Final work : Do ESG factors predict credit rating downgrades ? Evidence from S&P-rated European public companies

Auteur : Zhou, Yuemin
Promoteur(s) : Blanchard, Gildas
Faculté : HEC-Ecole de gestion de l'Université de Liège
Diplôme : Master de spécialisation en gestion des risques financiers
Année académique : 2025-2026
URI/URL : http://hdl.handle.net/2268.2/25413

Do ESG factors predict credit rating downgrades?

Evidence from S&P-rated European public companies

Promoteur :	Travail de fin d'études présenté par
Gildas Blanchard	**Yuemin ZHOU**
Lecteur(s) :	en vue de l'obtention du diplôme de
David Suetens	Master de spécialisation en Gestion des risques financiers, Module Sustainable and Climate Finance
	Année académique 2025/2026

Table of Contents

Do ESG factors predict credit rating downgrades?
Evidence from S&P-rated European public companies

Abstract:

This study examines the relationship between environmental, social, and governance (ESG) factors and corporate credit rating downgrades, distinguishing between input-based ESG scores and output-based ESG metrics. Using a panel of publicly listed corporates in European developed markets rated by S&P Global, the analysis investigates whether ESG-related variables provide incremental predictive information for downgrade risk beyond traditional financial, macroeconomic, and structural controls. The results show that output-based indicators including CO_2 emissions, ESG controversies, and country-level physical climate vulnerability, are consistently associated with higher downgrade likelihood, whereas commonly used input-based ESG scores are not. Firms with missing CO_2 disclosures exhibit lower observed downgrade risk, consistent with disclosure-driven selection effects. The study provides robust evidence that realized ESG outcomes contain relevant information for downgrade risk, with implications for credit risk assessment and IFRS 9 SICR frameworks.

Keywords: credit risk, SICR, ESG factors, credit rating downgrades

1. Introduction

Under IFRS9 framework, banks should integrate forward-looking perspective when estimating expected credit loss (expected loan loss provisioning) in addition to historical events and current situation (BIS, 2017). ECB has addressed that the impact of climate change and environmental degradation on economic activity can result in structural changes that affect the financial system and its stability (ECB, 2020). Regulators are putting increasing emphasis on banks ingratiating novel risks which include climate-related and environmental risks (CER) to ensure adequate accounting provisions. Nevertheless, most banks are unable to capture novel risks in a fully functional and validated statistical model due to insufficient data (ECB, 2024).

Under IFRS9 Framework, banks must classify financial assets into three stages to measure expected credit loss:

- Stage 1: performing, no significant credit risk increase
- Stage 2: performing but significant increase in credit risk
- Stage 3: credit impaired / defaulted

Significant Increase in Credit Risk (SICR), a concept introduced by the IFRS9 framework, is a forward-looking approach to identify the timing of classification of Stage 2. When SICR is recognized, loans initially classified as Stage 1 transit to Stage 2, which requires recognizing lifetime expected credit loss (ECL) instead of only 12-month ECL. Several factors could be applied to define SICR, such as probability of default (PD) increase, rating downgrades, macroeconomic changes, and other possible qualitative indicators.

Rating downgrades are widely used as one of the observable proxies for SICR events, especially for non-systematic institutions with insufficient internal data. As ECB (2025) states, "Banks with weaker data and IT infrastructures tend to overly rely on ratings in their early warning systems and credit risk processes". While this study cannot directly observe banks' internal bank ratings, probability of default models or actual defaults, it helps us to understand better what's behind credit risk ratings. To specify, this paper will investigate whether ESG performance is correlated with company-level credit rating downgrades. It aims to investigate whether ESG performance is reflected in publicly observable rating downgrades, thus whether downgrade-based staging already reflects ESG risks. Since rating agencies incorporate forward-looking credit assessments, a significant relationship would suggest that ESG factors may serve as a proxy for early credit risk signals, providing insights for IFRS9 SICR monitoring. For banks that use external ratings as an input to SICR, this study can help banks to anticipate the rating actions of credit risk agencies, to understand the systematic biases embedded in credit ratings, and thus to improve the effectiveness of SICR identification. In another word, understanding ESG's predictive power on downgrades can help banks enhance forward-looking monitoring for Stage 2 recognition even before external downgrading events become available. Even for banks that rely primarily on their internal models, the findings of this study could serve as a reference: if rating agencies explicitly consider ESG factors, banks should also consider the necessity of integrating ESG risks into its internal models of credit risk.

The main research question is thus whether ESG factors influence the likelihood of company level rating downgrades (i.e. whether ESG riskier firms are more likely to be downgraded).

On top of the main research question, we are also interested in exploring several sub-questions including:

1) How does ESG interact with traditional financial metrics (such as leverage, profitability, tangibility, firm size) in predicting rating downgrades?
2) Which ESG dimensions (governance, environmental, social) are most strongly associated with rating downgrades?
3) Would deterioration in ESG performance increase the probability of downgrades?
4) Does ESG-rating downgrade relationship vary over time, reflecting evolving ESG integration in rating agency methodologies?
5) What implications can the relationship between ESG information and rating downgrades indicate for banks to monitor forward-looking credit risk under IFRS9?

To answer the above questions, we will make use of logit regression to explore the relationship between downgrading events and company-level ESG variable of the publicly listed companies whose primary operations are in the European developed markets. We base our analysis on S&P ratings and will make use of LSEG Refinitiv to source ESG related data, which include both raw and factual ESG variables and ESG-related scores.

The paper is organized as follows. Section 2 reviews literature to see the existing research on exploring the relationship between ESG risks and credit risk including its components. Section 3 also reviews exiting literature to understand the relationship between ESG factors and credit risks, which will help us to choose the types of variables to be considered and to determine the hypotheses before fitting the regression model. We will discuss Environmental (E), Social (S), and Governance (G) factors separately. As for the environmental dimension, we focus mainly on transition climate risk. For physical climate risk, as accurate firm-level data is not easily accessible, we will study it using country-level data in the section of robustness analysis. Next, Section 4 explains the methodologies and variables that are or are potentially used by the rating agencies, particularly S&P, to decide credit risk ratings. These variables would thus be used as control variables. Section 5 presents the methodology used in this study, the logit regression. Furthermore, the data sources, data treatment and cleansing processes are described in detail in Section 6. Section 7 provides the results from different models which could be interpreted to draw conclusions on the above-mentioned research questions. What's more, with robustness analysis in section 8, we will explore potential overfitting issues in our models. We will also discuss the implications from the results and the limitations confronted by the study together with the possible future developments in section 9 and 10 respectively. Lastly, our conclusions will be drawn in section 11.

2. Literature Review

Traditional models of corporate credit risk largely focus on the risk of bankruptcy or default, dating back to the Z-score model of Altman in 1968. The model incorporates financial ratios to predict a corporate's likelihood of bankruptcy by the estimated Z score. It basically answers the question: how risky the firm is now (Altman, 1968). Shumay extended the study with a dynamic hazard rate model, which estimates the conditional probability that a firm will default in a given period (Shumway, 2001).

The earliest research on the link between ESG factors and corporates was initially from an equity investor's perspective. Early in 1978, Spicer (Spicer, 1978) provided empirical evidence on a moderate to strong association between company equity's investment value and its social performance. Sharfman and Fernando studied firms' environmental risk management and cost of capital and found a negative correlation between them (Sharfman & Fernando, 2008). Only since recent fifteen to twenty years has there been an increasing amount of research exploring the relationship between ESG factors and corporate credit risk. One of the earliest empirical research projects on ESG and corporate credit risk is from the study of Bauer and Hann in 2010. They have found that environmental concerns are associated with a higher cost of debt financing and lower credit ratings (Bauer & Hann, 2010), based on evidence from 582 US public corporations between 1995 and 2006. The paper demonstrated that environmental concerns constitute a meaningful risk for bank lenders. In the same year, Schneider addressed that a firm's environmental performance is a significant component of its bankruptcy risk (Schneider, 2010). Later, Chava has shown that environmental profile of a firm is a proxy for an omitted component of its default risk (Chava, 2011).

Bessis breaks down credit risk into three dimensions: default, exposure, and recovery, reflecting in expected credit loss as being a product of probability of default (PD), loss given default (LGD), and exposure at default (EAD) (Bessis, 2010). A substantial number of studies used probability of default as a target variable to explore the impact of ESG factors. Ferriani and Pericoli provided robust evidence from European listed non-financial corporates on the negative correlation between firms' sustainability profiles and their creditworthiness and on the significant time-varying relationship between ESG risks and PDs (Ferriani & Pericoli, 2024). Instead of directly using PDs provided by rating agencies (Ferriani and Pericoli used PDs from Moody's Expected Default Frequency), Wattanatorn developed its own time-varying model based on Duan et al.'s forward-intensity probability scheme using a wider range of economies and horizons, confirming the risk-reducing role of firms' ESG performance (Wattanatorn, 2025).

Another metric, Merton's distance-to-default is a widely used proxy for PD. It measures the number of standard deviations the firm's asset value must fall for default to occur. Atif and Ali found, from a sample of US nonfinancial institutions, a positive relationship between ESG disclosure and Merton's distance-to-default (Atif & Ali, 2021). Ma and Hu shed light from emerging markets that ESG performance significantly reduces corporate default and violation risk, measured from Merton's distance-to-default (Ma & Hu, 2026).

