Raghunandan and Rajgopal (2022), who argue that ESG ratings reflect disclosure volume rather than performance. A negative and significant β_3 would suggest that the additional information disclosed as per the mandatory disclosure frameworks in developing countries—shaped by domestic priorities—does not fully align with the criteria embedded in international ESG ratings, which are grounded in institutional norms of developed economies. An insignificant β_3 would imply that mandatory disclosure requirements did not affect the ESG scores of developing country firms.

In the *MESG* regressions, a significant β_3 would invalidate our claims of a methodology-driven "*score gap*" and would rather imply the existence of a disclosure bias as identified by Drempetic *et al*. (2020) and Raghunandan and Rajgopal (2022). As the *MESG* scores address the global benchmarking issue in the construction of the ESG scores, if β_3 is insignificant in the *MESG* regressions, this would indicate that lower scores are not attributable to disclosure bias but instead reaffirm that they result from methodology-driven "*score gap*" embedded in Refinitiv's global benchmarking process.

<Insert Table 8 Here>

The regression results testing these assertions are presented in Table 8. In the *ESG* regressions, the coefficient on *DVPG* × *MAN* remains negative and significant (column 2). This suggests that the disclosures in developing countries are reflective of the domestic priorities rather than in conformity with the rating framework adopted by the ESG rating agencies. As a result, the additional disclosures were considered unfavourable, leading to lower scores. In the *MESG* regressions, the coefficient on *DVPG* × *MAN* is insignificant. Since *MESG* scores adjust for measurement issues in ESG construction, the evidence suggests that lower scores are not driven by disclosure bias. Instead, they reinforce the interpretation that the lower scores reflect methodology-driven "*score gap*" embedded in Refinitiv's global benchmarking process.

Next, to validate our findings, we assess the parallel trends assumption underlying our staggered DiD estimation following Fauver, Hung, Taboada, and Wang (2024). Specifically, we include the *DVPG* dummy, year indicators for the event window (*Year <= –5, Year –4, Year –2, Year 0, Year 1, Year 2, Year 3, Year 4, and Year >=5*), with *Year 0* being the reform year and *Year –1* as the benchmark year, and the interaction of year indicators and the *DVPG* dummy in the regression model, along with the control variables and industry and year fixed effects in the regression models. Our main variables of interest are the interaction terms between the year indicators and developing and developed status *(DVPG)*. For brevity, we plot the coefficients of the interaction terms and their respective statistical significance in Figure 1 for *ESG* and

MESG as the dependent variables, respectively.[26] The coefficients of the interaction terms are insignificant during the pre-reform period for both *ESG* (Figure 1, A) and *MESG* (Figure 1, B). The coefficients become significantly negative for year 1 and onwards for *ESG,* but remain insignificant for *MESG*. This suggests that Refinitiv's global benchmarking allocates further lower ESG scores to the firms in developing countries, mainly due to increased disclosure by the firms in the developed countries after the reforms. However, in *MESG,* we see no significant difference in the ESG scores among developing and developed country firms before and after mandatory ESG disclosure laws. This shows that laws effectively increased disclosures in the mandating countries despite their development status. This becomes evident when firms are compared with their respective country and industry peers using the modified ESG scores.

<Insert Figure 1 Here>

4.5 Robustness Checks

To further ensure the robustness of our findings, we perform several additional tests. First, we have argued that firm-fixed effects or country-fixed effects are perfectly correlated with the *DVPG* dummy and hence cannot be included in our analysis. Thus, we control for time-varying industry trends using *Industry* × *Year* fixed effects, and the results are qualitatively unchanged, confirming the robustness of our findings to alternate fixed effects. Second, we have relied on the IMF classification in our analysis as it reflects financial market realities. However, for robustness, we repeat our tests with the United Nations and World Bank classification of firms into developed and developing.[27] The results consistently show that developing country firms have lower ESG scores, reflecting both differing institutional priorities and methodology-driven "*score gap*". Third, as Table A3 shows, India and China represent more than 50% of the developing-country firms, and the U.S. represents around 40% of developed-country firms. This raises concerns that our results may reflect country-specific differences rather than group-level differences. To address this, we repeat our analysis excluding these countries and find results consistent with our main findings.[28]

5. Conclusion and Limitations

Using a global dataset of firms from 46 developed and developing countries, we demonstrate that corporate ESG scores are consistently lower in developing countries compared to developed ones. Our analysis reveals that this "*score gap*" is not solely due to differences in

[26] The tabulated regression results are available at request.
[27] See, UN classification. Accessed on August 18, 2024. See, WB classification. Accessed on August 18, 2024.
[28] Results of robustness checks are available upon request.

national priorities and poorer institutional settings but also stems from methodology-driven features of the ESG scoring process itself rather than from any systematic "*rating bias*".

By identifying Refinitiv's global benchmarking issue, our modified ESG (*MESG*) score addresses the conceptual limitation in applying a global benchmarking procedure across the various indicators to construct the ESG scores. Hence, *MESG* could serve as an appropriate measure of ESG performance that can be reliably used by investors and decision makers in assessing sustainability performance. We further suggest that Refinitiv modify the scoring procedure so as to eliminate the systemic differences arising in the corporate ESG scores of developed and developing country firms that are unrelated to actual sustainability performance.

While we provide evidence of a "*score gap*" in corporate ESG ratings in developing economies and provide a viable solution, we acknowledge several limitations of this study. First, our objective is not to examine specific ESG rating procedures or investigate detailed issues within rating processes that may contribute to this "*score gap*". Second, although differences exist in how ESG raters handle adverse versus missing disclosures, examining the impact of these discrepancies on score bias is beyond the scope of this study. Third, while ESG scores are derived from both internal (firm-generated) and external (third-party) sources, our analysis does not explore how these information sources might influence the variation in ESG scores between developed and developing countries. Instead, our focus is on the potential disparity in final ESG scores—the outcomes that ultimately guide investment decisions and government policies, regardless of the methodologies used to determine them.

Future research could explore how data sources and rating agency practices, including the handling of disclosures, contribute to "*score gap*" and/or "*rating bias*" affecting firms in developing countries. Since our conclusions are based on Refinitiv ESG scores, one may argue that the results may differ from other databases due to variation across ESG rating providers (Berg *et al.* 2022; Christensen *et al.* 2022). However, we believe that our findings with Refinitiv are more reliable due to its more comprehensive global coverage of firms, compared to other databases (Drempetic *et al.,* 2020; Basu *et al.,* 2022). Moreover, local ESG raters, with greater insight into the domestic institutional landscape, may be less prone to biases observed in global raters, who often emphasise comparability over context. Exploring how biases vary across global raters and between global and local raters could therefore offer a valuable direction for future research.

