A Legal Origins Perspective on ESG Ratings Disagreement

Abstract

Environmental, social, and governance (ESG) scores of firms vary significantly across different rating providers. This article considers the Legal origins theory as a potential factor influencing ESG disagreement. By comparing ESG scores from five reputable ratings providers for a sample of 2,392 public firms across 53 countries, we find that "dispersion disagreement" among rating providers is lower for firms incorporated in common law countries, while "correlation disagreement" is lower for firms incorporated in common law countries. This suggests, firstly, that firms in common law countries engage in more independent and firm-specific ESG efforts, resulting in lower ESG dispersion, and secondly, that firms in civil law countries are influenced more by common factors such as national regulations and industry practices, leading to higher correlation in ESG scores between rating providers.

Key words: ESG, ESG disagreement, dispersion, correlation, legal origin

1. Introduction

Environmental, Social, and Governance (ESG) scores, provided by independent rating agencies, have become a key measure of a firm's sustainability performance. Rating agencies provide ESG scores to financial market participants who progressively integrate them into their investment strategies and decision making (Pedersen et al., 2021). However, even trusted providers using robust methodologies often assign different ESG scores to the same firm. Inconsistent ESG scores undermine their usefulness in reducing sustainability-related information asymmetries. In support of the so-called ESG ratings disagreement, several studies observe only low to moderately positive correlations of ESG scores between different rating providers (Avramov et al., 2022; Berg et al., 2022; Chatterji et al., 2016; Dorfleitner et al., 2015).

Regulatory frameworks should in principle enhance ESG ratings transparency by improving the quality and/or quantity of disclosed information, thereby reducing informational asymmetries. Building on this rationale with the legal origins theory (La Porta et al., 2008), we would expect ESG ratings of firms in civil law countries, which adhere to codification of sustainability-related regulatory frameworks (with the EU being the primary example), to exhibit lower levels of disagreement between different ratings providers compared to firms in common law countries. Whereas the *level* of ESG ratings has already been linked to countries' legal origins (Liang & Renneboog, 2017), the nexus between the *disagreement* of ESG ratings and legal origins has not yet been explored in literature to the best of our knowledge.

In this paper, we examine the ESG ratings disagreement through the lens of legal origins of countries where the firms are incorporated. Specifically, we explore two dimensions of this disagreement: *correlation disagreement*, which refers to the *sample-level* correlations of ESG scores between different providers, and *dispersion disagreement*, which refers to the *firm-level* variability in these scores, focusing on the differences between civil and common law countries. For the purpose of our analysis, we construct a cross-sectional database of current (2024) ESG scores for 2,392 firms from 53 countries across five reputable ESG ratings providers (Bloomberg, MSCI, S&P, Sustainalytics, and LSEG) and combine it with firm-specific (LSEG) and country-specific (World Bank Open Data) variables, including the legal origins variable, obtained from La Porta et al. (2008).

We find supporting evidence linking each dimension to the legal origins, and, in the case of the dispersion disagreement, to firm-, industry- and country-specific characteristics. Our findings only partially support our initial expectations. On the one hand, we find (as expected) higher correlations of ESG scores between ratings providers in civil law countries in comparison with common law countries. On the other hand, civil law countries exhibit higher dispersion of ESG scores relative to common law countries. The implication could be that increased regulation of sustainability in civil law countries (such as demands for mandatory ESG disclosures) may, on the one hand, help ESG ratings providers better agree on the relative ESG ranking of *different* firms. However, at the same time, it might encourage them to exploit legislative ambiguities in sustainability standards and metrics, leading to further disagreement on the exact ESG score of the *same* firm, when compared to firms in common law countries.

In support of this explanation, Christensen et al. (2022) showed that the increase in ESG disclosures between 2004 and 2016 has actually led to a greater dispersion (measured as standard deviation) of ESG ratings. As underscored by Berg et al., (2022), Billio et al. (2021) and Chatterji et al. (2016), the ESG ratings disagreement has been driven by a lack of a common scope, attributes and standards. Insofar as the regulatory approaches of civil law countries do not touch upon these issues, they may in fact accentuate firm-level variability of ESG ratings. There is, however, a silver lining to our findings: regulatory frameworks are changing, and have recently shifted (particularly in the EU) towards the development of common sustainability reporting standards and increased transparency of ESG ratings activity. This could alter the current state of ESG ratings disagreement between civil and common law countries, making this research area also an intriguing field for future exploration.

Our paper contributes to the existing literature in several ways. Firstly, it connects the legal origins theory with the ESG ratings disagreement literature, empirically demonstrating the effect of different legal environments on ESG disagreement. This intersection is critical for policymakers aiming to design effective sustainability policies that promote transparency in reporting, assessing, and interpreting ESG performance. Enhancing the credibility of ESG scores is crucial not only for researchers studying ESG-related topics but also for firms and investors. Secondly, it distinguishes between *correlation disagreement* and *dispersion disagreement* as two separate dimensions of the ESG ratings disagreement. This distinction is particularly relevant in the context of legal origins when considering different factors beyond methodological differences that may influence ESG ratings. We show that each of these

dimensions behave differently, highlighting the need for a nuanced interpretation of the effects of legal origins on the ESG ratings disagreement. Finally, this paper seeks to improve understanding whether regulatory approaches to sustainability add value by reducing asymmetry of ESG information. Our findings indicate that industries that are directly related to sustainability regulations and issues, such as energy, exhibit substantially lower ESG information asymmetry. Understanding the relationship between legal origins and ESG disagreement among rating providers has implications for investors and policymakers. It could help investors better assess ESG risks and opportunities across different jurisdictions, while policymakers may use this knowledge to strengthen regulatory frameworks and promote global ESG harmonization efforts.

The rest of the paper is organised as follows: Section 2 presents the literature review; Section 3 describes the data and methodology; Section 4 presents key results with discussion; and Section 5 concludes.

2. Literature review

Our paper builds on the studies by Chatterji et al. (2016) and Liang & Renneboog (2017). Chatterji et al. (2016) found substantial disagreement in SRI ratings among six providers (KLD, Asset4, Calvert, FTSE4Good, DJSI, and Innovest), even after controlling for different definitions of CSR among providers, which they attribute to differences in "theorization". Furthermore, they demonstrated that the consistency of these ratings did not improve between 2003 and 2006, based on samples from KLD and Asset4. However, they suggested that rating providers based in the same country or geographical region tend to have a higher (tetrachoric) correlation in SRI ratings than the average correlation among the six providers. This leads to two conclusions. First, rating providers located in the same country or geographical region may share a common belief about *what* it means to be socially responsible, leading them to measure the same attributes. Second, rating providers may also agree on how to measure these overlapping attributes, a concept they refer to as "commensurability".

Liang & Renneboog (2017) demonstrated the significance of the legal origins theory in evaluating the level of firm's sustainability performance. They show a strong correlation between firms' level of CSR ratings and legal origins, indicating that firms in civil law countries have higher CSR ratings than those in common law countries, with highest scores are observed for firms from Scandinavian civil law. They argue that "legal origin appears to be the strongest predictor of CSR adoption and performance at the firm level, stronger than

alternative factors such as political institutions, regulations, social preferences, and a firm's financial and operational performance".

This paper addresses the intersection gap between these two bodies of literature by integrating the aspects considered in both studies. For example, Liang & Renneboog (2017) applied the legal origin theory to compare the sustainability performance of firms in civil law countries with those in common law countries. They used CSR scores from a single provider (MSCI IVA) and did not consider the *correlation disagreement* nor *dispersion disagreement*. Chatterji et al. (2016), on the other hand, focused on methodological differences and examined the convergence of ratings using tetrachoric correlations between providers. They indicated that the location of a rating provider affects ESG rating disagreement in the context of *correlation disagreement*, however, they did not account for legal origins, nor did they consider *dispersion disagreement*. This paper, however, shifts the focus to the location of firms, prioritizing the reporting aspect over methodological considerations when analysing ESG rating disagreement.

More recent papers investigating ESG rating convergence (Chatterji et al., 2016; Dorfleitner et al., 2015; Widyawati, 2021) or divergence (Berg et al., 2022) aim to explain methodological differences, thereby focusing on *correlation disagreement*. Those papers have identified (Chatterji et al., 2016) and quantified (Berg et al., 2022) the sources of *correlation disagreement.* Berg et al. (2022) developed a unified taxonomy of categories integrating methodologies from six providers (KLD, Sustainalytics, Moody's ESG, S&P Global, Refinitiv, and MSCI). Using a sample of 924 firms' ESG scores from 2014, they quantified three distinct sources of ESG rating divergence: measurement divergence accounts for 56% of the total divergence, scope divergence for 38%, and weight divergence for the remaining 6%. Earlier work (Chatterji et al., 2009; Chatterji & Toffel, 2010) examined methodology of KLD (now part of MSCI), one of the earliest specialized providers of ESG information and primary dataset used in most existing academic research on ESG to date. According to Berg et al. (2022), divergence (i.e., *correlation disagreement*) in relation to other providers is particularly pronounced for KLD, which underscores the challenges of relying on a single rating provider for comprehensive ESG analysis and highlights the need for a more nuanced approach when interpreting firms' ESG performance across various rating providers.