In contrast, not many studies are performed on LGD, EAD or SICR related topics mainly due to lack of data. Scandizzo and Hughes provided instructions on calculating climate-related loss given default (Scandizzo & Hughes, 2025). Zhang et al. are one of the limited empirical studies on ESG and LGD modeling. The paper has demonstrated a significant and temporal relationship between LGD (estimated by Moody's Default and Recovery Database) and ESG factors. Besides, the predictive power of ESG factors strengthens during adverse macroeconomic environment (Zhang, Andreeva, & Dong, 2025).

While the studies discussed above help us to understand how ESG factors affect various dimensions of credit risk, we perform also a brief review of literature on the determinants of credit risk ratings. Credit ratings, different from parameters such as PD or LGD, represent a forward-looking opinion from the rating agencies on the credit risk of a counterparty or security. Rating agencies such as S&P Global, Moody's, and Fitch perform credit assessments based on both qualitative and quantitative approaches, which may not be 100% transparent. Understanding the determinants of credit risk ratings can help users understand potential systematic bias behind the ratings and how to mitigate risks more effectively. Early research on credit ratings used financial ratios as determinants (Horrigan, 1966). Even today, the validity of this finding remains. Oliveira and Basso have listed 18 literature references relative to the determinants of credit ratings dated from 2006 to 2023, among which the most frequently tested variables are leverage, profitability, size, liquidity related financial metrics. They further revealed that interest coverage, profitability, Tobin's Q and Total Shareholder Return (measuring company's market performance), and Altman's Z-score (measuring bankruptcy risk) were significant factors in explaining credit ratings. However, ESG factors are not explicitly tested by all these existing literatures (Oliveira & Basso, 2023).

On the other hand, credit rating downgrades may in turn affect the market, especially large downgrades and those degrading from investment to speculative grades. Kladakis and Skouralis found that market's perspectives are influenced by credit rating changes, which have a significant effect on institution's contribution to systemic risk. This adverse impact of downgrades can be partly mitigated by stronger accounting-based stability, such as higher profitability and capital (Kladakis & Skouralis, 2023).

From the above literature review, we have seen a theoretical gap existing in the intersection between ESG factors and credit risk ratings, especially in the European countries. More importantly, no literature yet connects ESG factors to the IFRS9 SICR identification. This study, therefore, aims at addressing these gaps by focusing on European developed markets' public companies and thereby informing whether ESG factors can be incorporated into SICR frameworks with banks that use external credit ratings as inputs.

3. ESG and credit risks

The following section discusses in detail the relationships between each pillar of ESG risks – environmental, social, and governance - and credit risk, based on existing empirical and theoretical research. We will show the channels through which ESG factors can influence firm's risk profile. This section also aims to identify a set of measurable proxies of each ESG dimension that can be considered within our empirical model. Meanwhile, it seeks to establish clear and testable hypotheses regarding the expected direction of these relationships. Such structured approach will provide the theoretical and conceptual foundation for our empirical model.

3.1 E (Environmental)

The emerging novel risks addressed by ECB to be considered in the ECL calculation include mainly systematic and macro risk drivers such as energy supply, geopolitical stability, high interest rates, inflation, and climate change, among which climate related risks have become a key focus of many banks today (ECB, 2024).

In 2021, BCBS issued a report discussing climate-related risk drivers and their transmission channels. By definition, climate risk drivers represent climate-related changes, including physical and transition risks, that could give rise to financial risks. Physical risks arise from the changes in weather and climate that impact economies. Transition risk drivers arise as a result of transitioning an economy that is reliant on fossil fuels to a low-carbon economy. While physical risks could be distinguished between acute physical hazards and chronic physical hazards, transition risks are further divided into four categories including policy and legal risks, technology, market changes, and reputation risk (BCBS, 2021).

Climate risk, both physical and transitional climate risks, can affect corporate credit risk through both direct and indirect channels. Directly, climate-related physical risks such as extreme weather events or transition risks arising from changes in regulations can impair companies' asset values and increase the probability of financial distress. Indirectly, through macro- and micro-economic transmission channels, climate risk drivers could be translated to an increase in credit risk and cause financial impact to banks. The impact of transmission is firm specific as macro risks do not impact all firms equally. Both physical and transition risks are transmitted to credit risk through firm-level balance sheet and income statement channels. Climate risk can influence key financial fundamentals that are vital to credit rating assessment such as profitability, leverage, and liquidity.

Extensive empirical and theoretical studies have investigated the dynamic relationship between climate risk and credit risk. Capasso et al. showed in their study that the distance-to-default, a widely used market-based measure of corporate default risk, is negatively associated with a firm's exposure to climate risks which is measured by the amount of carbon emissions and carbon intensity (Capasso, Gianfrate, & Spinelli, 2020). Bauer and Hann demonstrated that poor environmental performances are associated with worse credit ratings (Bauer & Hann, 2010).

Other papers provide evidence on the indirect effects of climate risk to credit risk through financial fundamentals. The paper of Nguyena and Phanb found that after Kyoto Protocol ratification, heavy carbon emitting firms started to decrease their level of financial leverage. Such reduction in leverage was especially strong for firms that already face difficulty accessing finance. The reason is that increased carbon risk leads to higher financial distress risk, making firms more cautious about taking financial leverage (Nguyen & Phan, 2020). Hugon and Law showed that firms' earnings, on average, are negatively impacted by an unusually warm climate (Hugon & Law, 2018). Huang et al. found that firms in countries with higher climate risk have poorer economic performance as measured by return-on-assets (ROA) and cash-flows from operations over assets (CFO) as well as more volatile earnings (Huang, Kerstein, & Wang, 2018).

Physical climate risk is by nature location-specific and ideally requires detailed information on the geographic distribution of a firm's tangible assets and operations. However, such granular data is not always available for small and multinational firms. Some industry practices use country-level climate vulnerability indicators based on the location of firm's headquarters as a proxy of physical climate risk. However, this approach is based on strong assumptions that the countries of headquarters are where firms primarily operate. Therefore, it does not fully capture multinational exposure, especially firms with headquarters in low-risk countries but operations are in higher-risk regions. Besides, this proxy omits within-country variation in physical climate risk. Since the proxy is at country level, it also does not reflect firm-specific risks. We deem that to have a meaningful conclusion on the link between physical risk and credit risk of corporates, more granular data and more detailed analyses are necessary. Due to these limitations, our main model focuses exclusively on transition risks as a proxy for the environmental dimension of ESG risks. However, in the robustness analysis, we will test the impact of physical climate risk using a country-level physical climate vulnerability index, serving as a coarse proxy of the physical climate risk of companies.

In this study, we will also measure firm's exposure to transition climate risk by the amount of carbon emission and carbon intensity, which is the ratio between the carbon emission and revenues. And we develop the following null hypothesis:

H1: Firm's exposure to transition climate risk is negatively associated with its probability of being downgraded.

3.2 S (Social)

The social dimension of ESG encompasses a variety of factors that reflect the relationship between a firm and its stakeholders, such as employees, customers, and suppliers. Among these, human capital, which comprises employee satisfaction, training, health and safety, and ethical labor practices, is critical to operational stability and reputation of the firm. Another key aspect is diversity and inclusion. More and more companies are actively promoting gender diversity and fostering open communication channels between employees and the top management team. Ahold Delhaize has introduced its "100/100/100" aspiration: 100% gender equity, 100% reflective of communities its brand serves, and 100% inclusive culture (Ahold Delhaize, 2026). BNP Baribas Group has developed transparent and seamless dialogue channels for each type of stakeholder. Additionally, customers and community relations are also vital aspects of the social dimension. It concerns particularly product safety, data privacy, and supply chain management. In recent years, as firms seek to align with various standards (such as UN Guiding Principles on Business and Human Rights, Modern Slavery Act, and EU Corporate Sustainability Due Diligence Directive, etc.), large corporates make efforts to enhance audit and monitoring mechanisms across their global supply chains. The aim is to avoid human right abuses and environmental misconduct in supply chains (BNP Paribas Group, 2025).

The social component of ESG is currently not explicitly emphasized by regulators in terms of ECL calculation. Direct research on the impact of the social dimensions of ESG factors on credit risk is extensively less than for those for environmental or governance dimensions. However, it does not mean social dimension does not contribute to firms' credit risk.

Social factors could affect credit risk through operational, human capital, and social capital channels. Firms with weaker human rights performance are more subject to operational disruptions, reputational damages, regulatory and legal risks. For example, poor labor practices such as high turnover and high employee fatal or injury rates could disrupt operations and reduce productivity. Social controversies such as use of child labor, data breaches, and lawsuits can harm firms' reputation, leading to a loss of revenue and possible higher financing costs, let alone higher legal costs including fines from the regulators. These issues could finally deteriorate creditworthiness and increase the likelihood of credit rating downgrades.