References

Adams, R.B., Hermalin, B.E. and Weisbach, M.S., 2010. The role of boards of directors in corporate governance: A conceptual framework and survey. *Journal of Economic Literature* 48(1), pp.58-107. https://doi.org/10.1257/jel.48.1.58

Agoraki, M. E. K., Kouretas, G. P., Wu, H., and Zhao, B. 2025. Imbalanced ESG investing?. *Journal of Corporate Finance* 93: 102810. https://doi.org/10.1016/j.jcorpfin.2025.102810

Almeida, H., I. Cunha, M. A. Ferreira, and F. Restrepo. 2017. The Real Effects of Credit Ratings: The Sovereign Ceiling Channel. *The Journal of Finance* 72 (1): 249–290. https://doi.org/10.1111/JOFI.12434

Apergis, N., T. Poufinas, and A. Antonopoulos. 2022. ESG scores and cost of debt. *Energy Economics* 112: 106186. https://doi.org/10.1016/J.ENECO.2022.106186

Asante-Appiah, B. and Lambert, T.A., 2023. The role of the external auditor in managing environmental, social, and governance (ESG) reputation risk. *Review of Accounting Studies*, *28*(4), pp.2589-2641. https://doi.org/10.1007/s11142-022-09706-z

Ayers, B. C., S. Ramalingegowda, and P. Eric Yeung. 2011. Hometown advantage: The effects of monitoring institution location on financial reporting discretion. *Journal of Accounting and Economics* 52 (1): 41–61. https://doi.org/10.1016/j.jacceco.2011.03.003

Ball, R., A. Robin, and J. S. Wu. 2003. Incentives versus standards: properties of accounting income in four East Asian countries. *Journal of Accounting and Economics* 36 (1–3): 235–270. https://doi.org/10.1016/J.JACCECO.2003.10.003

Basu, S., J. Vitanza, W. Wang, and X. R. Zhu. 2022. Walking the walk? Bank ESG disclosures and home mortgage lending. *Review of Accounting Studies* 27 (3): 779–821. https://doi.org/10.1007/s11142-022-09691-3

Berg, F., J. F. Kölbel, and R. Rigobon. 2022. Aggregate Confusion: The Divergence of ESG Ratings*. *Review of Finance* 26 (6): 1315–1344. https://doi.org/10.1093/rof/rfac033

Berg, F., K. Fabisik, and Z. Sautner. 2021. Is history repeating itself? The (un)predictable past of ESG ratings. *SSRN*. http://ssrn.com/abstract_id=3722087

Bhattacharya, C. B., S. Sen, L. M. Edinger-Schons, and M. Neureiter. 2023. Corporate Purpose and Employee Sustainability Behaviors. *Journal of Business Ethics* 183 (4): 963–981. https://doi.org/10.1007/s10551-022-05090-5

Bhattacharya, U., H. Daouk, and M. Welker. 2003. The World Price of Earnings Opacity. *The Accounting Review* 78 (3): 641–678. https://doi.org/10.2308/accr.2003.78.3.641

Black, B., A. G. de Carvalho, V. Khanna, W. Kim, and B. Yurtoglu. 2017. Corporate governance indices and construct validity. *Corporate Governance: An International Review* 25 (6): 397–410. https://doi.org/10.1111/corg.12215

Borensztein, E., K. Cowan, and P. Valenzuela. 2013. Sovereign ceilings "lite"? The impact of sovereign ratings on corporate ratings. *Journal of Banking & Finance* 37 (11): 4014–4024. https://doi.org/10.1016/j.jbankfin.2013.07.006

Boumparis, P., Florackis, C., Guedhami, O., and Sainani, S. 2025. Backing away from ESG? The effect of sovereign rating downgrades on corporate sustainability. *Journal of Corporate Finance* 94: 102856. https://doi.org/10.1016/j.jcorpfin.2025.102856

Cai, Y., C. H. Pan, and M. Statman. 2016. Why do countries matter so much in corporate social performance? *Journal of Corporate Finance* 41: 591–609. https://doi.org/10.1016/j.jcorpfin.2016.09.004

Campbell, J. L. 2007. Why Would Corporations Behave in Socially Responsible Ways? An Institutional Theory of Corporate Social Responsibility. *The Academy of Management Review* 32 (3): 946–967. https://doi.org/10.5465/amr.2007.25275684

Carrieri, F., V. Errunza, and K. Hogan. 2007. Characterizing World Market Integration through Time. *Journal of Financial and Quantitative Analysis* 42 (4): 915–940. https://doi.org/10.1017/S0022109000003446

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 (8): 1597–1614. https://doi.org/10.1002/smj.2407

Christensen, D. M., G. Serafeim, and A. Sikochi. 2022. Why is Corporate Virtue in the Eye of The Beholder? The Case of ESG Ratings. *The Accounting Review* 97:147–175. https://doi.org/10.2308/TAR-2019-0506

Cornaggia, J.N., Cornaggia, K.J. and Israelsen, R.D., 2020. Where the heart is: Information production and the home bias. *Management Science*, 66(12), pp.5532-5557. https://doi.org/10.1287/mnsc.2019.3438

De Moor, L., P. Luitel, P. Sercu, and R. Vanpée. 2018. Subjectivity in sovereign credit ratings. *Journal of Banking & Finance* 88: 366–392. https://doi.org/10.1016/j.jbankfin.2017.12.014

Deci, E. L., and R. M. Ryan. 1985. Intrinsic Motivation and Self-Determination in Human Behavior. Springer US.

Didier, T., R. Rigobon, and S. L. Schmukler. 2013. Unexploited gains from international diversification: Patterns of portfolio holdings around the world. *The Review of Economics and Statistics* 95 (5): 1562–1583. https://doi.org/10.1162/REST_a_00351

Doidge, C., G. Andrew Karolyi, and R. M. Stulz. 2007. Why do countries matter so much for corporate governance? *Journal of Financial Economics* 86 (1): 1–39. https://doi.org/10.1016/j.jfineco.2006.09.002

Drempetic, S., C. Klein, and B. Zwergel. 2020. The Influence of Firm Size on the ESG Score: Corporate Sustainability Ratings Under Review. *Journal of Business Ethics* 167 (2): 333–360. https://doi.org/10.1007/s10551-019-04164-1

Eccles, R. G., I. Ioannou, and G. Serafeim. 2014. The impact of corporate sustainability on organizational processes and performance. *Management Science* 60 (11): 2835–2857. https://doi.org/10.1287/mnsc.2014.1984

European Commission: Directorate-General for Financial Stability, Financial Services and Capital Markets Union and ERM, *Study on sustainability-related ratings, data and research*, Publications Office of the European Union, 2021, https://data.europa.eu/doi/10.2874/14850

Fauver, L., Hung, M., Taboada, A.G. and Wang, E.J., 2024. Boardroom gender diversity reforms and institutional monitoring: global evidence. *Review of Accounting Studies*, *29*(1): 621-664. https://doi.org/10.1007/s11142-022-09710-3

Fuchs, A., and K. Gehring. 2017. The Home Bias in Sovereign Ratings. *Journal of the European Economic Association* 15 (6): 1386–1423. https://doi.org/10.1093/jeea/jvx009

Goar, M. 2024. COP29: Alone at the podium, developing countries call for action and funding. *Le Monde*. https://www.lemonde.fr/en/environment/article/2024/11/13/cop29-developing-countries-call-for-action-and-funding_6732664_114.html

Grossman, G. M., and A. B. Krueger. 1991. Environmental Impacts of a North American Free Trade Agreement. *NBER Working Paper*.