This paper delves into the underlying factors beyond these methodological differences by integrating the role of legal origins (La Porta et al., 2008) as a significant factor influencing the ESG rating disagreement across rating providers. It is reasonable to expect that legal origins might play a role in examining ESG ratings inconsistency, as firms in countries with different legal origins are subject to inherent differences in their regulatory environments and enforcement mechanisms. According to the legal origins theory, countries' legal systems adhere either to *civil law* or *common law* principles (with civil law further subdivided into French, German, Socialist and Scandinavian civil law). Civil law relies heavily on *ex-ante* legislation while common law provides broader frameworks with *ex-post* setting up mechanisms.[1] In another words, civil law is *policy implementing*, while common law is *dispute resolving* (Damaska, 1991). La Porta et al., (2008) demonstrate that legal origins significantly influence legislation and regulation, which in turn have a profound impact on various economic outcomes. These outcomes include financial development, international trade, unemployment rates, industry growth, resource allocation, investment, size of unofficial economy, productivity and growth rate of GDP per capita (La Porta et al., 2008; Mahoney, 2001).

By analysing ESG rating disagreement from the perspective of legal origin, we can identify the distinct effects of *correlation disagreement* and *dispersion disagreements*—critical aspects that prior studies did not adequately clarify. Studies examining the differences on the supply side of the ESG rating market (Berg et al., 2022; Chatterji et al., 2016; Dorfleitner et al., 2015; Widyawati, 2021), namely ESG rating convergence/divergence among rating providers, focus on *correlation disagreement*, sometimes referred to as *ESG risk* (Dorfleitner et al., 2015), while studies on in the demand side of the ESG rating market (Avramov et al., 2022; Christensen et al., 2022; Gibson Brandon et al., 2021), namely the variability in firms' ESG performance, focus on *dispersion disagreement*, sometimes referred to as *ESG uncertainty* (Avramov et al., 2022).

Dorfleitner et al. (2015) show that distribution of a ESG ratings from three providers (ASSET 4, KLD and Bloomberg) with sample of 8,500 firms do not coincide, but they found moderate positive correlation between Bloomberg and ASSET 4. They were the first to examine year-over-year change in the level of firms' ESG performance among rating providers and called it

[1] EU as a whole follows civil law principles.

ESG risk[2]. They argue that a negative year-over-year change in ESG scores is more significant for investors in terms of *ESG risk*. To assess this, they examined *expected loss* (Fishburn, 1977), over three consecutive years. Their analysis employed the concept of the lower partial moment of order one, introduced by Bawa, (1975) and Jean (1975). They concluded that *expected loss* varies among rating providers, with no correlation in *ESG risk* among the three rating providers. However, they did not estimate the dispersion of ESG scores and therefore only addressed *correlation disagreement*.

Studies on the demand side of the ESG rating market reveal that greater *dispersion disagreement* is associated with higher stock returns, higher return volatility, larger absolute price movements and a reduced probability of external financing (Avramov et al., 2022; Christensen et al., 2022; Gibson Brandon et al., 2021), suggesting that ESG rating disagreement in the context of *dispersion disagreement* affects social impact and economic welfare (Avramov et al., 2022).

3. Methodology

3.1. Correlation disagreement

We measure *correlation disagreement* using pairwise Pearson correlation coefficients of ESG scores across the five ESG ratings providers for two subsamples, one for civil law countries and another one for common law countries. Correlation coefficients help to understand how aligned ESG scores are between pairs of different providers within each subsample of firms. Specifically, they indicate to what extent the providers agree on the relative ESG ranking of firms in civil and common law countries, with higher correlation coefficients indicating lower levels of disagreement (and, conversely, lower coefficients indicating higher levels of disagreement).

The estimated coefficients for the same pairs of ESG ratings providers are then compared between the civil law and common law subsamples, testing for statistical significance of the differences between them using the Fisher's Z-transformation statistics. Specifically, if *a* and *b* are two ratings providers, transformation of correlation coefficients from the civil law subsample (r_{ab}^{civ}) and the common law subsample (r_{ab}^{com}) can be written as:

[2] Dorfleitner et al. (2015) define *ESG risk* as "the risk of changes in the average ESG level of an investors' portfolio" over time.

$$Z_{ab}^{civ} = \frac{1}{2}\ln\left(\frac{1+r_{ab}^{civ}}{1-r_{ab}^{civ}}\right) \text{ and } Z_{ab}^{com} = \frac{1}{2}\ln\left(\frac{1+r_{ab}^{com}}{1-r_{ab}^{com}}\right)$$

(1)

The Z-score for the difference between two correlations is given by:

$$Z_{\text{diff},ab} = \frac{Z_{ab}^{\text{civ}} - Z_{ab}^{\text{com}}}{SE}, \text{ where } SE = \sqrt{\frac{1}{n^{civ}-3} + \frac{1}{n^{com}-3}}$$

(2)

If the differences are found to be significant, this is an indication of different levels of *correlation disagreement* in civil vs. common law countries.

3.2. Dispersion disagreement

Drawing from the approaches used by Liang & Renneboog (2017) and Christensen et al. (2022), we employ an ordinary least squares (OLS) regression model to examine the link between the *dispersion disagreement* and the legal origins of the country where the firm operates.

Specifically, we estimate the following two equations:

$$ESGdisp_{ij} = \beta_0 + \beta_1 LegOr_j + \sum_l \beta_l CC_{l,j} + \sum_k \beta_k FC_{k,ij} + \varepsilon_{ij}$$

(3)

$$ESGdisp_{ij} = \beta_0 + \beta_1 LegOr_j + \sum_l \beta_l CC_{l,j} + \sum_l \beta_l LegOr_j \times CC_{l,j} + \sum_k \beta_k FC_{k,ij} + \varepsilon_{ij}$$

(4)

Where $ESGdisp_{ij}$ represents the firm-level dispersion of ESG scores for firm *i* in country *j*. $LegOr_j$ is the legal origin dummy variable for country *j*, representing civil law (0) and common law (1) countries. $CC_{l,j}$ denotes *l* country-level control variables for country *j*. The following country controls are considered: GDP per capita ($GDPpc$), the Government climate risk score ($CliRScr$), the Freedom House score ($FHScr$), the Regulatory quality index ($RegQ$), and the Control of corruption index ($CCInd$). $FC_{k,ij}$ represents *k* firm-level control variables for firm *i* in country *j*. The following firm controls are considered: firm size ($logTA$), profitability (ROE), industry sector ($GCIS$) and ESG controversies score ($ESGCrvScr$). ε_{ij} is the random error term. The key difference between Eqs. (1) and (2) is in

the inclusion of interaction terms between our main variable of interest, $LegOr$, and the country control variables ($LegOr \times CC$). By adding interaction terms, we test whether the effect of legal origin operates through a specific country-level channel. We include control variables in the regression incrementally. Firstly, we include country level variables, secondly, we include firm level variables, and lastly, we include sector dummy variables.

Two measures of dispersion of ESG scores, an absolute and a relative one, are considered as our dependent variable, $ESGdisp$. We use standard deviation (SD) of ESG scores across the five ESG ratings providers for each firm as our *absolute* measure of firm-specific dispersion of ESG scores. Conversely, we employ coefficient of variation (CV) across the five ESG ratings providers for each firm as our *relative* measure of firm-specific dispersion of ESG scores. In particular, our firm-specific dependent variable for a particular firm *i* can be written as:

$$ESGDisp_i = \begin{cases} SD_i = \sqrt{\frac{1}{5-1} - \sum_1^5 (ESG_i - \overline{ESG_i})} \\ CV_i = \frac{SD_i}{\overline{ESG_i}} \end{cases} \tag{5}$$

Although previous studies typically used one of these measures (Avramov et al., 2009, 2022; Christensen et al., 2022), our analysis considers both to ensure a richer and more detailed assessment of the factors influencing ESG score dispersion. The main advantage of using the coefficient of variation is that it is not affected by the units of measurement of the data. Including both measures serves as a robustness check of our results and ensures that the conclusions drawn are not driven by the choice of dispersion measure.

4. Data

The data for this study is sourced from Bloomberg, LSEG Workspace, and World Bank Open Data. Information regarding the legal origin of countries is obtained from La Porta et al. (2008). The analysis encompasses a cross-sectional sample of 2,392 public firms from 53 countries. Of these, 1,084 firms (45.32%) are incorporated in countries with common law, while 1,308 firms (54.68%) are incorporated in countries with civil law. Within the civil law category, 929 firms (38.84%) are from countries with German civil law, 306 firms (12.79%) from countries with French civil law, and 73 firms (3.05%) from countries with Scandinavian

civil law. We use the most recently available data (2024 in the case of ESG, firm-specific and legal origins data and 2022 in the case of other country-specific data).