As for the empirical evidence on the relationship between social factors and credit risk, several studies have revealed the impact of human capital on firm value and credit risk. The study of Edman implies that employee satisfaction is strongly positively correlated with shareholder returns and need not represent managerial slack, which is consistent with human capital-centered theories (Edmans, 2011). One indicator of employee satisfaction is employee turnover, which measures the loss of employees over a period of time (March & Simon, 1958). An annual rate is calculated as the ratio between the number of employees who quit the firm and the total number of employees of that year. Hossain et al. studied microfinance institutions (MFIs) and found that employee turnover raises the credit risk of MFIs (Hossain, Mia, & Hooy, 2023). Moreover, Hasnaoui and Hasnaoui studied commercial banks in the GCC and concluded that human capital efficiency is negatively related to the credit risk profile and banks with higher human capital efficiency tend to have lower credit risk (Hasnaoui & Hasnaoui, 2022). In terms of gender diversity, existing studies mainly focus on the relation between board gender diversity and credit risk, which will be further discussed in the next section. As for social capital, Yeboah et al. emphasized the importance of community connections by addressing that Individuals embedded in strong social networks were consistently perceived as more creditworthy (Yeboah, Mogre, Menzo, & Prince, 2024).

Rating agency S&P Global has addressed that social dimension is harder to quantify than the other dimensions and default observations linked to social issues are still limited (Sindaco, 2021). Thus, the agency put in place a scorecard methodology: a ranking system of economic activities based on firms' social sustainability. The social factors considered are safety management, human right management, consumers and demographic trends, and social cohesion and community. These factors will be integrated into the analysis of a company's management and governance, industry risk, competition position, and financial analysis. In other words, social factors are implicitly transmitted to credit risk via these analyses.

In our study, to empirically assess the relationship between social factors and credit rating downgrade risks, we decided to test three distinct yet complementary metrics: workforce score, human capital score, and ESG controversies score. All of them are provided by LSEG workspace. According to LSEG definitions (LSEG, 2026), the workforce score measures a company's effectiveness towards job satisfaction, healthy and safe workplace, maintaining diversity and equal opportunities, and development opportunities for its workforce. Human right score measures a company's effectiveness towards respecting the fundamental human rights conventions. ESG controversies score measures a company's exposure to environmental, social and governance controversies and negative events reflected in global media. This metric could partially cover the aspects of consumer and supply chain relations. We have found empirical evidence showing that ESG controversies have a negative effect on credit ratings (Samaniego-Medina & Giraldez-Puig, 2022). According to this paper, ESG controversies act as a relevant negative factor in occurring an upgrading event in future credit assessments.

Finally, we may also consider a metric called labor intensity. It is the ratio between labor cost and total revenue to capture firm exposure to workforce-related risks, which may amplify the impact of social factors on credit risk.

With these quantified metrics, we are able to assess whether the current credit rating assessment could capture the "S" related risks.

Furthermore, based on the above literature review and theoretical analysis, we could make the following null hypotheses:

H2a: Firms with higher workforce scores (indicating stronger labor practices, employee relations, and workplace safety) are more likely to experience credit risk downgrades.

H2b: Firms with better human right performance are more likely to face credit risk downgrades.

H2c: Firm is involved more in ESG controversies are less likely to be downgraded.

3.3 G (Governance)

The inclusion of the "G" dimension within ESG reflects a worldwide understanding that environmental and social outcomes ultimately depend on how a company is organized and controlled. Governance provides a framework that ensures environmental and social commitments are not merely symbolic but are effectively implemented and monitored.

Corporate governance has become increasingly central to credit risk assessment as both regulators and market participants have recognized that weak governance structures can amplify financial vulnerabilities and could lead to credit deterioration. Pirson and Turnbull pointed out that failures in corporate governance were one of the origins of the 2008 financial crisis (Pirson & Turnbull, 2018) (Ballester, González-Urteaga, & Beatriz, 2020).

A number of regulatory frameworks and international initiatives explicitly incorporate governance as a core pillar of ESG, reinforcing its importance beyond the United Nations Sustainable Development Goals. Reforms such as enhanced board oversight, stricter disclosure requirements, and strengthened shareholder rights were introduced to improve transparency and accountability of firms. For instance, the G20/OECD Principles of Corporate Governance provides a main international benchmark for sound corporate governance practices. Also take the banking industry as an example, regulators have significantly strengthened their oversight over financial institutions through structured supervisory processes such as Supervisory Review and Evaluation Process (SREP). Under SREP, supervisors conduct comprehensive assessment of banks' business model, solvency and liquidity positions, as well as internal governance.

Before turning to empirical evidence, it is important to first discuss the theoretical foundations of corporate governance. The Principal-Agent (P-A) relationship operates when a party, the principal, contracts or instructs the other party, the agent, to perform a delegated task (Bo & Driver, 2012). In modern corporate governance, principal refers to shareholders and agent refers to managers. Shareholders delegate decision making authority to managers. Adam Smith has already identified in 1776 that "the directors of such companies, however, being the managers rather of other people's money than of their own, it cannot well be expected, that they should watch over it with the same anxious vigilance with which the partners in a private copartnery frequently watch over their own." (Smith, 1776) Such tension between principal and agents is formalized by the well-known Agency Theory. It creates agency costs in monitoring and incentivizing managers to act in shareholders' best interests. Two key manifestations of this misalignment are Moral Hazard and Adverse Selection problems. The former arises when managers may engage in more aggressive risk-taking or speculative behavior because they do not fully bear the downside risk of their actions. The latter happens when managers take advantage of information asymmetry (i.e. information held only by managers). These issues are particularly relevant in the context of credit risk, as weak governance can lead to excessive risk-taking and omission of risk exposures. Mechanisms such as board oversight, board independence settings, executive compensation design, and disclosure requirements are often adopted to mitigate P-A problems. Hence, a sound corporate governance acts as a critical line of defense against management misconduct and excessive risk-taking. As Garcia et al., put, firms' risk-taking policies ultimately depend on the corporate board (García, Herrero, & Morillas, 2022).

Empirical research has also proven strong association between corporate governance and credit risks and indicated different metrics representing corporate governance. Switzer et al. studied the relationship between corporate governance and default risk in financial firms over the post-financial crisis period (Switzer, Wang, & Tu, 2016). They considered five key internal governance mechanisms: institutional ownership, insider ownership, board independence, board size, and CEO duality. These measures could also be found in other papers (e.g. (García, Herrero, & Morillas, 2022)). Ballester, et al. has summarized 68 literatures focusing on the role that internal corporate governance plays in firm's credit risk (Ballester, González-Urteaga, & Beatriz, 2020). The paper mainly analyzed three governance components developed by S&P Global for assessing corporate governance: ownership structure, board structure, and financial stakeholders' rights and relations. Ownership structure could be split into ownership concentration, foreign, state and family ownership. The paper marked that the decreasing effect of ownership concentration on credit risk is not found in European market. Besides, family ownership reduces credit risk only in European countries, where the use of excess control rights, which are found to increase credit risk. The paper also found through a systematic literature review that, in terms of board structure, CEO duality (the practice of a single individual serving as both CEO and board chair) has a positive effect on credit risk while female and foreign directors show negative effects. Another paper from Adusei et al. provided evidence from microfinance institutions (MFIs) that board gender diversity is negatively and significantly related to MFIs' financial performance. Surprisingly, an MFI that combines 50% or higher female representation on its board is likely to experience a tumble in its financial performance (Adusei, Akomea, & Poku, 2017). Yeboah et al primarily focusing on SME companies in developing countries, recommended analyzing cautiously the role of women in financial decision-making to avoid underestimating creditworthiness of corporates (Yeboah, Mogre, Menzo, & Prince, 2024). Although we are not only focusing on MFIs or SMEs, we shall still be cautious when making our hypothesis towards the metric.

Moreover, Ballester, et al. concluded that the board size does not show clear effect on credit risk in Europe and there is no consensus on the role played by board independence (Ballester, González-Urteaga, & Beatriz, 2020). Indeed, generally higher board independence is expected to be associated with lower credit risk. Regulatory reforms in response to Enron-like governance failures concentrate heavily on board independence (Wang & Lin, 2010). However, in some cases, overly independent boards may lack firm-specific knowledge and expertise, potentially weakening decision-making in complex environments.

Credit rating agencies such as Moody's ratings and S&P explicitly include board independence, management quality and internal controls in their credit risk assessment process. Therefore, in our study, subject to constraints of data availability, we will test the Governance pillar on board independence, board gender diversity, and management quality.

The null hypotheses shall be drawn cautiously:

H3a: Board independence is not correlated with the probability of downgrading.

H3b: Board gender diversity is not correlated with the probability of downgrading.

H3c: Firms with better management quality are more likely to be downgraded.

4. Traditional factors impacting creditworthiness

4.1 S&P Global rating methodologies

The following section introduces the methodologies applied by S&P Global when determining credit risk ratings for corporates. The narratives are based on the General Corporate Method issued in Jan 2024. We aim to understand the key drivers of credit risk ratings to determine a list of control variables.