Grossman, G. M., and A. B. Krueger. 1995. Economic Growth and the Environment. *The Quarterly Journal of Economics* 110 (2): 353–377. https://doi.org/10.3386/w3914

Hail, L., and C. Leuz. 2006. International Differences in the Cost of Equity Capital: Do Legal Institutions and Securities Regulation Matter? *Journal of Accounting Research* 44 (3): 485–531. https://doi.org/10.1111/j.1475-679X.2006.00209.x

Hainmueller, J. 2012. Entropy Balancing for Causal Effects: A Multivariate Reweighting Method to Produce Balanced Samples in Observational Studies. *Political Analysis* 20 (1): 25–46. https://doi.org/10.1093/pan/mpr025

Hair, Joseph F., Black, William C., Babin, Barry J., and Anderson, Rolph E. 2018. *Multivariate Data Analysis*. 8th ed. Cengage Learning.

Hartzmark, S. M., and A. B. Sussman. 2019. Do Investors Value Sustainability? A Natural Experiment Examining Ranking and Fund Flows. *The Journal of Finance* 74 (6): 2789–2837. https://doi.org/10.1111/jofi.12841

Heinkel, R., A. Kraus, and J. Zechner. 2001. The Effect of Green Investment on Corporate Behavior. *Journal of Financial and Quantitative Analysis* 36 (4): 431–449. https://doi.org/10.2307/2676219

Hill, A. D., S. G. Johnson, L. M. Greco, E. H. O'Boyle, and S. L. Walter. 2021. Endogeneity: A Review and Agenda for the Methodology-Practice Divide Affecting Micro and Macro Research. *Journal of Management* 47 (1): 105–143. https://doi.org/10.1177/0149206320960533

IMF. 2021. Global Financial Stability Report: Covid-19, Crypto, And Climate: Navigating Challenging Transitions. https://digitallibrary.un.org/record/4019919?v=pdf

IMF. 2023. *World Economic Outlook Database - Groups and Aggregates*. Available at: https://www.imf.org/en/Publications/WEO/weo-database/2023/April/groups-and-aggregates

Ioannou, I., and G. Serafeim. 2012. What drives corporate social performance the role of nation-level institutions. *Journal of International Business Studies* 43 (9): 834–864. https://doi.org/10.1057/jibs.2012.26

Isidro, H., D. Nanda, and P. D. Wysocki. 2020. On the Relation between Financial Reporting Quality and Country Attributes: Research Challenges and Opportunities. *The Accounting Review* 95 (3): 279–314. https://doi.org/10.2308/accr-52607

Jackson, G., and A. Apostolakou. 2010. Corporate social responsibility in Western Europe: An institutional mirror or substitute? *Journal of Business Ethics* 94 (3): 371–394. https://doi.org/10.1007/s10551-009-0269-8

Kaiser, H. F. 1960. The Application of Electronic Computers to Factor Analysis. *Educational and Psychological Measurement* 20 (1): 141–151. https://doi.org/10.1177/001316446002000116

Kentikelenis, A. E., T. H. Stubbs, and L. P. King. 2016. IMF conditionality and development policy space, 1985–2014. *Review of International Political Economy* 23 (4): 543–582. https://doi.org/10.1080/09692290.2016.1174953

Kim, I., S. Miller, H. Wan, and B. Wang. 2016. Drivers behind the monitoring effectiveness of global institutional investors: Evidence from earnings management. *Journal of Corporate Finance* 40: 24–46. https://doi.org/10.1016/j.jcorpfin.2016.06.006

Krueger, P., Z. Sautner, D. Y. Tang, and R. Zhong. 2024. The Effects of Mandatory ESG Disclosure Around the World. *Journal of Accounting Research* 62: 1795–1847. https://doi.org/10.1111/1475-679X.12548

Larcker, D. F., B. Tayan, and E. M. Watts. 2022. Seven myths of ESG. *European Financial Management* 28 (4): 869–882. https://doi.org/10.1111/eufm.12378

Leuz, C., D. Nanda, and P. D. Wysocki. 2003. Earnings management and investor protection: An international comparison. *Journal of Financial Economics* 69 (3): 505–527. 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 (2): 853–910. https://doi.org/10.1111/jofi.12487

LSEG. 2024. Environmental, Social and Governance scores from LSEG. https://www.lseg.com/content/dam/data-analytics/en_us/documents/methodology/lseg-esg-scores-methodology.pdf

Maignan, I. 2001. Consumers' perceptions of corporate social responsibilities: A cross-cultural comparison. *Journal of Business Ethics* 30 (1): 57–72. https://doi.org/10.1023/A:1006433928640

Maignan, I., and D. A. Ralston. 2002. Corporate Social Responsibility in Europe and the U.S.: Insights from Businesses' Self-presentations. *Journal of International Business Studies* 33 (3): 497–514. https://doi.org/10.1057/palgrave.jibs.8491028

Maslow, A. H. 1943. A theory of human motivation. *Psychological Review* 50 (4): 370–396. https://doi.org/10.1037/h0054346

Maslow, A. H. 1954. Motivation and personality. 3rd Edition. New York: Harper & Row.

Michelson, G., N. Wailes, S. Van Der Laan, and G. Frost. 2004. Ethical investment processes and outcomes. *Journal of Business Ethics* 52 (1): 1–10. https://doi.org/10.1023/B:BUSI.0000033103.12560.be

Mitton, T. 2024. Economic Significance in Corporate Finance. *Review of Corporate Finance Studies* 13 (1): 38–79. https://doi.org/10.1093/rcfs/cfac008

Nofsinger, J. and Varma, A. 2014. Socially responsible funds and market crises. *Journal of Banking & Finance*, *48*, pp.180-193. https://doi.org/10.1016/j.jbankfin.2013.12.016

Official Journal of European Union. 2024. *Regulation (EU) 2024/3005 of the European Parliament and of the Council of 27 November 2024 on the transparency and integrity of Environmental, Social and Governance (ESG) rating activities, and amending Regulations (EU) 2019/2088 and (EU) 2023/2859*. http://data.europa.eu/eli/reg/2024/3005/oj

Oster, E. 2019. Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business and Economic Statistics* 37 (2): 187–204. https://doi.org/10.1080/07350015.2016.1227711

Pagell, M., Z. Wu, and M. E. Wasserman. 2010. Thinking differently about purchasing portfolios: An assessment of sustainable sourcing. *Journal of Supply Chain Management* 46 (1): 57–73. https://doi.org/10.1111/j.1745-493X.2009.03186.x

Panayotou, T. 1993. Empirical tests and policy analysis of environmental degradation at different stages of economic development. *International Labour Organization (ILO) Working Papers*.

Parguel, B., Benoît-Moreau, F. and Larceneux, F. 2011. How sustainability ratings might deter 'greenwashing': A closer look at ethical corporate communication. *Journal of Business Ethics* 102(1), pp.15-28.

Pinto, T. N. 2024. Africa has little use for ESG. GIS Reports Online. https://www.gisreportsonline.com/r/africa-esg/

Porter, M.E. and Kramer, M.R. 2006. The link between competitive advantage and corporate social responsibility. *Harvard Business Review*, *84*(12), pp.78-92.

Powell, W. W., and P. J. DiMaggio. 1991. The new institutionalism in organizational analysis. The University of Chicago Press.

Raghunandan, A., and S. Rajgopal. 2022. Do ESG funds make stakeholder-friendly investments? *Review of Accounting Studies* 27 (3): 822–863. https://doi.org/10.1007/s11142-022-09693-1

Renneboog, L., J. Ter Horst, and C. Zhang. 2008. Socially responsible investments: Institutional aspects, performance, and investor behavior. *Journal of Banking & Finance* 32 (9): 1723–1742. https://doi.org/10.1016/j.jbankfin.2007.12.039

Roberts, M.R. and Whited, T.M. 2013. Endogeneity in empirical corporate finance. In *Handbook of the Economics of Finance* (Vol. 2, pp. 493-572). Elsevier. https://doi.org/10.1016/B978-0-44-453594-8.00007-0

Ryan, R. M., and E. L. Deci. 2017. Self-Determination Theory: Basic Psychological Needs in Motivation, Development, and Wellness. Guilford Press.