4.1. ESG scores

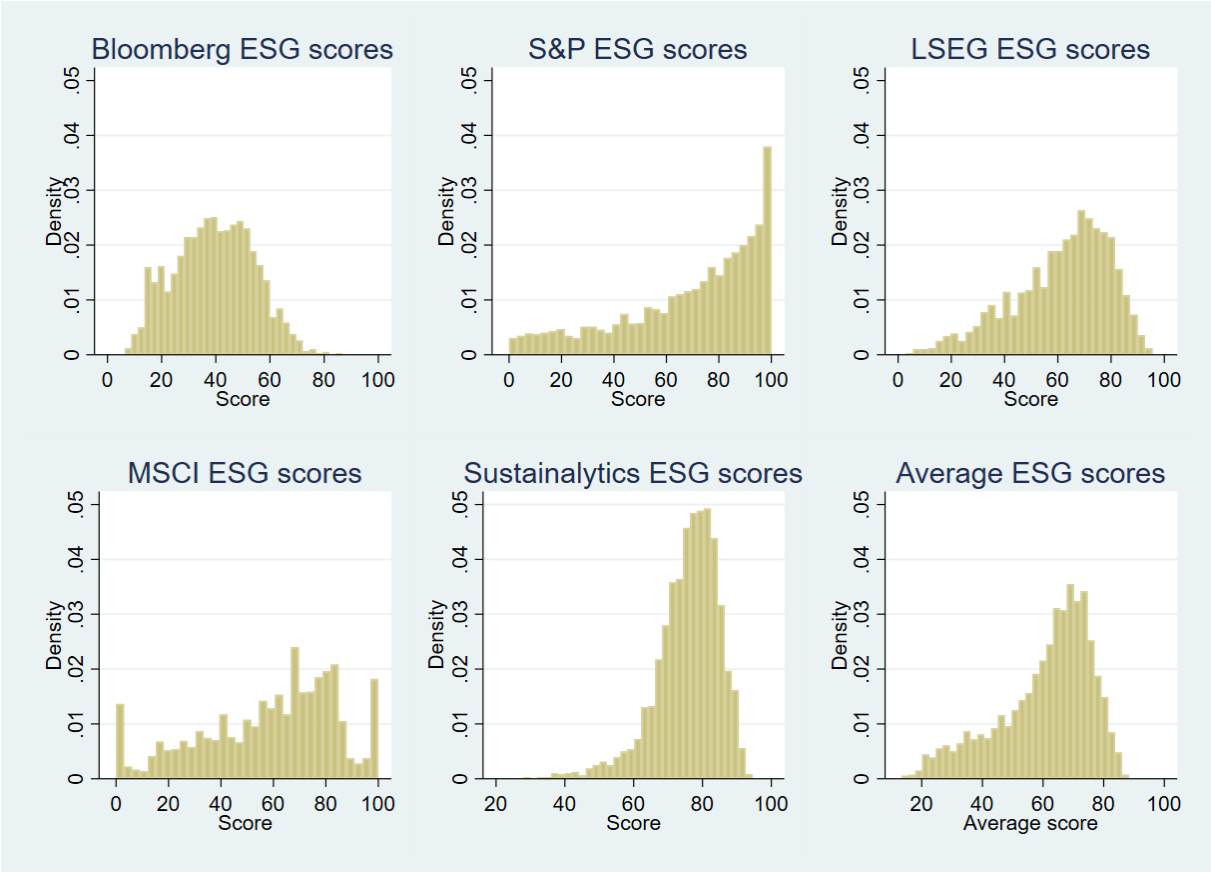
To compile our final sample of ESG scores, we utilized two sources: LSEG Workspace and Bloomberg Terminal. The Bloomberg Terminal aggregates data from third-party providers, including MSCI, S&P Global, and Sustainalytics. Since ESG ratings providers vary in their coverage of firms, our sample is limited to firms that have been rated by all five providers. We obtained the following number of firm-specific ESG scores, published in April 2024: 14,475 scores from Bloomberg, 10,197 scores from S&P Global, 4,131 scores from Sustainalytics, 2,672 scores from MSCI, and 11,776 scores from LSEG Workspace. After merging and matching these scores, our final sample consists of 2,392 firms, each scored by all five ESG ratings providers.

The next thing to consider when comparing ESG scores across different providers are different scoring scales. The Bloomberg ESG score is a weighted generalized mean of E, S, G pillar scores, where the weights are determined by the pillar priority ranking. The scores range from 0 to 10, with 10 being the best score. The MSCI ESG score is calculated by normalizing the weighted average key issue score to the industry peer set, adjusted to reflect any Ratings Review Committee overrides. The scores range from 0 to10, with 10 being the best score. The S&P Global ESG score is based on individual questions that roll up into criteria. The types and weights of individual questions and criteria are adjusted for each industry-specific questionnaire. The score ranges from 0 to 100, with 100 being the best score. Sustainalytics ESG risk score applies the concept of risk decomposition, which is assigned to one of five risk categories. The score ranges from 0 and 100, with 0 indicating that ESG risks have been fully managed, and 100 indicating the highest level of unmanaged ESG risk. Importantly, the sentiment of this score is reverse in comparison to the scores of the other providers, with higher values associated with a negative ESG sentiment. It is also unique in the sense that it specifically focuses on ESG risks and does not consider ESG opportunities and impacts. It is calculated as the difference between a company's overall exposure score and its overall managed risk score, or alternatively by adding the Corporate Governance unmanaged risk score to the sum of the company's issue unmanaged risk scores. LSEG ESG score is a relative sum of the category weights, which vary per industry for the environmental and social categories, for governance the weights remain the same across all industries. The pillar

weights are normalised to percentages ranging between 0 and 100. The scores range from 0 to 100, with 100 being is the best score.

The Bloomberg, MSCI and Sustainalytics ESG scores were adjusted to a common scale ranging from 0 to 100, with 100 being the best. The adjustment to the common scale was done by conversion, where each original score was multiplied by a factor of 10 for Bloomberg and MSCI scores. For Sustainalytics, the scores were inverted due to their reverse sentiment. Figure 1 illustrates the distribution of adjusted ESG scores from the five ESG ratings providers alongside the average ESG score. Each distribution shows variability, with scores generally clustering in the middle to upper ranges, reflecting a moderate to high ESG performance across the sample of 2,392 public firms. Except for distribution of the Bloomberg ESG scores, the scores of all other ratings providers (as well as the average scores) are skewed to the left (negative skew).

Figure 1. Distribution of ESG scores and average firm scores across providers for a sample of 2,392 public firms.



Note: The figure presents the density distributions of ESG scores from five different providers as well as the average ESG score across these providers, on common scale from 0 to 100, 100 is best. The figure is organized into two rows, each containing three plots. The top row displays the distributions

for Bloomberg, S&P Global, and LSEG. The bottom row shows the distributions for MSCI, Sustainalytics, and the average ESG scores. Each subplot illustrates the distribution of ESG scores for a sample of 2,392 firms, highlighting the variations and concentration of scores assigned by each provider.

4.2. Country-level and firm-level variables

We use country-level and firm-level control variables in our analysis. Country-level data include GDP per capita (measured in thousands of USD), the Government climate risk score, the Freedom House score, the Regulatory quality index, and the Control of corruption index. All country-level data, except for the Government climate risk score (sourced from Bloomberg in April 2024), were obtained from World Bank Open Data and pertain to the year 2022.

The Government climate risk score measures the progress of a country in meeting the global goals set out in the Paris Agreement and ranges from 0 to 10, with 10 being the best. We control for the Government climate risk score because countries with better climate risk management are likely to foster environments where companies are more attuned to ESG considerations, potentially affecting the dispersion of ESG scores. By accounting for this variable, we can better understand how legal origin influences the broader impact of environmental policy and performance through government climate risk scores, especially for firms in civil law countries.

The Freedom House score evaluates the level of freedom and democratic governance in a country. This score ranges from 0 to 100, with 100 being the highest level of freedom. We control for the Freedom House score to account for the broader political and governance context in which firms are incorporated. A higher level of freedom and democratic governance often correlates with stronger institutions, transparency, and regulatory quality. Controlling for this variable ensures that our analysis of legal origin's effect on ESG score dispersion is not confounded by varying levels of political freedom and governance quality across countries.

The Regulatory quality and Control of corruption indices are part of the Worldwide Governance Indicators (WGI). The Regulatory quality index measures the perceptions of how well the government formulates and enforces policies and regulations that enable and foster private sector development. The Control of corruption index measures perceptions regarding the extent to which public power is used for personal gain. This includes both minor and major forms of corruption, as well as the domination of the state by elites and private

interests. Both indices range from -2.5 to 2.5, with higher values indicating better governance outcomes. Controlling for regulatory quality and corruption is crucial to accurately assess the impact of legal origin on the dispersion of ESG ratings. It ensures that the analysis considers key governance factors that can significantly influence ESG practices and reporting consistency among firms.

Consistent with the approaches of previous studies, we consider firm-level financial data that may be associated with firm ESG efforts (e.g. Avramov et al., 2022; Berg et al., 2021; Christensen et al., 2022; Drempetic et al., 2020; Dumrose et al., 2022; Pedersen et al., 2021). Firm-level financial data was sourced from LSEG Workspace in April 2024 for the latest available fiscal year. We use firm characteristics including firm size, measured by log of total assets (USD), profitability from the shareholders' perspective, measured by return on equity (ROE), and sector membership according to the Global Industry Classification Standard (GICS). In addition to financial variables, we also include firm-level ESG controversies score provided by LSEG. ESG controversies score measures a firm's exposure to environmental, social and governance scandals and negative events reported in the global media. It ranges from 0 to 100, with 100 being the best. We control sector membership using sector-specific dummies, with the Energy sector representing the baseline.

5. Results and discussion

Our results are reported in two parts. In the first part, we analyse the *correlation disagreement*, using pairwise Pearson correlation coefficients and comparing them between common law and civil law firms. In the second part, we present the results of the regression analysis, focusing on the *dispersion disagreement*.