According to S&P Global, credit ratings are forward-looking opinions about an issuer's relative creditworthiness. Credit risk agencies evaluate creditworthiness through a combination of quantitative financial metrics (such as debt ratios, cash flow, and liquidity) and qualitative factors (such as competitive position, industry dynamics, and management effectiveness). These financial measures are assessed both in absolute terms and relative to industry peers, considering both historical performance trends and future outlook (S&P Global, 2026).

S&P publishes its up-to-date methodology of determining credit risk ratings for corporates regularly. Take the corporate industrial companies and utilities as an example, the following framework issued in July 2025 is used to grant credit risk ratings.[1]

The assessment begins with analyzing the company's business risk and financial risk profiles. Each type of profile has its own scale of scores. The combination of the scores from these two risks profiles of a company helps to determine its initial anchor.

Additional six rating factors are then considered to modify the anchor, which are diversification/portfolio effect, capital structure, financial policy, liquidity, and management together with governance.

The following graphs extracted from the General Corporate Method of S&P Global illustrate such framework and the modification process (S&P Global, 2024).

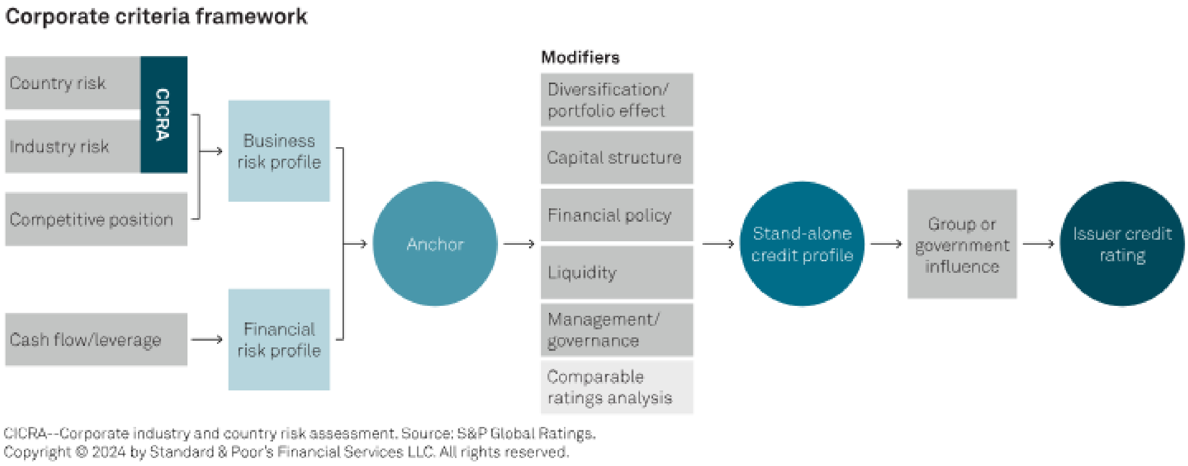


Figure 1: Corporate criteria framework (source: S&P Global)

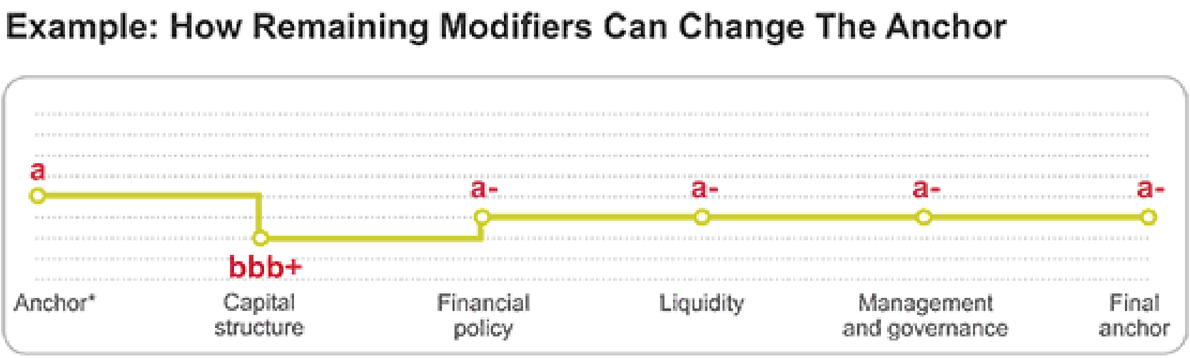


Figure 2: Rating modifiers (source: S&P Global)

[1] Note that this framework does not apply to project finance entities, project developers, commodity trading and investment holding companies, certain NGOs, and other entities whose cash flows are primarily derived from partially owned equity holdings.

Business risk assessment is mainly a qualitative assessment with a small portion of quantitative assessment where the industry risk, country risk, and the company's competitive position are analyzed. The analysis of competitive position could be further divided into four aspects: 1) analyzing competitive advantages; 2) scale, scope, and diversity; 3) operating efficiency; and 4) profitability and its volatility. There is flexibility in determining the subfactors of these aspects and the weight of each subfactors across different companies. For example, depending on types of companies, the following weights are applied according to S&P's General Corporate Method.

Standard Competitive Position Group Profiles (CPGPs) And Category Weightings

	--(%)--					
Component	Services and product focus	Product focus/scale driven	Capital or asset focus	Commodity focus/cost driven	Commodity focus/scale driven	National industries and utilities
1. Competitive advantage	45	35	30	15	10	60
2. Scale, scope, and diversity	30	50	30	35	55	20
3. Operating efficiency	25	15	40	50	35	20
Total	100	100	100	100	100	100
Weighted-average assessment*	1.0-5.0	1.0-5.0	1.0-5.0	1.0-5.0	1.0-5.0	1.0-5.0

*1 (strong), 2 (strong/adequate), 3 (adequate), 4 (adequate/weak), 5 (weak).

Figure 3 : Examples of category weightings on competitive positions (source: S&P Global)

It is worth mentioning that S&P Global do consider environmental and social credit factors, but only as one of the subfactors of the obligor's competitive position. Besides, it is not written explicitly how much importance is given to these factors and whether they are considered in each of the companies rated.

This assessment inspires us of several quantitative indicators that we could consider in our control variables. For example, the scale could be measured by sizes of markets and revenues or market share.

Diversity could be measured by number of business lines and global market presence. Financial metrics such as EBITDA margin and return on capital could be used to measure profitability. (S&P Global, 2024)

Next, the financial risk assessment analyzes mainly the cash flow and leverage of the company. Funds from operations (FFO) to debt and debt to EBITDA could be used to assess the cash flow and leverage of a company. One of the issues of the debt to EBITDA ratio is that, normally the greater the ratio is, the higher leverage the firm is taking. However, when EBITDA becomes negative, the ratio loses its economic meaningfulness. Therefore, we prefer using total liability over total assets to act as a proxy to firm's leverage.

Among the six modifiers mentioned above, the diversification/portfolio effect modifier is partially covered by the subfactor Diversity of Business risk assessment. The capital structure modifier could be measured by weighted average of debt maturity and company's exposure to interest rate and currency risks. Liquidity is another key factor of credit risk rating as a lack of liquidity could push an otherwise healthy company to default. The main quantitative measures of a company's liquidity used by S&P Global are the ratio and the difference between the liquidity sources and uses. The sources include cash and liquid investment, positive FFO, positive forecasted working capital inflows, and proceeds of asset sales, etc. Whereas the uses include negative FFO, expected capital spending, negative forecasted working capital inflows, and all near-mature debts, etc. (S&P Global, 2024) While these data are not usually easily accessible, two common proxies of a company's liquidity would be the current ratio and quick ratio. The former is the ratio between current assets and current liabilities, and the latter is the ratio between the quick assets (cash, marketable securities, account receivables) and current liabilities.

The financial policies modifier is highly qualitative-based, and it is correlated to the management's financial discipline and whether the company is controlled by a financial sponsor. A financial sponsor-owned company is defined as "nonfinancial corporate entities in which one or more financial sponsors own at least 40% of the entity's common equity or retain the majority of the voting rights and control through preference shares, and where the sponsors exercise control of the company is considered as either solely or jointly." (S&P Global, 2024) Nevertheless, this information is not easily accessible.

As for the management together with governance modifier, S&P Global has identified five main subfactors which are: 1) ownership structure 2) Board structure, composition, and effectiveness; 3) risk management, internal controls, and audit 4) Transparency and reporting 5) Management (S&P Global, 2024). The assessment is mainly expert judgment based. Research has shown that while greater board independence can mitigate agency problems, it may also reduce the board's overall expertise (Expert Director Hypothesis), weaken the quality of its advisory role (Excessive Monitoring Hypothesis), as well as increase information acquisition costs (Information Cost Hypothesis) (Chen, 2014). Therefore, the direction of the effect of independent directors remains unknown. The common indicator of board independence is the ratio of the number of independent directors to total board size.

As we can see, while the G (Governance) component of ESG is explicitly considered in credit risk ratings through the evaluation of management effectiveness, the E (Environment) and S (Social) components are not.