Schmidt, C. G., K. Foerstl, and B. Schaltenbrand. 2017. The supply chain position paradox: Green practices and firm performance. *Journal of Supply Chain Management* 53 (1): 3–25. https://doi.org/10.1111/jscm.12113

SEBI. 2022. Consultation Paper on Environmental, Social and Governance (ESG) Rating Providers for Securities Markets. https://www.sebi.gov.in/reports-and-statistics/reports/jan-2022/consultation-paper-on-environmental-social-and-governance-esg-rating-providers-for-securities-markets_55516.html

SEBI. 2023. Consultation Paper on ESG Disclosures, Ratings and Investing. https://www.sebi.gov.in/reports-and-statistics/reports/feb-2023/consultation-paper-on-esg-disclosures-ratings-and-investing_68193.html

Semenova, N., and L. G. Hassel. 2015. On the Validity of Environmental Performance Metrics. *Journal of Business Ethics* 132 (2): 249–258. https://doi.org/10.1007/s10551-014-2323-4

Shafik, N., and S. Bandyopadhyay. 1992. Economic Growth and Environmental Quality: Time-series and Cross-country Evidence. World Bank Publications.

Shroff, N., R. S. Verdi, and B. P. Yost. 2017. When does the peer information environment matter? *Journal of Accounting and Economics* 64 (2–3): 183–214. https://doi.org/10.1016/j.jacceco.2017.03.005

Sparkes, R., and C. J. Cowton. 2004. The maturing of socially responsible investment: A review of the developing link with corporate social responsibility. *Journal of Business Ethics* 52 (1): 45–57. https://doi.org/10.1023/B:BUSI.0000033106.43260.99

Statman, M., and D. Glushkov. 2009. The Wages of Social Responsibility. *Financial Analysts Journal* 65 (4): 33–46. https://doi.org/10.2469/faj.v65.n4.5

Stern, D. I., M. S. Common, and E. B. Barbier. 1996. Economic growth and environmental degradation: The environmental Kuznets curve and sustainable development. *World Development* 24 (7): 1151–1160. https://doi.org/10.1016/0305-750X(96)00032-0

United Nations. 2021. World economic situation and prospects 2021. *United Nations Publications*. https://www.un.org/en/world-economic-situation-and-prospects-2021

World Bank. 2022. World Development Indicators DataBank. https://databank.worldbank.org/source/world-development-indicators

List of Tables and Figures

Table 1
ESG Categories

Pillars	Categories	Number of indicators
(1)	(2)	(3)
Environment	Emission	28
	Innovation	20
	Resource Use	20
Social	Community	14
	Human Rights	8
	Product Responsibility	10
	Workforce	30
Governance	CSR Strategy	9
	Management	35
	Shareholders	12
Total		186

Notes. This table reports the number of indicators under each category for environmental, social and governance pillar scores.

Table 2
Variables Description

Variables	Definitions
ESG	Firm-level overall ESG score obtained from Refinitiv.
ENV	Firm-level environmental pillar score obtained from Refinitiv.
SOC	Firm-level social pillar score obtained from Refinitiv.
GOV	Firm-level governance pillar score obtained from Refinitiv.
MESG	Firm-level modified ESG score calculated using *MENV*, *MSOC*, and *MGOV* scores following Refinitiv's aggregation approach.
MENV	Firm-level modified environmental score constructed by rescaling the 68 indicators and category weights of *ENV* within each country-industry-year.
MSOC	Firm-level modified social score constructed by rescaling the 62 indicators and category weights of *SOC* within each country-industry-year.
MGOV	Firm-level modified governance score constructed by rescaling the 56 indicators and category weights of *GOV* within each country-industry-year.
DVPG	Indicator variable equals one if the country is classified as a developing country by the International Monetary Fund' 2023 classification, and zero otherwise.[29]
LIQ	The ratio of the difference between current assets and current liabilities over total assets at the end of the fiscal year.
RDEXP	The ratio of R&D expenditure over total assets at the end of the fiscal year.
LEV	The ratio of total debt to total assets at the end of the fiscal year.
SIZE	Firm size is calculated as the natural log of the firm's assets at the end of the fiscal year.
RSIZE	The total assets of a firm in a given year divided by the sum of the total assets of all firms in the same industry and year.
ROA	Return on assets (income before extraordinary items) divided by average total assets.
GROWTH	Change in sales scaled by lagged total sales.
CAPEX	The ratio of capital expenditure to total assets at the end of the fiscal year.
DIV	An indicator that equals one (and zero otherwise) if the firm has dividend payout at the end of fiscal year.

[29] Available at: IMF Classification, 2023.

Table 3
Difference of ESG Scores between Developed and Developing Countries

Variables	Mean		Difference
	Developing	Developed	(2) – (3)
(1)	(2)	(3)	(4)
ESG	0.452	0.475	−0.022***
MESG	0.516	0.524	−0.008***

Notes: This table presents the *t*-test results comparing *ESG* and *MESG* between developed and developing countries. The sample is based on the annual data of firms in 46 developed and developing countries from 2009 to 2023. Significance levels are indicated by ***, **, and * for the 1%, 5%, and 10% levels, respectively. Standard errors are clustered at the firm level and reported in parentheses. All variables are defined in Table 2.

Table 4
Univariate Logit Regressions of Country Factors Across Developing-Developed Classifications

Factors	*DVPG IMF*		*DVPG World Bank*		*DVPG United Nations*	
	Coefficient	Pseudo R^2	Coefficient	Pseudo R^2	Coefficient	Pseudo R^2
(1)	(2)	(3)	(4)	(5)	(6)	(7)
Factor 1	**−11.172***	**0.886**	**−3.288***	**0.570**	**−3.429***	**0.629**
	(4.112)		**(1.166)**		**(0.876)**	
Factor 2	−0.215	0.008	−0.073	0.001	0.057	0.001
	(0.298)		(0.302)		(0.323)	
Factor 3	0.253	0.011	0.298	0.015	0.430	0.030
	(0.306)		(0.342)		(0.329)	
Factor 4	−0.380	0.017	−0.174	0.004	1.346***	0.078
	(0.318)		(0.245)		(0.651)	
Factor 5	0.041	0.000	−0.667	0.045	0.356	0.019
	(0.308)		(0.743)		(0.271)	
Factor 6	0.036	0.000	−0.058	0.001	−0.331	0.018
	(0.312)		(0.378)		(0.331)	
Factor 7	−0.161	0.005	−0.423	0.030	−0.224	0.009
	(0.311)		(0.368)		(0.302)	
Factor 8	−0.144	0.004	0.065	0.001	0.266	0.012
	(0.308)		(0.363)		(0.289)	
Factor 9	−0.458	0.033	−0.459	0.030	−0.020	0.000
	(0.364)		(0.367)		(0.292)	
Factor 10	−0.324	0.018	0.140	0.003	−0.477	0.037
	(0.325)		(0.307)		(0.325)	
Factor 11	−0.124	0.003	0.003	0.000	−0.060	0.001
	(0.302)		(0.238)		(0.334)	
Factor 12	−0.079	0.001	−0.373	0.022	−0.085	0.001
	(0.306)		(0.331)		(0.304)	
Factor 13	0.119	0.002	0.069	0.001	0.308	0.016
	(0.306)		(0.342)		(0.291)	
Factor 14	−0.081	0.001	−0.034	0.000	0.029	0.000
	(0.305)		(0.345)		(0.304)	

Notes: This table presents the results for the univariate logit regressions with the robust standard errors presented in parentheses. The sample includes the 46 developing and developed countries. The independent variables are the 14 country-level latent factors obtained through the principal component analysis of 64 country-level variables, and the dependent variable is *the DVPG* indicator, according to the IMF (columns 2 & 3), World Bank (columns 4 & 5), and United Nations (columns 6 & 7).