5.1. Correlation analysis and *correlation disagreement*

The pairwise Pearson correlation matrix of five ESG score providers for our full sample, considering all firms, regardless of the legal origins of the countries where they operate, is presented in Table 1. The majority of correlations fall within the moderate range, with an average correlation coefficient of 0.5307. For example, the correlation between S&P and LSEG (0.7381) is quite strong, indicating a low "correlation disagreement" between S&P and LSEG. Bloomberg's ESG scores show moderate correlations with the other providers, with the highest correlation being with LSEG's ESG scores (0.6255) and the lowest with Sustainalytics's ESG scores (0.2665). Sustainalytics's has the lowest correlations with all of the other providers, but particularly with Bloomberg (0.2665) and LSEG (0.3587). This might

be related to the distinct nature of Sustainalytics's ESG ratings, focusing primarily on ESG risks, and indicates that Sustainalytics employs a distinct methodology and criteria in their ESG risk scoring, leading to the highest "correlation disagreement" with the other ratings providers. It furthermore underscores the need to closely scrutinize the scope of ratings, as already mentioned by many authors.

Table 1. Correlation matrix of ESG scores.

	Bloomberg	S&P	LSEG	MSCI	Sustainalytics
Bloomberg	1				
S&P	0.6012***	1			
LSEG	0.6255***	0.7381***	1		
MSCI	0.5185***	0.6476***	0.5752***	1	
Sustainalytics	0.2665***	0.4837***	0.3587***	0.4920***	1

Note: The table presents the pairwise Pearson correlation matrix of ESG scores between the five providers, namely Bloomberg, S&P Global, LSEG, MSCI, and Sustainalytics, based on the sample of 2,392 public firms. Statistical significance is indicated with *** at 0.00001.

Next, we considered correlations of ESG scores between providers separately for firms incorporated in common law countries (Table 2) and civil law countries (Table 3). The analysis shows that correlation coefficients are substantially higher, and, hence, "correlation disagreement" on ESG performance lower for firms incorporated in civil law countries in comparison to firms from common law countries. Specifically, the average correlation of ESG scores equals 0.5788 for firms in civil law countries, whereas for firms in common law countries it stands at 0.3891.

Table 2. Correlation matrix of ESG scores for common law.

	Bloomberg	S&P	LSEG	MSCI	Sustainalytics
Bloomberg	1				
S&P	0.5121***	1			
LSEG	0.5805***	0.6191***	1		
MSCI	0.4246***	0.4491***	0.4296***	1	
Sustainalytics	0.1119***	0.3178***	0.1798***	0.2665***	1

Note: The table presents the pairwise Pearson correlation matrix of ESG scores between the five providers based on the sample of 1,084 public firms from common law countries. Statistical significance is indicated with *** at 0.00001.

Table 3. Correlation matrix of ESG scores for civil law.

	Bloomberg	S&P	LSEG	MSCI	Sustainalytics
Bloomberg	1				
S&P	0.6453***	1			
LSEG	0.6565***	0.7895***	1		
MSCI	0.5391***	0.7070***	0.6363***	1	
Sustainalytics	0.3118***	0.5283***	0.4265***	0.5474***	1

Note: The table presents the pairwise Pearson correlation matrix of ESG scores between the five providers based on the sample of 1,308 public firms from civil law countries. Statistical significance is indicated with *** at 0.00001.

We specifically tested the significance of differences in correlation coefficients of the same pairs of ESG ratings providers between civil and common law subgroups using the Fisher's Z-transformation statistics. The differences are found to be statistically significant for all pairs, leading us to conclude that there is clear evidence of lower "correlation disagreement" in civil law countries in comparison to common law countries. This lower "correlation

disagreement" could be attributed to the more centralized and standardized regulatory environments typically found in civil law jurisdictions. The results indicate the importance of considering legal origins when examining ESG scores.

5.2. Regression analysis and *dispersion disagreement*

Table 4 presents the results of the regression analysis where ESG dispersion is measured in absolute terms (SD). Conversely, Table 5, shows the results of the regression analysis where ESG dispersion is measured in relative terms (CV). In both tables, Model I presents the simple bivariate OLS regression with Legal origin as the only explanatory variable. Model II extends Model I by including country-level controls, whereas models IV and VI additionally also include firm-level control variables and industry dummy variables. Models III, V, and VII furthermore include interaction terms between the legal origin dummy variable and the country-level control variables ($CliRScr$, $FHScr$, $RegQ$, and $CCInd$). Following the recommendation of Balli & Sørensen (2013), the variables included in the interaction terms in Models III, V, VII, as well as further models VIII-X, with the exception of the legal origin dummy, have been demeaned (cantered) to ensure clarity in interpreting the coefficients.

Table 4: Regression analysis of factors affecting the standard deviation of ESG scores across models I-VII.

Model	I	II	III	IV	V	VI	VII
LegOr	-2.316***	-2.054***	-2.407	-1.842***	-2.123	-1.628***	-1.302
	(0.2092)	(0.3293)	(1.2908)	(0.3522)	(1.3882)	(0.3323)	(1.2989)
CliRScr		0.098	-0.388	0.026	-0.458*	0.119	-0.336
		(0.1355)	(0.2069)	(0.1400)	(0.2158)	(0.1319)	(0.2011)
FHScr		-0.036***	-0.005	-0.029***	0.006	-0.023***	0.011
		(0.0063)	(0.0170)	(0.0065)	(0.0183)	(0.0061)	(0.0171)
RQInd		1.617**	-0.145	1.422*	-0.630	0.850	-0.933
		(0.5588)	(1.1427)	(0.5981)	(1.2805)	(0.5657)	(1.1929)
CCInd		-1.468***	-0.545	-1.430**	-0.350	-1.248**	-0.505
		(0.4011)	(0.6456)	(0.4330)	0.7361	(0.4079)	(0.6872)
GDPpc		-0.012	0.002	-0.009	0.003	-0.0109	0.002
		(0.0066)	(0.0080)	(0.0067)	(0.0082)	(0.0062)	(0.0076)
LegOr x CliRScr			0.926**		0.908*		0.784*
			(0.3474)		(0.3601)		(0.3361)
LegOr x FHScr			-0.035		-0.039*		-0.039*
			(0.0181)		(0.0194)		(0.0181)
LegOr x RQInd			-0.106		0.661		-0.587
			(1.5863)		(1.7188)		(1.6079)
LegOr x CCInd			-0.634		-0.928		0.381
			(1.1483)		(1.2219)		(1.1443)
ESGCovScr				0.013**	0.011**	0.008	0.006
				(0.0040)	(0.0040)	(0.0039)	(0.0039)
logTA				-0.227**	-0.230**	-0.338***	-0.344***
				(0.0813)	(0.0815)	(0.0932)	(0.0931)
ROE				-0.313*	-0.310*	-0.315**	-0.315**
				(0.1299)	(0.1297)	(0.1210)	(0.1208)
Commun. services						6.896***	6.796***
						(0.6747)	(0.6742)
Consumer discretionary						5.774***	5.670***
						(0.6103)	(0.61025)
Consumer staples						3.101***	3.029***
						(0.6269)	(0.6259)
Financials						6.134***	6.067***
						(0.6087)	(0.6076)
Health care						4.715***	4.736***
						(0.6225)	(0.6226)
Industrials						3.997***	3.991***
						(0.5788)	(0.5778)
Information technology						6.159***	6.185***
						(0.6099)	(0.6097)
Materials						1.735**	1.685**
						(0.6110)	(0.6193)
Real estate						6.074***	6.126***
						(0.7126)	(0.7118)
Utilities						1.837**	1.763*
						(0.6806)	(0.6811)
Intercept	19.984***	22.086***	21.796***	26.133***	25.765***	24.284***	23.870***
	(0.1408)	(0.5374)	(0.7967)	(2.1261)	(2.2366)	(2.5162)	(2.5841)
Observations	2,392	2,275	2,275	2,005	2,005	2,005	2,005
R-squared	0.0488	0.0989	0.1061	0.1173	0.1232	0.2409	0.2460

Note: The table reports results from regression of ESG dispersion, measured as standard deviation, on explanatory variables. Model I includes only the Legal origin as the explanatory variable, model II adds country controls to model I, model IV adds firm controls model II, model VI includes eleven sectors to model IV with Energy sector as benchmark. Model III adds three interaction terms to model II, model IV adds firm controls model II, model V three interaction terms to model IV, model VI includes eleven sectors to model IV with Energy sector as benchmark, Model VII, V and VII adds three interaction terms to model II, IV and VI. *, **, and *** indicate statistical significance at the 0.05, 0.01, and 0.001.

A first look at the results of the simple bivariate OLS regression with *standard deviation as the dependent variable* (Table 4, Model I) indicates that firms in common law countries have on average 2.32 points lower standard deviation of ESG scores (on the 1 to 100 ESG score scale) of the analysed five ESG ratings providers compared to firms in civil law countries. Extending our view to the other models that include country-, firm- and industry- specific controls, but do not include interaction terms (Table 4, Models II, IV and VI) confirm our initial findings as the $LegOr$ regression coefficient remains negative and statistically significant, indicating lower absolute dispersion of ESG scores for firms in common law countries. This suggests a lower level of ESG dispersion disagreement between the rating agencies in common law countries in comparison with civil law countries.