4.2 Potential control variables

Altman proposed to combine several measures into a meaningful model to predict firm bank solvency issues (Altman, 1968). We have therefore shortlisted the following aspects of information that could be relevant to downgrading events:

- Industry fixed effects

  Industry fixed effects (FE) shall be included to control unobserved and time-invariant heterogeneity across industries that may systematically influence credit risk. Firms operating in different industries face different industry-specific characteristics including business models, competitive dynamics, capital structure, and regulatory environments. All these differences could affect their financial and risk profiles and thus their likelihood of experiencing credit rating downgrades. For instance, industries such as information technology and financial services tend to be less exposed to environmental risks, while manufacturing and energy companies are more exposed to both physical and transition risks. Utility firms have more stable cash flows while cyclical industries such as energy companies are more impacted by economic fluctuations.

Therefore, failing to control such differences may bias the estimated effects of firm-level variables and ESG factors.

Besides, controlling industry effect is a standard approach in empirical corporate finance and credit risk studies. For example, the ECB working paper (Schwaab, Koopman, & Lucas, 2016) included 6 industry-specific factors in their model predicting firms' systematic default risk. Dumrose and Höck has included industry FE to study the relationship between corporate carbon risk and credit risk (Dumrose & Höck, 2023). Dang et al. has performed panel regression analysis by including industry FE (Dang, Li, & Yang, 2018). The paper of Bauer and Hann on ESG and credit risk has addressed the importance of accounting for the heterogeneity across industries (Bauer & Hann, 2010).

In our final dataset, we made use of the variable "TRBC Economic Sector Name" of a company to capture its industry FE. This variable contains 10 distinct values: Financials, Industrials, Basic Materials, Technology, Consumer Cyclicals, Consumer Non-Cyclicals, Energy, Real Estate, Healthcare, and Utilities. Using economic sectors instead of more granular industry subcategories attempts to limit the number of dummy variables in our main model.

- Year fixed effects

Year fixed effects shall be included to capture time-specific shocks that affect all firms systematically. Also, as ESG environment evolves due to regulatory changes or increasing market attention and ESG reporting quality improves over time, we need to add this control variable to account for such general evolution experienced by all firms. A symbolic evolution is the Paris Agreement in 2015, which marked a major shift in global climate regulations and policies and significantly increasing market attention among investors, regulators, and credit rating agencies. In the paper of Mehmood et al. which studies the impact non-interest income on bank credit risk, year fixed effects are applied to control the model (Mehmood & Luca, 2023). Similarly, Dumrose and Höck included year FE to study the relationship between corporate carbon risk and credit risk. (Dumrose & Höck, 2023)

- Size effect

S&P Global has explicitly stated that company size is one of the important elements for credit risk assessment (Lui, 2021). Existing studies have shed evidence that larger firms are likely to use capital efficiently and generate more profitability (Supiyadi & Novita, 2023) and small companies have a higher default risk (Ting, 2011). Rapposelli et al. also found that larger companies perform better than medium-sized and smaller companies in terms of efficiency (across all business profiles), including default on bank loans (Rapposelli, Birindelli, & Modina, 2023). However, banks that are too big may experience negative returns due to bureaucratic complexity and other factors (Supiyadi & Novita, 2023). Firm complexity has a negative effect on corporate credit ratings. Negative effects diminish for more transparent and better-governed firms and intensify during periods of high policy uncertainty. Note that board composition also contributes to firm complexity (Dang, Puwanenthiren, Mazur, Hoang, & Nadarajah, 2025). Other studies showed that the firm size may not have a significant impact on credit risk (Faisal & Wenno, 2023). Therefore, we cautiously assume that size effect is negatively associated with the probability of downgrading events.

Dang et al. has investigated 100 empirical papers from top finance, accounting, and economics journals that use firm size measures on the topics of empirical corporate finance and concluded three most popular size effect proxies: total assets, total sales, and market value of equity (Dang, Li, & Yang, 2018). Among the 87 papers that use single measures, 49 of them use total assets. After all, these three measures are empirically highly correlated with each other as shown in Dang's paper. In our study, we also use one single measure even though we understand no measure can capture all-characteristics of a firm’s size. As Dang et al. states, total assets measure total firm resources, market capitalization involves firm growth opportunities and equity market condition, and total sales measures product market competition and is not forward looking.

In our case, total assets are more suitable as this variable represents operational scale and is less volatile than total revenue or sales. Besides, as section 6 puts, the total revenues from business activities extracted our data source could become negative (likely due to accounting corrections, discontinued operations or restructuring). Market capitalization is even more volatile, and it is more relevant when studying equity market.

- Business diversity

Number of business lines could be an indicator of the diversity of a firm’s business. However, such information is not easily available. After all, firm size and industry classification already capture a large part of diversification.

- Profitability

Firm’s profitability provides important information on the probability for a firm to fail (Beaver, 1968). Asset Profitability is a key metric in differentiating the financial health and operational efficiency of companies in this analysis (Demirhan, 2024). Wang showed that the growth rate of net profit is negatively correlated with credit risk (Wang L. , 2020). Therefore, we assume that profitability effect is negatively associated with the probability of downgrading events.

While traditional accounting metrics such as Return on Asset (ROA) or Return on Equity (ROE) are widely used in the literature, these metrics rely on net income, which can be volatile due to non-operating items, tax effects, or accounting anomalies, particularly in firms with negative revenue or those undergoing financial distress. EBITDA, by excluding interest, taxes, depreciation, and amortization, provides a more stable and comparable measure of profitability.

Given the presence of negative revenue values in our sample, we decided to use EBITDA to asset ratio, which is a common proxy for profitability. It is a measure of the true productivity of the firm's assets, abstracting from any tax or leverage factors (Altman, 1968). Altman also stated that since a firm's ultimate existence is based on the earning power of its assets, this ratio appears to be particularly appropriate for studies dealing with corporate failure. Shumway has included this variable in predicting bankruptcy risk (Shumway, 2001). Dumrose and Höck used EBIT asset ratio as a proxy of the probability control variable to study the relationship between the corporate carbon risk and credit risk (Dumrose & Höck, 2023).

Therefore, EBITDA to asset ratio offers a robust and practical way to capture core profitability and preserve sample integrity.

- Leverage

Credit risk increases for firms with high leverage and idiosyncratic volatility (Dodd, Kalimipalli, & Chan, 2021). To reflect leverage effect on corporate credit risks, S&P Global has explicitly stated the use of debt to EBITDA ratio. As explained in section 6, there are companies with negative EBITDA in our sample, which leads to negative ratios. However, a lower and negative debt to EBITDA ratio does not mean lower leverage. Therefore, we decided to use traditional measures such as the ratio between total liability and total assets to represent a firm’s leverage. This measure captures the overall financial obligations of the firm and provides a comprehensive view of its balance sheet structure. While some studies focus on interest-bearing debt, total liabilities offer a broader indicator of solvency, encompassing both financial and operating aspects.

- Liquidity

Current and quick ratios are crucial in assessing firms’ short-term financial health (Demirhan, 2024). Current ratio is the ratio between current liabilities and current assets. Quick ratio is equal to the difference between current liabilities and inventory divided by current assets. Demirhan showed, with the evidence from production and manufacturing sector of ISE public listed companies, that current ratio, as a critical financial metric, plays a significant role in distinguishing firms’ financial characteristics, particularly liquidity. Moreover, a declining current ratio may indicate rising liabilities or diminishing assets, which could be used as early warning signs of financial distress. Therefore, we assume that liquidity effects are negatively associated with the probability of downgrading events.

- Governance

The governance dimension of ESG has been explicitly considered by credit rating agencies. Therefore, we include governance-related metrics as control variables. Different from other control variables, we decide to test three metrics, as mentioned above: board independence, board gender diversity, and management quality measured by the management score (an input-based metric) provided by Refinitiv.

- Country-specific variables

Rating agencies assess country risk where a firm generates cash flows and is exposed to risk, not just legal registration. Therefore, country of headquarter is more relevant than country of incorporation. However, if a company operates in multiple countries, which is often the case for public companies, it is challenging to obtain such information from our databases. Furthermore, country-specific characteristics could be partially captured by the country’s macro-economic variables. Hossain et al. used GDP growth rate as country-specific control variable (Hossain, Mia, & Hooy, 2023). Therefore, since we will consider several macro-economic variables of the countries of firms’ headquarters, we have decided not to add another variable to represent country-specific variables (except the country vulnerability index, a coarse proxy of firms’ physical climate risk).

- Macro-economic variables

On top of the above factors that are considered to determine credit risk ratings, empirical evidence also shows strong correlation between macroeconomic factors and credit risk. For instance, evidence from Malaysia showed that unemployment rate and real GDP growth have strong impact on the level of non-performing financing for each financing portfolio (Isaev & Masih, 2017). Mcwayizeni examined the effects of macroeconomic variables on non-performing loans and found that factors such as GDP, interest rates, exchange rates, unemployment, and trade volumes significantly influence the dynamics of non-performing loans in emerging economies, but not significantly in developed economies (Mcwayizeni Thwala, 2016). Castro, focusing on Greece, Ireland, Portugal, Spain and Italy, found that credit risk is significantly influenced by the macroeconomic environment (Castro, 2013). More specifically, credit risk increases when GDP growth slows down and when unemployment rate rises. Besides, Figlewski et al. found that among all corporate issuers in Moody's corporate bond Default Research Database during the period 1981 - 2002, the intensity of occurrence of a credit event was strongly influenced by both macroeconomic and ratings-related factors such as real GDP growth, inflation, and unemployment rate (Figlewski, Frydman, & Liang, 2012). As inflation is usually highly correlated with unemployment rate, we decide to add only GDP annual growth and unemployment rate as control variables.