Table 5
Multivariate Regressions of ESG Scores

Variables	*ESG*	*MESG*
(1)	(2)	(3)
DVPG	**−0.034*** **	**−0.002**
	(0.006)	**(0.004)**
SIZE	0.060***	0.047***
	(0.002)	(0.001)
LEV	0.000	−0.008
	(0.016)	(0.012)
RDEXP	−0.058	0.144
	(0.123)	(0.096)
CAPEX	0.077	0.112***
	(0.048)	(0.038)
ROA	0.256***	0.153***
	(0.026)	(0.021)
GROWTH	−0.039***	−0.016***
	(0.005)	(0.004)
LIQ	−0.064***	−0.036***
	(0.015)	(0.012)
DIV	−0.002	0.001
	(0.006)	(0.004)
Constant	−0.830***	−0.505***
	(0.038)	(0.029)
Observations	50,327	50,327
Adj. R2	0.224	0.173
Year FE	Yes	Yes
Industry FE	Yes	Yes

Notes: This table presents entropy-balanced multivariate regression estimates, with *ESG* and *MESG* scores as dependent variables and the *DVPG* dummy variable as the independent variable of primary interest. All continuous variables have been winsorized at the 1st and 99th percentiles except ESG scores. The sample is based on the annual data of firms in 46 developed and developing countries from 2009 to 2023. Significance levels are indicated by ***, **, and * for the 1%, 5%, and 10% levels, respectively. Standard errors are clustered at the firm level and reported in parentheses. All variables are defined in Table 2.

Table 6
Multivariate Regression of Component Scores

Variables	Refinitiv's Scores			Modified Scores		
	ENV	*SOC*	*GOV*	*MENV*	*MSOC*	*MGOV*
(1)	(2)	(3)	(4)	(5)	(6)	(7)
DVPG	**−0.034*****	**−0.054*****	**−0.006**	**0.005**	**0.000**	**−0.011****
	(0.007)	**(0.007)**	**(0.006)**	**(0.005)**	**(0.005)**	**(0.005)**
SIZE	0.085***	0.059***	0.034***	0.057***	0.052***	0.029***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
LEV	−0.024	0.009	0.003	−0.027*	−0.013	0.012
	(0.020)	(0.020)	(0.017)	(0.015)	(0.014)	(0.014)
RDEXP	0.136	−0.417***	0.116	0.176	0.281**	−0.055
	(0.142)	(0.147)	(0.145)	(0.119)	(0.118)	(0.116)
CAPEX	0.040	0.092	0.088*	0.125***	0.154***	0.044
	(0.060)	(0.060)	(0.053)	(0.047)	(0.046)	(0.045)
ROA	0.226***	0.345***	0.161***	0.147***	0.194***	0.100***
	(0.033)	(0.033)	(0.026)	(0.025)	(0.026)	(0.023)
GROWTH	−0.051***	−0.047***	−0.019***	−0.018***	−0.016***	−0.012**
	(0.006)	(0.006)	(0.005)	(0.005)	(0.005)	(0.005)
LIQ	−0.085***	−0.083***	−0.021	−0.056***	−0.041***	−0.010
	(0.019)	(0.019)	(0.016)	(0.014)	(0.014)	(0.013)
DIV	0.010	−0.018**	0.007	−0.002	0.000	0.004
	(0.007)	(0.007)	(0.006)	(0.005)	(0.005)	(0.005)
Constant	−1.425***	−0.795***	−0.247***	−0.719***	−0.625***	−0.117***
	(0.046)	(0.050)	(0.040)	(0.036)	(0.036)	(0.034)
Observations	50,327	50,327	50,327	50,327	50,327	50,327
Adj. R2	0.278	0.171	0.070	0.164	0.149	0.051
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table presents entropy-balanced multivariate regression estimates, with ESG component scores as the dependent variables and the *DVPG* dummy variable as the independent variable of primary interest. All continuous variables have been winsorized at the 1st and 99th percentiles except ESG scores. Columns (2) to (4) show models using Refinitiv's ESG component scores as the dependent variable, columns (5) to (7) display models using modified ESG component scores. All continuous variables have been winsorized at the 1st and 99th percentiles except ESG scores. The sample is based on the annual data of firms in 46 developed and developing countries from 2009 to 2023. Significance levels are indicated by ***, **, and * for the 1%, 5%, and 10% levels, respectively. Standard errors are clustered at the firm level and reported in parentheses. All variables are defined in Table 2.

Table 7
Oster Coefficient Stability Test

	(1)	*ESG* (2)
(1)	*DVPG*	−0.034***
(2)	R–Squared	0.225
(3)	*Δ*	1
(4)	$R_{max} = 1.3 \times R^2_{baseline}$	0.293
(5)	**Bounds on Treatment effect**	**(−0.034, −0.039)**
(6)	Treatment effect excludes 0	Yes
(7)	**Oster's *δ***	**2.079**
(8)	$\mid \delta \mid > 1$	Yes