When interaction terms are added (Table 4, Models III, V and VII), the main effect of the legal origin variable[3] becomes statistically insignificant (but remains negative), whereas some of the interaction effects (in particular, interactions $LegOr \times CliRScr$ and $LegOr \times FHScr$) turn out significant. This suggests that legal origins may be affecting the dispersion of firm-level ESG scores through specific channels. Specifically, our results indicate that progress of countries towards meeting the Paris Agreement goals (a higher $CliRScr$ score) is somewhat counterintuitively associated with a higher firm-level dispersion of ESG scores in common law countries, with results consistent across Models III, V and VII (Table 4). On the other hand, there is some indication that increased levels of freedom and democratic governance in a country lowers ESG ratings dispersion disagreement in common law countries (statistically significant interaction coefficients in Models V and VII, Table 4). Since regression coefficients on the interaction term of legal origins with the Government climate risk score ($LegOr \times CliRScr$) is the only consistently statistically significant regression coefficient, we also consider models excluding interaction terms of legal origins with Regulatory quality and Control of corruption, and GDP per capita variable (Table 6, model VIII). This model indicates that the Government climate risk score interaction term is no longer statistically significant, while Legal origin increases in magnitude and remains statistically significant, even when controlling for firm-level factors (Table 6, model IX and X). Those findings suggest that the relationship between Legal origin and absolute ESG dispersion is more nuanced than initially observed.

[3] The main effect in this case measures the impact of legal origins on firm-level ESG score dispersion *at the mean* of respective interacted country-level control variables.

Table 6. Regression analysis of factors affecting the standard deviation of ESG scores across models VIII-X.

Model	VIII	IX	X
LegOR	-3.248**	-3.002***	-2.935**
	(1.0306)	(1.0937)	(1.0233)
CliRScr	0.043	-0.040	0.047
	(0.1525)	(0.1595)	(0.1496)
FHScr	-0.034***	-0.027***	-0.021***
	(0.0062)	(0.0063)	(0.0059)
RQInd	1.435**	1.342*	0.718
	(0.5493)	(0.5922)	(0.5602)
CCInd	-1.595***	-1.557**	-1.388**
	(0.4154)	(0.4519)	(0.4261)
LegOr x CliRScr	0.253	0.255	0.284
	(0.2450)	(0.2542)	(0.2377)
ESGCovScr		0.013*	0.008*
		(0.0039)	(0.0040)
logTA		-0.225**	-0.332***
		(0.0813)	(0.0936)
ROE		-0.314*	-0.316**
		(0.1299)	(0.1210)
Communication services			6.940***
			(0.6750)
Consumer discretionary			5.812***
			(0.6105)
Consumer staples			3.122***
			(0.6271)
Financials			6.141***
			(0.6089)
Health care			4.737***
			(0.6239)
Industrials			4.029***
			(0.5791)
Information technology			6.171***
			(0.6104)
Materials			1.778**
			(0.6202)
Real estate			6.089***
			(0.7131)
Utilities			1.877**
			(0.6813)
Intercept	22.073***	26.125***	24.168***
	(0.5357)	(2.1252)	(2.5036)
Observations	2,275	2,005	2,005
R-squared	0.0980	0.1171	0.2403

Note: The table reports results from regression of ESG dispersion, measured as standard deviation, on explanatory variables. Model VIII includes the Legal origin, (demeaned) Government climate risk score, the Freedom House score, Regulatory quality index, Control of corruption index and Government climate risk score interaction as explanatory variables. Model IX adds the ESG controversies score, (logarithm of) Total assets, and ROE to model VIII. Model X includes 11 sectors to model IX: Communication services, Consumer discretionary, Consumer staples, Energy (benchmark), Financials, Health care, Industrials, Information technology, Materials, Real estate and Utilities. Model VII adds three above mentioned interaction terms to model VI. *, **, and *** indicate statistical significance at the 0.05, 0.01, and 0.001.

With respect to the other country-, firm- and sector-specific variables included in our analysis, the signs of regression coefficients remain mostly consistent across all model specifications. As expected, we find that higher levels of freedom and democratic governance and better

control of corruption generally lower ESG dispersion disagreement. The results for regulatory quality and government climate risk score are less robust to model specification. On the other hand, higher firm levels of ESG controversies are associated with increased levels of ESG dispersion disagreement. Interestingly, we find that larger and more profitable firms generally exhibit lower levels of ESG dispersion disagreement, which is in line with the fact that these firms are on the one hand more exposed to market scrutiny and sustainability related regulatory requirements, whereas on the other hand, higher profitability means they can allocate more resources towards assuaging ESG ratings providers. The estimated regression coefficients on the industry-specific dummies also call for interesting interpretations. Specifically, given that the Energy sector is used as the baseline, Models VI and VII (Table 4) indicate that the Energy sector has by far the lowest levels of ESG dispersion disagreement. This is expected, as the energy sector is one of the most heavily regulated sectors both with respect to the green transition *and* ESG reporting and disclosure requirements, making the ESG ratings providers' job relatively easier in comparison with the other sectors. Moreover, the results for the other sectors show that sectors with higher levels of "tangibility" of sustainability issues, such as Materials, Industrials, Utilities, or Consumer Staples tend to exhibit less ESG dispersion disagreement (albeit still higher than the Energy sector) in comparison to the sectors with lower "tangibility" of sustainability issues (such as for example Finance, IT and Communication services). This indicates that ESG reporting and measurement challenges may be substantially different across the firm landscape.

We now turn to the results of regressions with *coefficient of variation as the dependent variable* (Table 5). Coefficient of variation (CV) normalizes variability by the mean, acting as a relative measure of dispersion. This measure can be more informative when comparing ESG scores across firms with varying means. The results largely corroborate our findings from the regression analysis with standard deviation as the dependent variable. Moreover, the regression coefficient of the legal origins variable remains negative and is highly statistically significant across all different model specifications, including country-, firm-, industry- level controls as well as interaction terms (Models I to VII, Table 5). This supports our earlier observations and indicates that common law countries have lower levels of ESG dispersion disagreement, not only in absolute, but also in relative terms.

Table 5. Regression analysis of factors affecting the coefficient of variation of ESG scores across models I-VII.

Model	I	II	III	IV	V	VI	VII
LegOr	-11.466***	-3.654**	-21.001***	-2.487*	-18.223***	-2.668*	-16.790***
	(0.8374)	(1.1506)	(4.5083)	(1.1654)	(4.7874)	(1.1574)	(4.5240)
CliRScr		-2.440***	-2.858***	-2.498***	-2.809***	-2.543***	-2.701***
		(0.4735)	(0.7225)	(0.4633)	(0.7136)	(0.4595)	(0.7006)
FHScr		-0.173***	-0.267***	-0.138***	-0.199**	-0.129***	-0.190**
		(0.0221)	(0.0594)	(0.0215)	(0.0606)	(0.0212)	(0.0594)
RQInd		-13.108***	-5.570	-14.927***	-10.195*	-15.771***	-11.046**
		(1.9523)	(3.9912)	(1.9791)	(4.2347)	(1.9708)	(4.1547)
CCInd		8.426***	3.929	9.735***	6.643**	10.171***	7.013**
		(1.4013)	(2.255)	(1.4327)	(2.4344)	(1.4209)	(2.3934)
GDPpc		0.007	0.044	0.010	0.029	0.011	0.002
		(0.0230)	(0.0281)	(0.0221)	(0.0271)	(0.0218)	(0.0076)
LegOr x			2.928*		2.366*		1.895
CliRScr			(1.2132)		(1.1910)		(1.1704)
LegOr x			0.105		0.067		0.068
FHScr							
			(0.0634)		(0.0642)		(0.0631)
LegOr x			-8.509		0.153		0.042
RQind			(5.5404)		(5.6842)		(5.6003)
LegOr x			4.150		-0.375		0.210
CCInd			(4.0108)		(4.0410)		(3.9856)
ESGCovScr				0.007	0.008	-0.0133	-0.012
				(0.0131)	(0.0131)	(0.0136)	(0.0137)
logTA				-1.987***	-1.968***	-2.864***	-2.824***
				(0.2693)	(0.2695)	(0.3248)	(0.3244)
ROE				-1.052*	-1.004*	-0.995*	-0.951*
				(0.4297)	(0.4288)	(0.4214)	(0.4206)
Commun.						11.114***	11.589***
services						(2.3504)	(2.3484)
Consumer						7.122**	7.533***
discretionary						(2.1260)	(2.1255)
Consumer						3.385	3.627
staples						(2.1840)	(2.1799)
Financials						12.377***	12.519***
						(2.1204)	(2.1164)
Health care						1.808	2.266
						(2.1687)	(2.1683)
Industrials						5.732**	5.978**
						(2.0164)	(2.0125)
Information						8.341***	8.584***
technology						(2.1245)	(2.1236)
Materials						1.410	1.796
						(2.1596)	(2.1570)
Real estate						8.284**	8.445**
						(2.4823)	(2.4792)
Utilities						4.127	4.611
						(2.3708)	(2.3722)
Intercept	40.758***	63.210***	66.193***	106.918***	107.379***	123.105***	122.057***
	(0.5637)	(1.8776)	(2.7827)	(7.0348)	(7.3968)	(8.7656)	(8.8171)
Observations	2,392	2,275	2,275	2,005	2,005	2,005	2,005
R-squared	0.0727	0.3509	0.3565	0.3866	0.3913	0.4153	0.4195

Note: The table reports results from regression of ESG dispersion, measured as standard deviation, on explanatory variables. Model I includes only the Legal origin as the explanatory variable, model II adds country controls to model I, model IV adds firm controls model II, model VI includes eleven sectors to model IV with Energy sector as benchmark. Model III adds three interaction terms to model II, model IV adds firm controls model II, model V three interaction terms to model IV, model VI includes eleven sectors to model IV with Energy sector as benchmark, Model VII, V and VII adds three interaction terms to model II, IV and VI. *, **, and *** indicate statistical significance at the 0.05, 0.01, and 0.001.