To summarize, the selected control variables aim to capture the main structural, financial, and macroeconomic determinants of credit rating downgrades identified in the theoretical and empirical literature. Industry and year fixed effects are included to account for unobserved heterogeneity and common time-specific factors, while firm-level financial characteristics such as size, profitability, leverage, and liquidity, control for differences in operational scale, financial performance, and balance-sheet strength. Governance-related metrics are incorporated to reflect qualitative aspects of firm oversight and those explicitly considered by rating agencies. Finally, country-level macroeconomic conditions are included to capture the economic environment in which firms operate. This comprehensive set of controls is designed to avoid possible bias due to omission and ensure that the estimated effects of ESG-related variables reflect incremental explanatory power for downgrade risk.

5. Methodology

Logit models are widely used in the credit risk literature as they ensure predicted probabilities lie within the range between zero to one and allow for nonlinear relationships between explanatory variables and the target variable (measures of credit risk). Previous studies such as Campbell et al. (Campbell, Hilscher, & Szilagyi, 2008) and Shumay (Shumway, 2001) demonstrate the suitability of logit-type models for corporate default prediction. Same as default or bankruptcy events, downgrading or not is a binary variable.

King and Zeng introduced a rare-event logit regression model for binary dependent variables with dozens to thousands of times fewer ones than zeros (King & Zeng, 2003). In our final dataset, there are in total 698 downgrading events out of 9406 observation points. The overall probability of downgrading is around 7.5%. Thus, we use the standard logit regression model to study the relationship between ESG factors and downgrade events.

The target variable is set as follows.

$$Downgrade_{i,t} = \begin{cases} 1 & if\ firm\ i\ experiences\ a\ rating\ downgrade\ in\ year\ t \\ 0 & otherwise \end{cases}$$

As decided in the previous literature review, the candidate explanatory variables will include macroeconomic variables such as GDP and unemployment rate, and firm-level financial controls such leverage, profitability, size, industry to account for temporal and sectoral differences, and year fixed effect. Petersen (Petersen, 2008) mentioned that panel data sets are data sets where observations can be grouped into clusters (e.g. multiple observations per firm, per industry, per year, or per country). Among the papers with approaches for estimating the coefficients and standard errors in the presence of within cluster correlation, 29% included dummy variables for each cluster (e.g. fixed effects or within estimation). It is also important to control at year level to address unobserved time effect, since the year fixed effects can help control the evolving ESG-related environment.

As for ESG factors, we split certain variables into factors at absolute level and change in these factors. The level factors may include variables such as carbon intensity, controversy score, employee turnover, and board independence (% of independent directors at the board), depending on the data availability.

The logit regression model is:

$$P\left(Downgrade_{i,t} = 1\right) = \frac{1}{1 + e^{-Z_{i,t}}}$$

Where

$$Z_{i,t} = constant + \beta_1 * ESG_{i,t-1} + \beta_2 * \Delta ESG_{i,t-1} +$$
$$\beta_3 * Control_{i,t-1} + \beta_4 * Macro_{i,t-1} +$$
$$\beta_5 * Industry\ FE + \beta_6 * Year\ FE + \varepsilon_{i,t}$$

and

$$\Delta ESG_{i,t-1} = ESG_{i,t-1} - ESG_{i,t-2}$$

The dataset used to fit the logit model has a panel structure, with yearly observations per firm over time, which may induce correlation in the error terms within firms. Clustered standard errors adjust for correlation of errors within a firm group, thus mitigate the bias caused by non-independent residuals (Breuer & deHaan, 2024). Without clustering, standard errors are biased and the significance could be false. For example, if a firm's bad governance persists, its relevant risk persists too, leading to correlated errors across years. Petersen found that compared to OLS, the Fama-MacBeth standard errors and Newey-West standard errors, only clustered standard errors are unbiased as they account for the residual dependence created by the unobserved firm effect (Petersen, 2008).

6. Data

The data set used to fit the logit model has a panel structure, with yearly observations per firm over time. The following section introduces how we obtain the final dataset from data sourcing to treatment and to the final variable selection.

5.1 Data sourcing

Our data is extracted from two data sources: S&P Capital IQ Pro and LSEG Refinitiv Workspace. The former is used to retrieve a list of active public companies from the European developed market that are rated by S&P as well as a list of companies that have been downgraded in the past since 1995. The latter is used to retrieve historical financial and ESG indicators. Indeed, Refinitiv does provide issuer-rating-related information, but it mixes different rating scales from different rating agencies (S&P, Moody's, Fitch, ICRA, etc.) with a considerable number of missing values. Thus, we stick to S&P in order to be consistent. In a word, the integration of these databases allows for the alignment of firm fundamentals, ESG characteristics, and credit risk events over time.

The first step is to build the target variable: flag downgrade as 1 if the downgrading happens in the next cohort year and 0 otherwise.

A key challenge in the dataset is the absence of continuous rating histories. However, with our current Capital IQ license, we can download a full history of discrete rating actions including new ratings, upgrade, downgrade, credit watch and non-rated actions, together with a list of companies that have latest credit ratings. In the Company Screening tab, we filtered company type as public company, geographical location as European developed markets, and key development events including the above-mentioned rating actions. The company's geographical location is determined based on a comprehensive market classification methodology and in most cases refers to the company's primary market or operational focus.

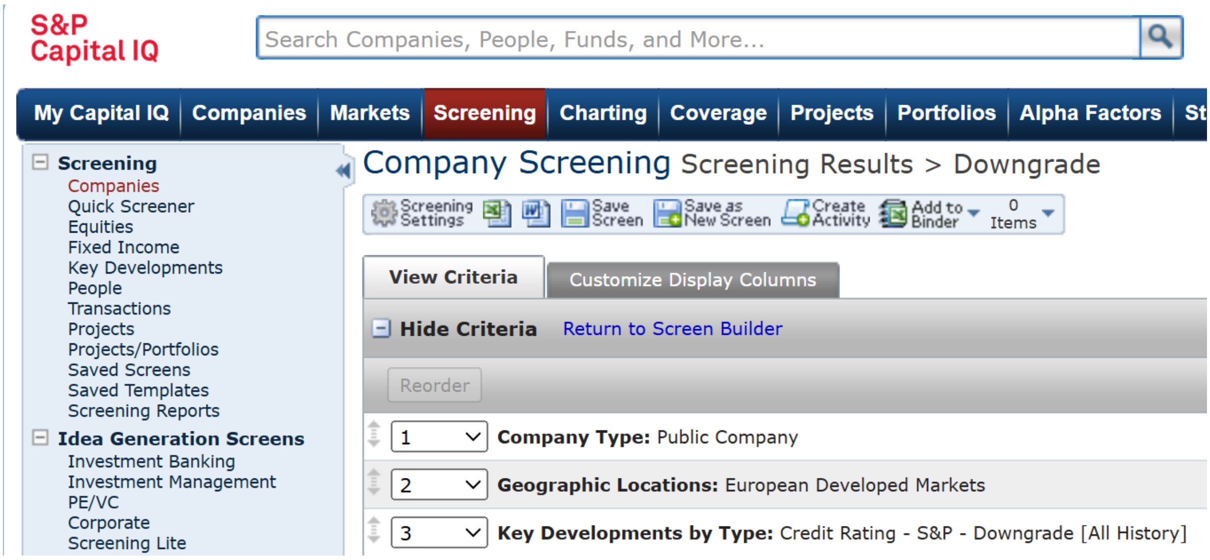


Figure 4 : Screenshot of S&P Capital IQ platform on company screening process (source: Capital IQ)

The key columns downloaded include company name, legal entity identification number (LEI/CICI code), capital IQ ID, and rating actions. Take 3M Company (NYSE: MMM.N) as an example, it has experienced the following downgrading actions.

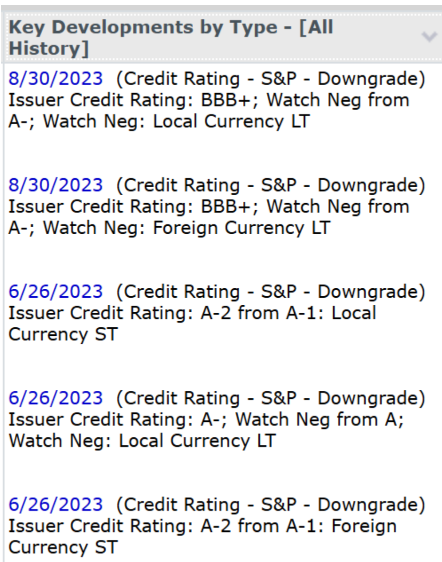


Figure 5 : Example of raw data on downgrade events (source: Capital IQ)

Note that credit ratings differ by local currency and foreign currency. Compared to local currency rating, the latter includes additional sovereign risk as well as transfer & convertibility risk. The BIS working paper by Amstad et al addressed that Sovereign debt in local currency has historically been considered safer than sovereign debt in foreign currency reflecting the view that sovereigns are much less likely to default on local currency obligations as they can always tax their subjects or print their own currency to service the debt (Amstad, Packer, & Shek, 2018). It also provided strong evidence that higher FX reserves, and less reliance on foreign currency borrowing. Therefore, to reflect pure credit risk, we only keep the actions on local currency ratings. Besides, we keep only long-term ratings instead of short-term ratings. This is because the former reflects structural credit risk, which is consistent with the characteristic of ESG risk. Moreover, we filtered out the ratings under the old rating scale (e.g. A-1 and A-2). Note that for each rating under the old rating scale, S&P provides their equivalent ratings under the new rating scale. Therefore, there is no loss of information if we perform such filter.