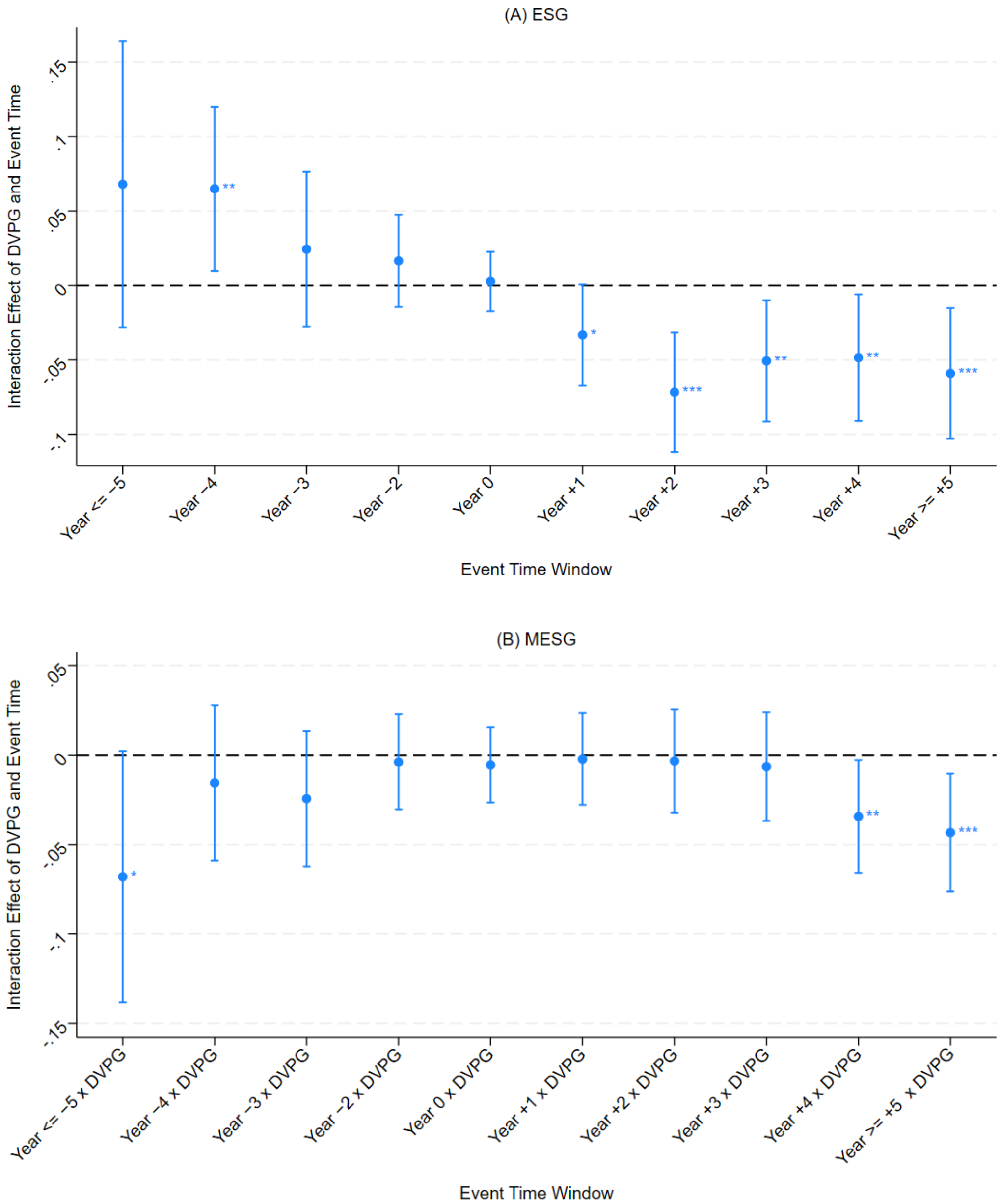
Notes: This table reports the results of Oster's (2019) approach to estimate the robustness to omitted variable bias. Rows (1) and (2) present the coefficient of *DPVG* and R-squared estimated from our baseline multivariate regressions. Row (3) and (4) present the assumption of δ and Rmax, we define Rmax as 1.3 times R-squared. Rows (5) and (6) report the bounds on the coefficient of *DPVG,* which is estimated using the Stata syntax psacalc. Rows (7) and (8) report the value of δ. Significance levels are indicated by ***, **, and * for the 1%, 5%, and 10% levels, respectively. Standard errors have been clustered at the firm level.

Table 8
Effect of Mandatory ESG Disclosures

Variables	*ESG*	*MESG*
(1)	(2)	(3)
DVPG	−0.002	0.010
	(0.023)	(0.016)
ESG MD	−0.024**	0.029***
	(0.011)	(0.008)
DVPG × ESG MD	**−0.079*** **	**−0.023**
	(0.023)	**(0.016)**
Constant	−0.002	0.010
	(0.023)	(0.016)
Observations	25,564	25,564
Adj. R2	0.239	0.148
Controls	Yes	Yes
Year FE	Yes	Yes
Industry FE	Yes	Yes

Notes: This table presents the results of a staggered difference-in-difference (DiD) analysis examining the effect of mandatory adoption of ESG disclosure with *ESG* and *MESG* as dependent variables and *DVPG*, *ESG MD,* and the interaction of both (*DVPG × ESG MD*) as the independent variables. Significance levels are indicated by ***, **, and * for the 1%, 5%, and 10% levels, respectively. All continuous variables have been winsorized at the 1st and 99th percentiles except ESG scores. The sample is based on the annual data of firms in 46 developed and developing countries from 2009 to 2023. Standard errors are clustered at the firm level and reported in parentheses. All variables are defined in Table 2.

Figure 1
Event Time Effects of Interaction Between ESG Mandates and Country Status



Notes: This table plots the coefficients of interaction effects of *DVPG* and even time dummies presented on the x-axis around the ESG mandates. The models include *DVPG* as the main effects, event time dummies and the interaction of event time dummies with *DVPG* as independent variables, and *ESG*, *MESG* as dependent variables in Fig. A and B, respectively. Significance levels are indicated by ***, **, and * for the 1%, 5%, and 10% levels, respectively. The bars above and below the coefficient reflect a 95% confidence interval. The sample spans between 2009 and 2023. All continuous variables have been winsorized at the 1st and 99th percentiles except ESG scores. All variables are defined in Table 2.

Appendix

Table A1
Sample Selection Procedure

Panel A: Sample Selection Procedure			
	Filters	Number of Observations	Unique Firms (RIC)
(1)	(2)	(3)	(4)
	Firms under Refinitiv ESG Universe (Private & Public, Active & Inactive between 2002 and 2023)	102,079	13,323
Less	Private Firms	(2,610)	(1,285)
		99,469	12,038
Less	Academic & Educational Services	(302)	(48)
Less	Financials	(15,717)	(1,755)
Less	Government Activity	(1)	(1)
Less	Real Estate	(6,860)	(826)
Less	Utilities	(4,081)	(412)
		72,508	8,996
Less	Country-Years less than 15 observations	(3,344)	(162)
Less	Countries not classified (Bermuda, Cayman Islands, Jersey)	(2,582)	(359)
Less	Missing Observations for Dependent, Independent and Lagged Control Variables	(12,048)	(505)
		54,534	7,970
Less	Firm-Year Observations before 2009	(4,207)	(1)
Total	**Sample**	**50,327**	**7,969**

Note: This table outlines the sample selection procedure.

Table A2

Sample Distribution Over Time

Year	Full Sample		Developing		Developed	
	Obs.	%	Obs.	%	Obs.	%
(1)	(2)	(3)	(4)	(5)	(6)	(7)
2009	1,232	2.450	30	0.280	1,202	3.020
2010	1,411	2.800	82	0.780	1,329	3.340
2011	1,771	3.520	223	2.110	1,548	3.890
2012	1,875	3.730	272	2.580	1,603	4.030
2013	1,987	3.950	329	3.120	1,658	4.170
2014	2,068	4.110	350	3.320	1,718	4.320
2015	2,132	4.240	365	3.460	1,767	4.440
2016	2,517	5.000	379	3.590	2,138	5.380
2017	2,924	5.810	433	4.100	2,491	6.260
2018	3,536	7.030	609	5.770	2,927	7.360
2019	4,119	8.180	722	6.840	3,397	8.540
2020	4,936	9.810	1,072	10.160	3,864	9.720
2021	5,935	11.790	1,339	12.690	4,596	11.560
2022	6,629	13.170	1,842	17.450	4,787	12.040
2023	7,255	14.420	2,508	23.760	4,747	11.940
Total	**50,327**		**10,555**		**39,772**	

Note: This table outlines the sample distribution over time across full sample and the developing and developed countries. The sample contains firm-year observations for 46 developed and developing countries from 2009 to 2023.