When interaction terms are considered in regressions, containing coefficient of variation as the dependent variable, the only interaction effect that remains statistically significant and maintains the same (positive) sign as in the previous analysis (with standard deviation as the dependent variable), is the interaction between legal origins and the Government climate risk score ($LegOr \times CliRScr$). This confirms our earlier finding that better (higher) country level climate risk scores are related with a higher level of ESG dispersion disagreement in common law countries. The results, however, are only statistically significant in two out of three model specifications that include interaction terms (Models III and V, Table 5). As a further robustness check, we adopted the same approach as in our previous analysis and considered three models excluding interaction terms of legal origins with Regulatory quality and Control of Corruption variables, as well as the GDP per capita variable, which are not found to be significant in the main models. In these alternative specifications (Table 7, models VIII, IX, and X), the Government climate risk score remains positive and highly statistically significant across all three models.

Country-specific controls are highly statistically significant across most model specifications and have expected signs: higher levels of climate policy efforts, freedom and democratic governance, and regulatory quality are associated with lower relative firm-level dispersion of ESG scores across the five ratings providers. Somewhat surprisingly and contrary to our previous analysis, the sign of the regression coefficient on the control of corruption index becomes positive and is highly statistically significant in all but one model specification (the exception is Model II, Table 5, which includes interaction terms). This could indicate some sort of collusion of ESG ratings providers in countries with worse corruption control performance, leading to more similar scores. Firm-specific control variables keep the same sign (as well as significance) as in our previous analysis, confirming the importance of firm size and profitability for reduced ESG information asymmetries. Finally, regression coefficients on the sector-specific dummy variables also largely remain the same (and significant) as before, with the Energy sector (the baseline) showing the lowest levels of ESG dispersion disagreement and tangibility of sustainability issues again playing an important role in reducing ESG related information asymmetries as in regressions with standard deviation of ESG scores as the dependent variable.

Table 7. Regression analysis of factors affecting the coefficient of variation of ESG scores across models VIII-X.

Model	VIII	IX	X
LegOR	-15.833***	-12.972***	-12.152**
	(3.5888)	(3.6099(	(3.5562)
CliRScr	-3.333***	-3.298***	-3.260***
	(0.5310)	(0.5265)	(0.5199)
FHScr	-0.184***	-0.150***	-0.140***
	(0.0216)	(0.0209)	(0.0206)
RQInd	-10.716***	-12.500***	-13.488***
	(1.9127)	(1.9546)	(1.9468)
CCInd	7.0853***	8.411***	8.945***
	(1.4466)	(1.4915)	(1.4808)
LegOr x CliRScr	3.065***	2.598***	2.358**
	(0.8531)	(0.8390)	(0.8260)
ESGCovScr		0.004	-0.016
		(0.0130)	(0.0136)
logTA		-2.003***	-2.856***
		(0.2685)	(0.3241)
ROE		-0.989***	-0.936*
		(0.4287)	(0.4207)
Communication services			11.255***
			(2.3460)
Consumer discretionary			7.197**
			(2.1215)
Consumer staples			3.418
			(2.1796)
Financials			12.382***
			(2.1161)
Health care			2.241
			(2.1684)
Industrials			5.843**
			(2.0126)
Information technology			8.577***
			(2.1215)
Materials			1.549
			(2.1556)
Real estate			8.474**
			(2.4781)
Utilities			4.369
			(2.3677)
Intercept	66.016***	110.145***	125.2717
	(1.8655)	(7.0144)	(8.7009)
Observations	2,275	2,005	2,005
R-squared	0.3546	0.3895	0.4176

Note: The table reports results from regression of ESG dispersion, measured as coefficient of variation, on explanatory variables. Model VIII includes the Legal origin, (demeaned) Government climate risk score, the Freedom House score, Regulatory quality index, Control of corruption index and Government climate risk score interaction as explanatory variables. Model IX adds the ESG controversies score, (logarithm of) Total assets, and ROE to model VIII. Model X includes 11 sectors to model IX: Communication services, Consumer discretionary, Consumer staples, Energy (benchmark), Financials, Health care, Industrials, Information technology, Materials, Real estate and Utilities. Model VII adds three above mentioned interaction terms to model VI. *, **, and *** indicate statistical significance at the 0.05, 0.01, and 0.001.

6. Robustness considerations

Although panel analysis is often desirable in economic analysis, we decided on a cross-sectional approach based on three reasons. Firstly, ESG scores do not exhibit substantial variability over time, which means our results may not be that different from earlier (recent) periods. Secondly, ESG metrics, including ESG scores and ratings, while not new, are still a relatively novel concept. As a consequence, the cross-section tends to decrease as we move back in time to create a panel data structure. Thirdly, sustainability policy (which includes regulatory frameworks) is a fast-evolving field, which is particularly relevant for the area of our study, relating ESG information asymmetries to countries' legal origins (and in turn, regulatory characteristics). Making conclusions based on relationships that may reflect past policies in a fast-moving world is of questionable use, which is why we believe it is better to take a snapshot of the present, while at the same time acknowledging that a panel approach might prove very useful with future data.

In support of the first two arguments, we examined, as a robustness check, past ESG scores from a single ESG ratings provider (LSEG). The level of ESG score does not change drastically year-over-year. For example, when we examined the change in the LSEG ESG score from 2022 to 2023 for our sample of 2,392 firms, 1.17% (28 firms) are not in the 2022 sample as they were scored for the first time in 2023. Consequently, the sample size was reduced to 2,364 firms (Appendix, Table 7, Figure 1). Among the 2,364 firms, the level of ESG score from 2022 to 2023 changed by less than 1,67 points on average. More than 65% (1,537 firms) showed a change of less than 5 points, while less than 35% (827 firms) showed a change of more than 5 points. This suggests that while there is some variability in the level of ESG scores over time, the majority of this variability comes from a small proportion of firms at the tails of the distribution, i.e., outliers. Going back in time further, the distribution, and consequently the absolute dispersion, of year-over-year changes in the level of ESG score has remained largely consistent over the past 9 years (Appendix, Figure 2). If the year-over-year variability in ESG scores is similarly low across the other rating providers, we can expect that the overall temporal dynamics of ESG scores are relatively stable for most firms with the main drivers of variability in the level of ESG scores coming from outliers.

In support of the third argument, the regulatory landscape in the area of sustainability is changing rapidly. The EU in particular, is on a quest to construct a comprehensive sustainable

finance framework, spanning company disclosures, reporting standards, environmental taxonomies and sustainability metrics, including ESG ratings. For example, larger firms in the EU with over 250 employees and a 40-million-euro turnover, along with non-EU firms with 150 million euro in annual turnover within the EU, will be gradually required to make sustainability related disclosures according to the EU's Corporate Sustainability Reporting Directive (CSRD) and European Sustainability Reporting Standards (ESRS) in a digital, machine-readable format, starting from the fiscal year 2024. Increasing availability of standardized and (hopefully) transparent data could improve the consistency and reliability of ESG ratings across ESG ratings providers, particularly for firms in civil law countries. Furthermore, the EU also recently (as of April 2024) adopted the Regulation on the Transparency and Integrity of ESG Rating Activities (ESGR) to regulate ESG ratings providers operating within the EU. This is the first regulation globally addressing the ESG ratings market. Under the new rules, ESG rating providers will need to be authorised and supervised by the European Securities and Markets Authority (ESMA) and adhere to transparency requirements with respect to their methodology and sources of information. Thus, the civil law regulatory mechanism aims to enhance transparency from both sides, firms and rating agencies. If these regulations prove effective, we could expect reduced ESG dispersion for firms in civil law countries in the future.

7. Conclusion

The (pairwise Pearson) correlation analysis of ESG performance of 2,392 firms from 53 countries reveals that firms are more consistently ranked in civil law countries compared to common law countries. The results highlight the importance of considering legal origins when examining ESG performance of firms from different jurisdictions. The lower "correlation disagreement" for firms in civil law countries suggest that rating providers agree more on the sustainability performance of firms relative to other firms. This could be attributed to the enhanced regulatory environments typically found in civil law jurisdictions.

However, regression analysis of the cross-sectional data with ESG cores available in April 2024 reveals that, despite the lower "correlation disagreement" among rating providers, firms in civil law countries exhibit higher "dispersion disagreement" compared to firms from common law countries. This indicates that even when relative ranking of firms in civil law countries is more consistent (lower "correlation disagreement"), the overall spread or variability of ESG scores for individual firms across providers is larger (higher "dispersion

disagreement") in both absolute and relative terms. These findings underscore the significance of legal origin in understanding ESG performance of firms but also call for cautious interpretation. Although the "correlation disagreement" among rating providers on firms' ESG performance is lower for firms in civil law countries, "dispersion disagreement" is, on average, lower for firms in common law countries.

This study underscores the importance of country-level sustainability practices for firms in civil law countries. They experience lower "dispersion disagreement" among rating providers when government climate risk management improves. While common law countries also see an improvement in "dispersion disagreement" with better government climate risk management, the effect is notably smaller.