Like this, we can extract the date of rating action, the ratings before downgrading and after downgrading. After combining all types of rating actions, we sort the data by rating events in chronological order, and a continuous rating window is defined between the first and last observed rating-related events. This approach relies on the assumption that rating agencies, such as S&P, do not frequently withdraw ratings without explicit notification. Then for the period after the non-rating action and before the next other types of actions, we mark it as non-rated period, which should be later removed from our dataset. Years without any rating action are treated as non-downgrading observations.

The output at this step contains all European developed market's public companies (924 in total), their sorted rating action dates, the next rating action date, and the flag of non-rated actions.

The financial and ESG related variables are sourced from the LSEG Refinitiv workspace through LSEG CodeBook, a cloud-hosted development environment integrated into LSEG Workspace. We use its Jupiter Notebook interface (where the necessary API libraries are preinstalled and configured) to get access to financial and ESG data.

Next, we used the LEI code (904 companies with valid LEI code) to retrieve financial and ESG variables from LSEG Refinitiv Workspace. We need to firstly map the LEI code into the Reuters Instrument Code (RIC), which further leads to a loss of 17 companies. So far, 4% of data is lost due to technical reasons, which is not deemed as material.

The table below summarizes the description of all variables downloaded from Refinitiv.

Variables	**Refinitiv code**	**Type**
Financial year	TR.F.EBITDA.Date	Date
EBITDA Margin - %	TR.F.EBITDAMargPct	Ratio
Earnings before Interest Taxes Depreciation & Amortization	TR.F.EBITDA	Number
Total Liabilities	TR.F.TotLiab	Number
Total Current Assets	TR.F.TotCurrAssets	Number
Total Current Liabilities	TR.F.TotCurrLiab	Number
Inventories - Total	TR.F.InvntTot	Number

Variables	Refinitiv code	Type
Current Ratio	TR.CurrentRatio	Ratio
Quick Ratio	TR.QuickRatio	Ratio
Revenue from Business Activities - Total	TR.F.TotRevenue	Number
Total Assets	TR.F.TotAssets	Number
Labor & Related Expenses - Total	TR.F.LaborRelExpnTot	Number
CO2 Equivalent Emissions Total	TR.CO2EmissionTotal	Number
ESG Controversies Score	TR.TRESGCControversiesScore	Number
Turnover of Employees	TR.TurnoverEmployees	Number
Independent Board Members	TR.AnalyticIndepBoard	Number
Board Gender Diversity, Percent	TR.AnalyticBoardFemale	Number
Workforce Score	TR.TRESGWorkforceScore	Number
Human Rights Score	TR.TRESGHumanRightsScore	Number
Management Score	TR.TRESGManagementScore	Number
TRBC Industry Name	TR.TRBCIndustry	Category
TRBC Economic Sector Name	TR.TRBCEconomicSector	Category
TRBC Industry Group Name	TR.TRBCIndustryGroup	Category
Country of Headquarters	TR.HeadquartersCountry	Category

Table 1 : Description of all variables extracted from Refinitiv (source: Refinitiv)

The variables from Refinitiv are downloaded for each end-of-year date over 30 years from 1995 to 2025. Then this dataset is merged with the dataset containing all rating actions based on the condition that the current end-of-year date is within the consecutive rating action dates. The dates that fall during the non-rated periods are removed.

In the final comprehensive panel dataset, we have, for each rated company, its financial and ESG indicators for each year when the company is rated. A flag of downgrade is added if a downgrade event happens in the next cohort of the current observation year. For instance, if the downgrade happens on 30/08/2023, then the year 2022 will be flagged as 1.

There are in total 15365 observation points, from which we observe the following evolution of the number of all rated and downgraded observations.

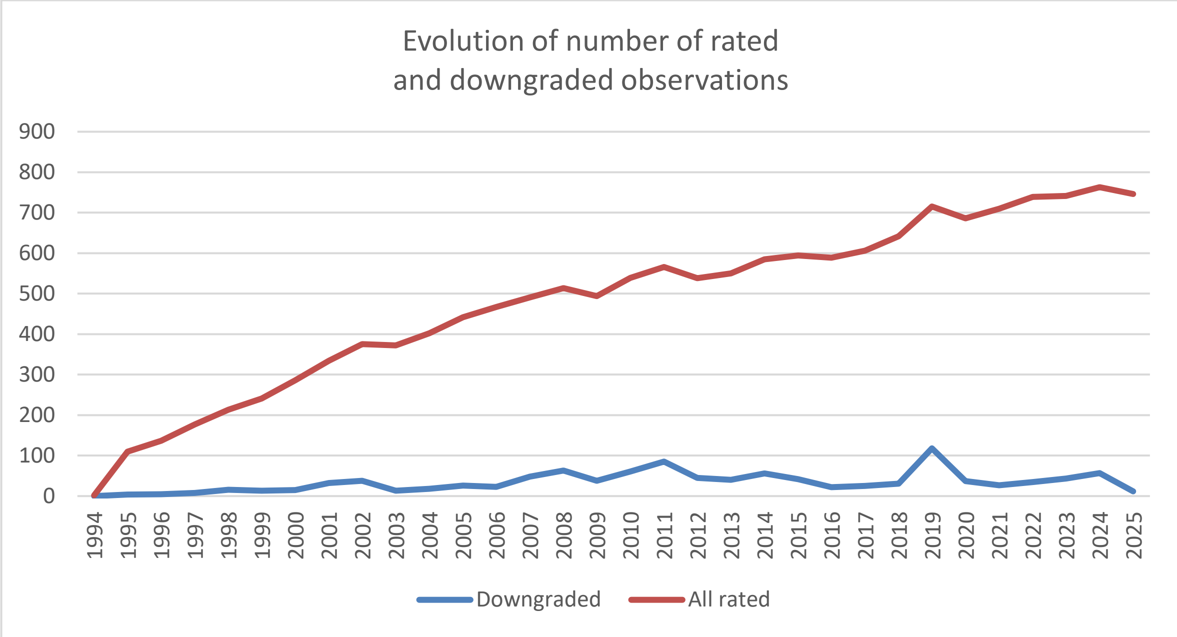


Figure 6 : Evolution of number of rated and downgraded observations

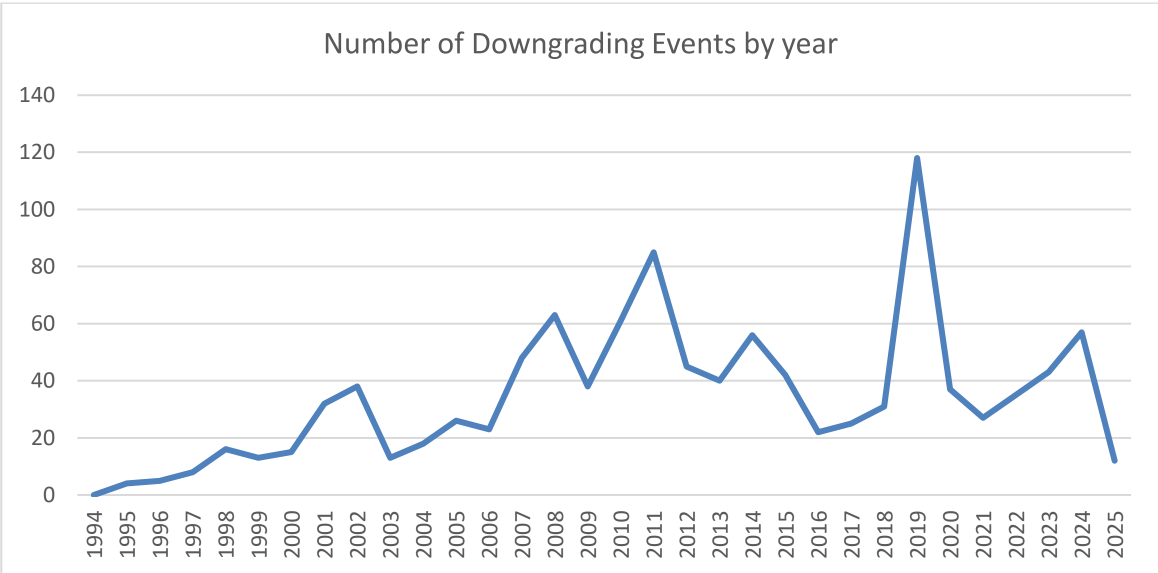


Figure 7 : Evolution of number of downgraded observations

As time evolves, the number of companies being rated steadily increases. Then by zooming into only the downgrade events, we observe a series that exhibits strong cyclical behavior, with downgrade activity concentrated in periods of economic and financial stress. From 1995 to 2012, the number of downgrade actions exhibits a broadly increasing pattern, until reaching peaks during major crisis periods. This phase is followed by a marked decline over the subsequent half-decade. During the pandemic period, downgrade activity rises sharply again, highlighting the strong concentration of downgrades during economic downturn periods.