Table A3
Country-wise Distribution of *ESG* and *MESG* Scores

Countries	Obs.	Firms	Mean		Median	
			ESG	*MESG*	*ESG*	*MESG*
(1)	(2)	(3)	(4)	(5)	(6)	(7)
Argentina	168	31	0.373	0.506	0.421	0.500
Brazil	650	83	0.543	0.510	0.567	0.500
Chile	162	26	0.474	0.516	0.520	0.500
China	3,747	956	0.379	0.517	0.358	0.514
Egypt	13	13	0.260	0.505	0.172	0.500
India	1,475	562	0.506	0.528	0.506	0.510
Indonesia	411	59	0.472	0.505	0.471	0.500
Malaysia	915	347	0.430	0.526	0.418	0.515
Marshall Islands	14	14	0.336	0.498	0.331	0.500
Mexico	392	63	0.493	0.513	0.529	0.500
Morocco	45	29	0.397	0.524	0.419	0.500
Peru	115	21	0.442	0.510	0.478	0.500
Philippines	37	19	0.516	0.500	0.531	0.500
Poland	98	22	0.489	0.499	0.506	0.500
Qatar	59	21	0.274	0.497	0.247	0.500
Russia	282	30	0.431	0.504	0.445	0.500
Saudi Arabia	148	52	0.353	0.518	0.376	0.500
South Africa	856	79	0.527	0.505	0.543	0.500
Thailand	524	124	0.558	0.515	0.577	0.500
Turkey	385	81	0.605	0.523	0.658	0.500
United Arab Emirates	59	34	0.341	0.499	0.329	0.500
Total Developing	**10,555**	**2,666**				
Australia	2,223	274	0.393	0.529	0.365	0.525
Austria	113	24	0.598	0.518	0.591	0.500
Belgium	257	36	0.567	0.512	0.570	0.500
Canada	2,519	302	0.430	0.523	0.414	0.522
Denmark	397	52	0.532	0.508	0.546	0.500
Finland	450	67	0.576	0.507	0.592	0.500
France	1,317	153	0.612	0.521	0.647	0.515
Germany	1,439	219	0.560	0.541	0.566	0.537
Hong Kong	294	29	0.463	0.500	0.474	0.500
Ireland	466	44	0.523	0.507	0.531	0.500
Israel	123	33	0.396	0.501	0.381	0.500
Italy	346	82	0.568	0.532	0.582	0.500
Japan	4,859	465	0.482	0.510	0.506	0.518
Luxembourg	119	29	0.600	0.512	0.608	0.500
Netherlands	618	73	0.585	0.526	0.620	0.500
New Zealand	256	40	0.417	0.508	0.390	0.500
Norway	282	59	0.535	0.508	0.542	0.500
Singapore	388	41	0.440	0.506	0.431	0.500
South Korea	1,384	165	0.493	0.512	0.561	0.514
Spain	463	51	0.631	0.515	0.657	0.500
Sweden	1,226	242	0.510	0.522	0.531	0.509

Switzerland	945	133	0.393	0.529	0.365	0.525
Taiwan	1,513	164	0.598	0.518	0.591	0.500
United Kingdom	2,884	391	0.567	0.512	0.570	0.500
United States	14,891	2,135	0.430	0.523	0.414	0.522
Total Developed	**39,772**	**5,303**				
Total Sample	**50,327**	**7,969**				

Notes: This table reports the country-wise main summary statistics for *ESG* and *MESG*. The sample contains firm-year observations for 46 developed and developing countries from 2009 to 2023. All variables are defined in Table 2.

Table A4
Univariate Regressions of ESG Scores

Variables	*ESG*	*MESG*
(1)	(2)	(3)
DVPG	−0.023***	−0.004*
	(0.002)	(0.002)
SIZE	0.062***	0.051***
	(0.000)	(0.000)
LEV	0.086***	0.080***
	(0.005)	(0.005)
RDEXP	−0.345***	−0.077***
	(0.015)	(0.016)
CAPEX	−0.127***	−0.085***
	(0.020)	(0.021)
ROA	0.291***	0.165***
	(0.007)	(0.007)
GROWTH	−0.031***	−0.011***
	(0.003)	(0.003)
LIQ	−0.181***	−0.129***
	(0.004)	(0.005)
DIV	0.110***	0.073***
	(0.002)	(0.002)

Notes: This table reports univariate regression estimates employing *ESG* and *MESG* as dependent variables. Significance levels are indicated by ***, **, and * for the 1%, 5%, and 10% levels, respectively. The sample comprises annual data of firms in 46 developed and developing countries from 2009 to 2023. All variables are defined in Table 2.

Table A5
Summary Statistics

Variables	Country Classification	Mean	SD	Median	Min	Max
(1)	(2)	(3)	(4)	(5)	(6)	(7)
			ESG Scores			
ESG	All	0.470	0.208	0.472	0.001	0.956
	Developing	0.452	0.197	0.453	0.001	0.948
	Developed	0.475	0.211	0.477	0.003	0.956
MESG	All	0.523	0.175	0.508	0.043	0.975
	Developing	0.516	0.161	0.500	0.043	0.946
	Developed	0.525	0.178	0.513	0.047	0.975
			Control Variables			
SIZE	All	21.765	1.736	21.819	17.026	25.772
	Developing	21.712	1.582	21.717	17.026	25.772
	Developed	21.779	1.775	21.846	17.026	25.772
LEV	All	0.248	0.182	0.231	0.000	0.855
	Developing	0.244	0.178	0.232	0.000	0.855
	Developed	0.249	0.183	0.230	0.000	0.855
RDEXP	All	0.022	0.060	0.000	0.000	0.440
	Developing	0.008	0.019	0.000	0.000	0.315
	Developed	0.026	0.066	0.000	0.000	0.440
CAPEX	All	0.050	0.046	0.037	0.001	0.252
	Developing	0.056	0.047	0.043	0.001	0.252
	Developed	0.048	0.045	0.035	0.001	0.252
ROA	All	0.028	0.137	0.044	−0.831	0.305
	Developing	0.064	0.085	0.057	−0.831	0.305
	Developed	0.018	0.146	0.041	−0.831	0.305
GROWTH	All	0.081	0.279	0.025	−0.620	1.661
	Developing	0.080	0.269	0.000	−0.620	1.661
	Developed	0.081	0.282	0.031	−0.620	1.661
LIQ	All	0.172	0.204	0.142	−0.266	0.860
	Developing	0.157	0.200	0.136	−0.266	0.860
	Developed	0.176	0.205	0.143	−0.266	0.860
DIV	All	0.724	0.447	1.000	0.000	1.000
	Developing	0.871	0.336	1.000	0.000	1.000
	Developed	0.685	0.464	1.000	0.000	1.000

Notes: This table reports summary statistics for all variables used in the multivariate analysis. All variables are winsorized at their 1st and 99th percentiles except ESG scores. The sample contains firm-year observations for 46 developed and developing countries from 2009 to 2023. All variables are defined in Table 2.