With the implementation of the EU's Corporate Sustainability Reporting Directive (CSRD), aimed at ESG reporting, certain firms in the EU are required to comply in the 2024 financial year, with their reports to be published in 2025. Over the following years all EU companies will gradually be obligated to disclose the data. Additionally, European Parliament adopted the Regulation on the Transparency and Integrity of ESG Rating Activities (ESGR), the first legislation at a global level to regulate the ratings market. It would be interesting to study effect of legal origin on ESG dispersion in the future, when this data becomes available.

Highlites

1. Rating providers have lower "correlation disagreement" on ESG performance of firms incorporated in civil law countries compared to firms incorporated in common law countries.
2. "Correlation disagreement" represents disagreement among rating providers on ranking firms based on ESG performance, measured by pairwise Pearson correlation coefficient.
3. Rating providers have lower "dispersion disagreement" on ESG performance of firms incorporated in common law countries compared to firms incorporated in civil law countries.
4. "Dispersion disagreement" is measured by standard deviation (absolute ESG dispersion) or coefficient of variation (relative ESG dispersion).
5. When ESG dispersion is measured in relative terms, the better the government manages climate risk, the lower the "dispersion disagreement" for firms incorporated in civil law countries.

Literature

Avramov, D., Cheng, S., Lioui, A., & Tarelli, A. (2022). Sustainable investing with ESG rating uncertainty. *Journal of Financial Economics*, *145*(2), 642–664. https://doi.org/10.1016/j.jfineco.2021.09.009

Avramov, D., Chordia, T., Jostova, G., & Philipov, A. (2009). Dispersion in analysts' earnings forecasts and credit rating. *Journal of Financial Economics*, *91*(1), 83–101. https://doi.org/10.1016/j.jfineco.2008.02.005

Balli, H. O., & Sørensen, B. E. (2013). Interaction effects in econometrics. *Empirical Economics*, *45*(1), 583–603. https://doi.org/10.1007/s00181-012-0604-2

Berg, F., Fabisik, K., & Sautner, Z. (2021). *Is History Repeating Itself? The (Un)Predictable Past of ESG Ratings* (SSRN Scholarly Paper 3722087). https://doi.org/10.2139/ssrn.3722087

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

Chatterji, A. K., Durand, R., Levine, D. I., & Touboul, S. (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

Chatterji, A. K., Levine, D. I., & Toffel, M. W. (2009). How Well Do Social Ratings Actually Measure Corporate Social Responsibility? *Journal of Economics & Management Strategy*, *18*(1), 125–169. https://doi.org/10.1111/j.1530-9134.2009.00210.x

Chatterji, A. K., & Toffel, M. W. (2010). How firms respond to being rated. *Strategic Management Journal*, *31*(9), 917–945. https://doi.org/10.1002/smj.840

Christensen, D. M., Serafeim, G., & Sikochi, A. (2022). Why is Corporate Virtue in the Eye of The Beholder? The Case of ESG Ratings. *The Accounting Review*, *97*(1), 147–175. https://doi.org/10.2308/TAR-2019-0506

Damaska, M. R. (1991). *The Faces of Justice and State Authority: A Comparative Approach to the Legal Process* (Revised ed. edition). Yale University Press.

Dorfleitner, G., Halbritter, G., & Nguyen, M. (2015). Measuring the level and risk of corporate responsibility – An empirical comparison of different ESG rating approaches. *Journal of Asset Management*, *16*(7), 450–466. https://doi.org/10.1057/jam.2015.31

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

Dumrose, M., Rink, S., & Eckert, J. (2022). Disaggregating confusion? The EU Taxonomy and its relation to ESG rating. *Finance Research Letters*, *48*, 102928. https://doi.org/10.1016/j.frl.2022.102928

Fishburn, P. C. (1977). Mean-Risk Analysis with Risk Associated with Below-Target Returns. *The American Economic Review*, *67*(2), 116–126.

Gibson Brandon, R., Krueger, P., & Schmidt, P. S. (2021). ESG Rating Disagreement and Stock Returns. *Financial Analysts Journal*, *77*(4), 104–127. https://doi.org/10.1080/0015198X.2021.1963186

La Porta, R., Lopez-de-Silanes, F., & Shleifer, A. (2008). The Economic Consequences of Legal Origins. *Journal of Economic Literature*, *46*(2), 285–332. https://doi.org/10.1257/jel.46.2.285

Liang, H., & Renneboog, L. (2017). On the Foundations of Corporate Social Responsibility. *The Journal of Finance*, *72*(2), 853–910. https://doi.org/10.1111/jofi.12487

Mahoney, P. G. (2001). The Common Law and Economic Growth: Hayek Might Be Right. *The Journal of Legal Studies*, *30*(2), 503–525. https://doi.org/10.1086/322053

Pedersen, L. H., Fitzgibbons, S., & Pomorski, L. (2021). Responsible investing: The ESG-efficient frontier. *Journal of Financial Economics*, *142*(2), 572–597. https://doi.org/10.1016/j.jfineco.2020.11.001

Widyawati, L. (2021). Measurement concerns and agreement of environmental social governance ratings. *Accounting & Finance*, *61*(S1), 1589–1623. https://doi.org/10.1111/acfi.12638

Appendix

Table 1. Countries by frequency, percent, cumulative and legal origin.

Country	Frequency	Percent	Cumulative	Legal origin
Australia	49	2.05	2.05	1
Austria	4	0.17	2.22	0
Belgium	11	0.46	2.68	0
Bermuda	15	0.63	3.30	1
Brazil	41	1.71	5.02	0
Canada	82	3.43	8.44	1
Cayman Islands	65	2.72	11.16	1
Chile	10	0.42	11.58	0
China	435	18.19	29.77	0
Colombia	2	0.08	29.85	0
Czech Republic	3	0.13	29.97	0
Denmark	14	0.59	30.56	0
Egypt, Arab Rep.	1	0.04	30.60	0
Finland	12	0.50	31.10	0
France	57	2.38	33.49	0
Germany	51	2.13	35.62	0
Greece	10	0.42	36.04	0
Hong Kong, China	36	1.51	37.54	1
Hungary	3	0.13	37.67	0
Indonesia	19	0.79	38.46	0
Ireland	21	0.88	39.34	1
Isle of Man	1	0.04	39.38	1
Israel	12	0.50	39.88	1
Italy	20	0.84	40.72	0
Japan	215	8.99	49.71	0
Jersey	5	0.21	49.92	1
Korea, Rep.	83	3.47	53.39	0
Kuwait	7	0.29	53.68	0
Liberia	1	0.04	53.72	1
Luxembourg	4	0.17	53.89	0
Malaysia	30	1.25	55.14	1
Mexico	23	0.96	56.10	0
Netherlands	37	1.55	57.65	0
New Zealand	6	0.25	57.90	1
Norway	12	0.50	58.40	0
Panama	1	0.04	58.44	0
Peru	1	0.04	58.49	0
Philippines	14	0.59	59.07	0
Poland	12	0.50	59.57	0
Portugal	3	0.13	59.70	0
Qatar	12	0.50	60.20	0
Saudi Arabia	34	1.42	61.62	1
Singapore	20	0.84	62.46	1
South Africa	15	0.63	63.09	1
Spain	17	0.71	63.80	0
Sweden	35	1.46	65.26	0
Switzerland	43	1.80	67.06	0
Taiwan	80	3.34	70.40	0
Thailand	34	1.42	71.82	1
Turkey	16	0.67	72.49	0
United Arab Emirates	9	0.38	72.87	1
United Kingdom	76	3.18	76.05	1
United States	573	23.95	100.00	1
Total	2,392	100.00		19/53

Note: The table shows the frequency, percentage, and cumulative percentage of countries in the sample of 2,392 firms, along with an indicator of their legal origin. The table includes 53 countries, with 19 countries identified as having a common legal origin (1).

Table 2. Descriptive statistics of ESG scores across providers.

ESG score	Observations	Mean	Standard deviation	Min	Max
Bloomberg	2,392	39.30	14.49	6.40	86.70
S&P Global	2,392	69.13	26.20	0.00	100.00
LSEG	2,392	61.33	18.10	2.90	95.57
MSCI	2,392	58.54	25.45	0.00	100.00
Sustainalytics	2,392	75.77	92.18	27.92	94.56

Note: The table presents the descriptive statistics for ESG scores from five providers. The statistics include the number of observations, mean scores, standard deviation (Std. Dev.), minimum scores (Min), and maximum scores (Max) for each of five ESG providers, namely Bloomberg, S&P Global, LSEG, MSCI and Sustainalytics.

Table 3. The Global Industry Classification Standard (GICS) sector by frequency, percent and cumulative.

GICS sector name	Frequency	Percent	Cumulative
Communication services	127	5.31	5.31
Consumer discretionary	237	9.91	15.22
Consumer staples	186	7.78	22.99
Energy	87	3.64	26.63
Financials	396	16.56	43.19
Health care	205	8.57	51.76
Industrials	394	16.47	68.23
Information technology	297	12.42	80.64
Materials	224	9.36	90.01
Real estate	117	4.89	94.90
Utilities	122	5.10	100.00
Total	2,392	100.00	

Note: There are 11 sectors in total, including communication services, consumer discretionary, consumer staples, energy, financials, health care, industrials, information technology, materials, real estate, and utilities.