Table A6

Correlation Matrix

Variables		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(1)	*DVPG*	1.000								
(2)	*SIZE*	−0.016	1.000							
(3)	*LEV*	−0.010	0.219	1.000						
(4)	*RDEXP*	−0.124	−0.262	−0.103	1.000					
(5)	*CAPEX*	0.073	0.001	0.038	−0.116	1.000				
(6)	*ROA*	0.135	0.267	−0.122	−0.513	0.067	1.000			
(7)	*GROWTH*	−0.002	−0.040	−0.033	0.070	0.044	0.072	1.000		
(8)	*LIQ*	−0.037	−0.352	−0.415	0.367	−0.198	−0.098	0.062	1.000	
(9)	*DIV*	0.169	0.356	−0.034	−0.301	0.015	0.374	−0.098	−0.143	1.000

Notes: This table reports the correlation matrix for all independent variables used in the multivariate regressions for hypothesis testing. All continuous variables have been winsorized at their 1st and 99th percentiles. The sample contains firm-year observations for 46 developed and developing countries from 2009 to 2023. All variables are defined in Table 2.

Table A7
Variation Explained by Country Latent Factors

Factor	Eigen Value	% of Variance	Cumulative %	Reliability Test (Cronbach's Alpha)
(1)	(2)	(3)	(4)	(5)
1	22.830	35.672	35.672	0.979
2	8.551	13.361	49.033	0.839
3	3.189	4.983	54.016	0.829
4	3.006	4.698	58.714	0.895
5	2.538	3.965	62.679	
6	2.375	3.711	66.390	
7	2.005	3.132	69.522	
8	1.925	3.008	72.530	
9	1.733	2.707	75.237	
10	1.518	2.372	77.609	
11	1.355	2.117	79.726	
12	1.190	1.860	81.586	
13	1.120	1.750	83.335	
14	1.029	1.607	84.943	
15	0.963	1.505	86.448	

Notes: This table presents the results for the principal component analysis on 64 country-level variables for 46 developing and developed countries. Column 1 reflects the country-level latent factors; column 2 shows the eigenvalues for the corresponding factor. Columns 3 and 4 present the variance and cumulative variance explained by each latent factor. Column 5 shows the Cronbach's Alpha for the first four latent factors.

Table A8
Country-Level Variable and Factor Solution

No.	Name	Extraction	Factor	Loading	No.	Name	Extraction	Factor	Loading
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1	Rule of law	0.970	1	0.962	33	Securities regulation: liability standards	0.797	2	0.729
2	Control of corruption	0.956	1	0.956	34	Public enforcement: securities regulation	0.866	2	0.726
3	Regulatory quality	0.921	1	0.926	35	Securities regulation: disclosure requirements	0.798	2	0.620
4	Low Political Risk	0.929	1	0.925	36	Number of analysts	0.777	2	0.464
5	Media	0.908	1	0.914	37	Ex post private control of self-dealing	0.816	2	0.446
6	Low repudiation of contracts by govt.	0.956	1	0.906	38	Ex ante private control of self-dealing	0.874	3	0.881
7	Law and order	0.909	1	0.906	39	Private control of self-dealing index	0.909	3	0.823
8	Corruption	0.884	1	0.888	40	Legal origin	0.743	3	0.735
9	Property rights	0.895	1	0.874	41	Uncertainty avoidance	0.758	3	-0.467
10	Judicial independence	0.885	1	0.853	42	Market Cap. to GDP	0.825	4	0.851
11	GDP per capita	0.855	1	0.850	43	Foreign investment to GDP	0.856	4	0.840
12	Judicial efficiency	0.862	1	0.836	44	Listed firms to population	0.843	4	0.769
13	Economic freedom	0.869	1	0.826	45	US institutional holdings	0.933	5	0.922
14	Information and knowledge	0.876	1	0.822	46	US cross-listing	0.892	5	0.868
15	Political stability	0.883	1	0.812	47	English proficiency	0.942	6	0.766
16	Political score	0.937	1	0.812	48	Language proximity to English	0.879	6	-0.704
17	Secrecy	0.930	1	-0.806	49	Buddhist	0.754	6	-0.675
18	Power distance	0.864	1	-0.789	50	Language fractionalization	0.904	7	0.865
19	Assessment of tax evasion	0.837	1	0.784	51	Ethnic fractionalization	0.808	7	0.767
20	Trust	0.912	1	0.762	52	Catholic	0.904	8	-0.729
21	Enforcement of audit standards	0.840	1	0.713	53	Other religion	0.891	8	0.528
22	Individualism	0.889	1	0.692	54	Long-term orientation	0.862	9	0.789
23	Creditor rights	0.874	1	0.650	55	Religion fractionalization	0.813	9	-0.517

24	Bank money in the private sector to GDP	0.860	1	0.628	56	Religiousness	0.802	10	0.503
25	Enforcement of accounting standards	0.782	1	0.609	57	Individualism in income	0.617	10	−0.487
26	Latitude	0.841	1	0.583	58	Masculinity	0.862	11	0.895
27	Ownership concentration	0.849	1	−0.555	59	Protestant	0.873	11	−0.644
28	Muslim	0.794	1	−0.508	60	Politically connected firms	0.793	12	0.715
29	Big4 market share	0.744	1	0.497	61	Public control of self-dealing	0.771	13	0.780
30	IPOs to GDP	0.735	1	0.471	62	Democracy	0.682	13	−0.490
31	Strength of securities regulation	0.953	2	0.877	63	Block premium	0.769	13	0.433
32	Class action lawsuit	0.800	2	0.740	64	Anti-director rights	0.752	14	0.680

Notes: This table reports the results for principal component analysis for 64 country-level variables for 46 developing and developed countries. Columns 2 & 7 show the 64 variable names. Columns 1 & 6 present the variable sequence with their primary factor membership reflected in Columns 4 & 9. The item-wise (variables) communalities are shown in columns 3 & 8, and factor loadings extracted from the rotated component matrix are shown in columns 5 & 10.

Table A9
Descriptive Statistics of Entropy Balanced Sample

Variables	Treat Group			Control Group		
	Mean	Variance	Skewness	Mean	Variance	Skewness
(1)	(2)	(3)	(4)	(5)	(6)	(7)
Before Entropy Balancing						
SIZE	21.840	2.398	−0.185	21.740	3.254	−0.179
LEV	0.240	0.030	0.610	0.251	0.034	0.801
RDEXP	0.011	0.001	5.146	0.027	0.005	3.963
CAPEX	0.055	0.002	1.558	0.048	0.002	2.204
ROA	0.060	0.007	−1.463	0.015	0.023	−3.069
GROWTH	0.075	0.069	2.249	0.084	0.081	2.515
LIQ	0.157	0.038	0.433	0.178	0.043	0.926
DIV	0.867	0.115	−2.166	0.668	0.222	−0.714
After Entropy Balancing						
SIZE	21.840	2.398	−0.185	21.840	2.398	−0.185
LEV	0.240	0.030	0.610	0.240	0.030	0.610
RDEXP	0.011	0.001	5.146	0.011	0.001	5.157
CAPEX	0.055	0.002	1.558	0.055	0.002	1.558
ROA	0.060	0.007	−1.463	0.060	0.007	−1.465
GROWTH	0.075	0.069	2.249	0.075	0.069	2.249
LIQ	0.157	0.038	0.433	0.157	0.038	0.433
DIV	0.867	0.115	−2.166	0.867	0.115	−2.166

Notes: This table reports the summary statistics before and after entropy-balanced matching. All continuous variables have been winsorized at the 1st and 99th percentiles. The sample comprises annual data of firms in 46 developed and developing countries from 2009 to 2023. All variables are defined in Table 2.