Table 4. Evaluation of linear combination of interaction coefficients and Legal origin coefficient for models III, IV and VII where dispersion is measured as standard deviation.

	Model III		Model V		Model VII	
Interaction coefficient	Coefficient	p-value	Coefficient	p-value	Coefficient	p-value
LegOr x CliRScr	-2.796* (1.342955)	0.037	-2.581 (1.441237)	0.073	-1.637 (1.348877)	0.225
LegOr x FHScr	-2.412 (1.29569)	0.063	-2.117 (1.394281)	0.129	-1.290913 (1.304606)	0.323
LegOr x RQInd	-2.553 (1.41423)	0.071	-2.753 (1.498972)	0.066	-2.235 (1.397556)	0.110
LegOr x CCInd	-2.952 (1.61116)	0.067	-2.473 (1.784172)	0.166	-1.807 (1.667637)	0.279

Note: There are no statistically significant linear combinations of interaction coefficients and Legal origin coefficient for models III, IV and VII when dispersion is measured as standard deviation, except for the Governance climate risk score interaction coefficient in model III. *, **, and *** indicate statistical significance at the 0.05, 0.01, and 0.001.

Table 5. Evaluation of linear combination of interaction coefficients and legal origin coefficient for models III, IV and VII where ESG dispersion is measured as coefficient of variation.

	Model III		Model V		Model VII	
Interaction coefficient	Coefficient	p-value	Coefficient	p-value	Coefficient	p-value
LegOr x CliRScr	-23.859*** (4.690562)	0.000	-21.032*** (4.766391)	0.000	-19.491*** (4.69807)	0.000
LegOr x FHScr	-21.269*** (4.525479)	0.000	-18.422*** (4.611102)	0.000	-16.980*** (4.543874)	0.000
LegOr x RQInd	26.672*** (4.939507)	0.000	-28.418*** (4.957332)	0.000	-27.836*** (4.867615)	0.000
LegOr x CCInd	-17.072*** (5.627329)	0.002	-11.580* (5.900529)	0.050	-9.777 (5.808294)	0.092

Note: There are statistically significant linear combinations of interaction coefficients and legal origin coefficient for models III, IV and VII when dispersion is measured as coefficient of variation, except for the Control of corruption interaction coefficient in model VII. *, **, and *** indicate statistical significance at the 0.05, 0.01, and 0.001.

Table 6. Evaluation of linear combination between Government climate risk score interaction coefficient and Legal origin coefficient for models VIII, IX and X where ESG dispersion is measured as standard deviation.

	Model VIII		Model IX		Model X	
Interaction coefficient	Coefficient	p-value	Coefficient	p-value	Coefficient	p-value
LegOr x CliRScr	-2.994*** (0.8007439)	0.000	-2.747** (0.8558104)	0.001	-2.651** (0.8011617)	0.001

Note: *, **, and *** indicate statistical significance at the 0.05, 0.01, and 0.001.

Table 7. Evaluation of linear combination between Government climate risk score interaction coefficient and Legal origin coefficient for models VIII, IX and X where ESG dispersion is measured as coefficient of variation.

	Model VIII		Model IX		Model X	
Interaction coefficient	Coefficient	p-value	Coefficient	p-value	Coefficient	p-value
LegOr x CliRScr	-12.767*** (2.788459)	0.000	-10.374*** (2.824663)	0.000	-9.794*** (2.784316)	0.000

Note: *, **, and *** indicate statistical significance at the 0.05, 0.01, and 0.001.

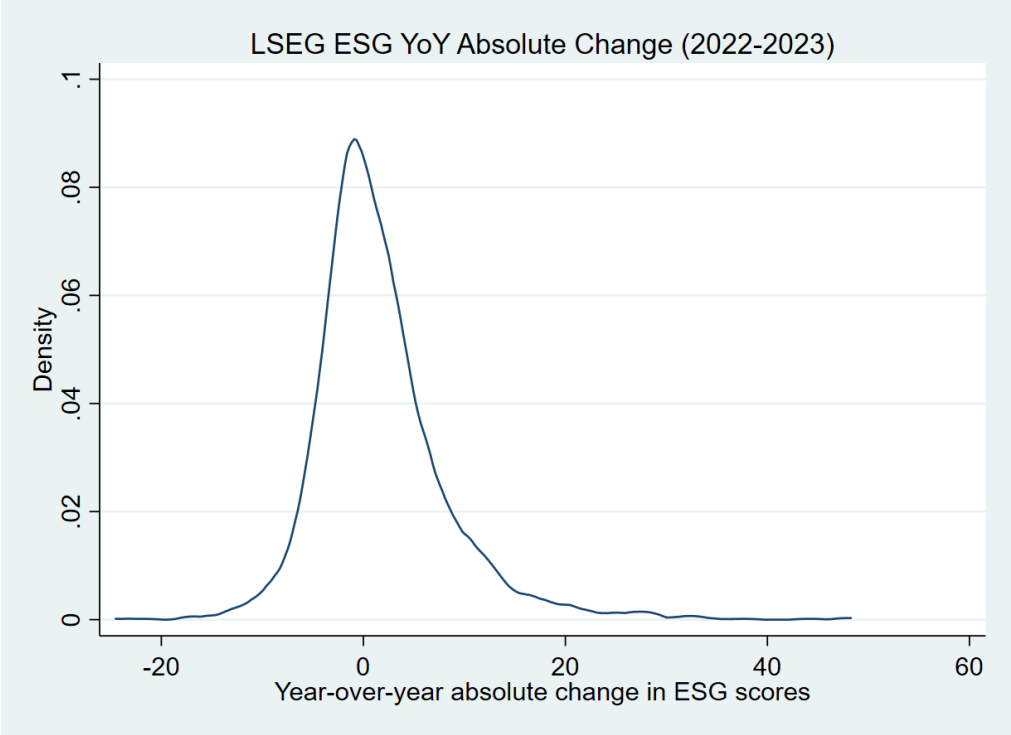
Table 8. Summary statistics of the year-over-year absolute change in LSEG ESG score (2022-2023).

Percentiles	YoY change	Smallest
1%	-11.2949	-24.51189
5%	-6.3220	-21.87914
10%	-4.7525	-17.36653
25%	-2.20437	-16.70483
50%	0.6379	
75%	4.3351	Largest
90%	9.4419	37.60362
95%	13.1732	44.20448
99%	25.4328	47.76462
		48.25695

Observations	Average	Sta. dev.	Variance	Skewness	Kurtosis
2,364	1.6727	6.560706	43.04287	1.501261	8.691184

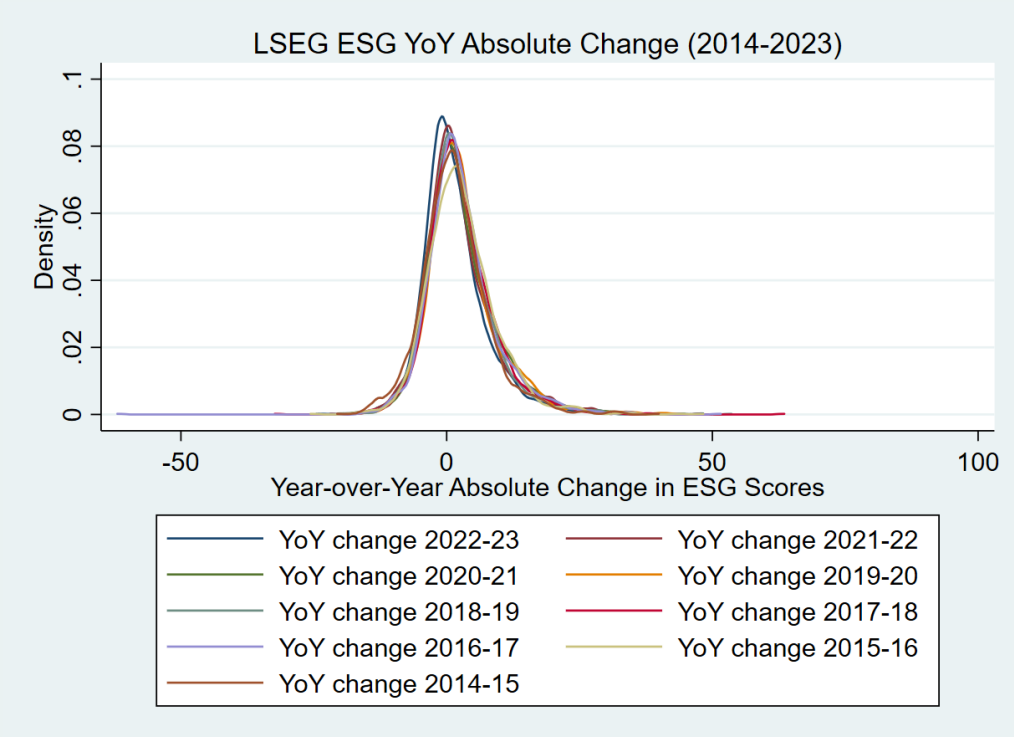
Note: Summary statistics of year-over-year absolute change in LSEG ESG score (2022-2023).

Figure 1. Density distribution of year-over-year absolute change in LSEG ESG score (2022-2023).



Note: Density distribution of year-over-year absolute change in LSEG ESG score (2022-2023)

Figure 2. Density distribution of year-over-year absolute change in LSEG ESG core (2014-2023).



Note: Density distribution of year-over-year absolute change in LSEG ESG core (2014-2023).