5.2 Data treatment

It is worth mentioning that all companies do not have available data for a 30-year period. Therefore, we need to decide on the appropriate window of observation. To understand the data availability of each candidate variables, we have plotted the following distribution of the earliest date of data availability.

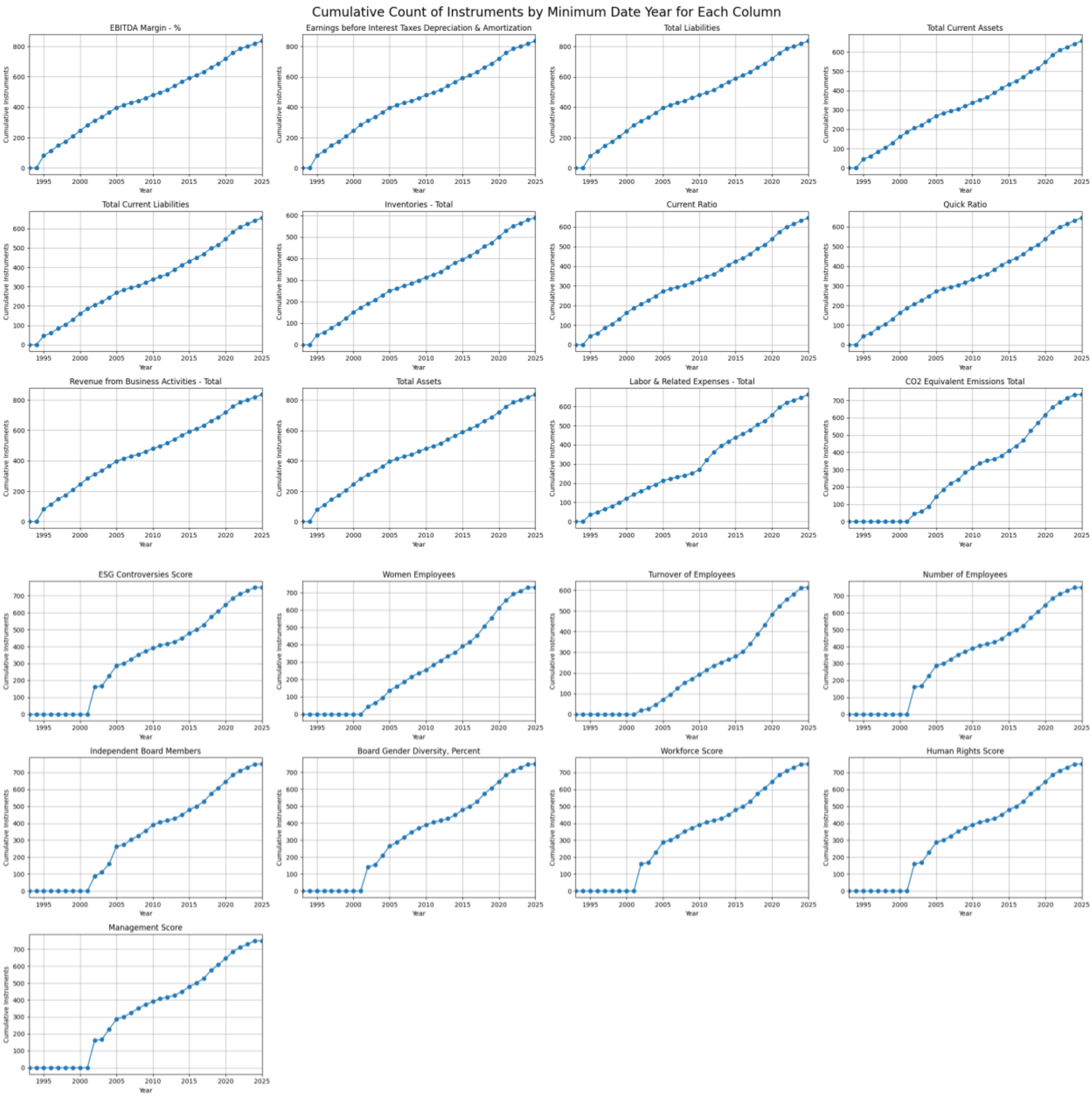


Figure 8 : Cumulative count of companies by year for each shortlisted variables

Indeed, the availability of financial data is much higher than the ESG variables. For example, EBITDA margin is available well before 2015 for most firms; while for CO2 emissions, there is no data before 2002 and meaningful coverage happens only after 2010. The choice of the observation window shall ensure sufficient number of companies, years, ESG variations, and of downgrading events. Therefore, in order to balance sample size and ESG data availability, the main analysis focuses on 2010-2024 (provided that the ESG variables are not yet available in 2025). A robustness analysis is conducted later using a shorter period from 2015 to 2024 with higher ESG coverage.

After filtering on the observation window, 9563 observation points remain. As we are dealing with public companies, or financial indicators of levels, the missing values are highly probably due to reporting gaps. We choose to interpolate maximum 2 consecutive missing values between two available observation points This approach assumes that financial variables evolve smoothly over short horizons and avoids introducing artificial volatility. Larger gaps and missing values at the beginning or end of a firm's presence in the dataset are not extrapolated, as such treatment would rely on stronger assumptions and could distort the underlying data.

For ESG indicators, we only treat CO2 emissions and board independence using forward filling for up to two consecutive years to account for unstructured disclosure gaps and persistence of ESG characteristics. Backward filling is not applied in order to avoid introducing look-ahead bias since future information would not have been available by the time of prediction. For these two variables, we create missing indicators to capture potential choice of non-disclosure. This is because, different from financial variables whose missing values are more frequently due to reporting gaps or database issues, missing ESG variables may reflect disclosure, transparency, and governance quality. As such, we do not risk having a smaller sample when the observation points with missing ESG data are directly removed.

Nevertheless, we remove the rows with key financial variables (EBITDA, Total revenue, Total debt, Total Assets) missing even after the above treatment given that this information is essential and the volume of missing values is not material. This results in a total of 9406 observations, concerning 801 distinct companies.

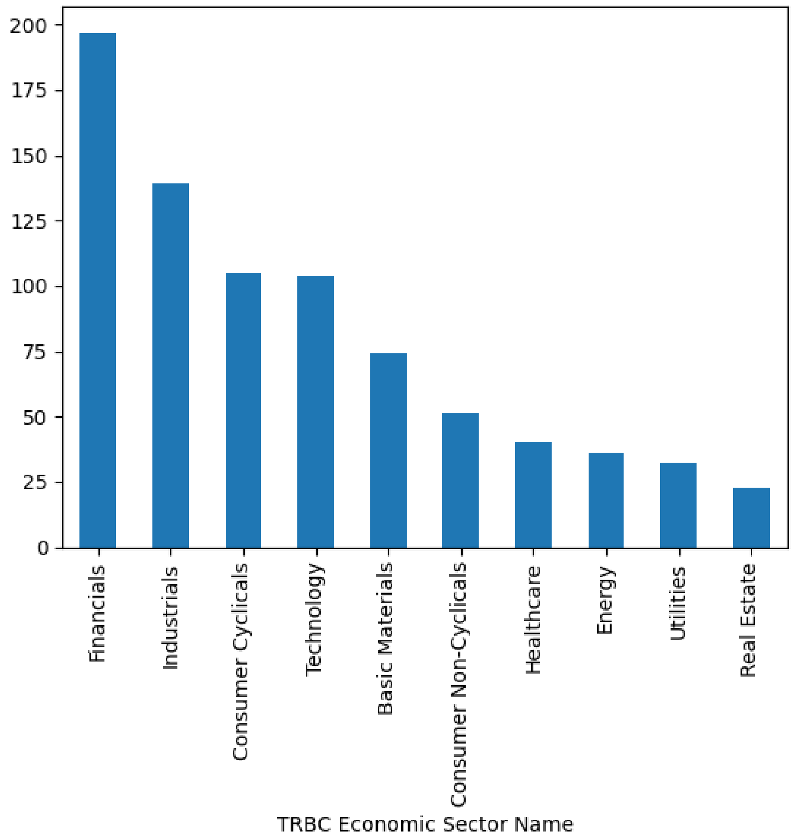


Figure 9 : Distribution of companies by economic sector in the modeling sample

The distribution of companies in the sample exhibits significant variation across economic sectors, with nearly 25% of companies concentrating on the financial industry. Industrials and consumer cyclicals also represent substantial portions of the sample. Technology follows closely. In contrast, sectors such as basic materials, consumer non-cyclicals, healthcare, energy, utilities, and real estate are less representative.

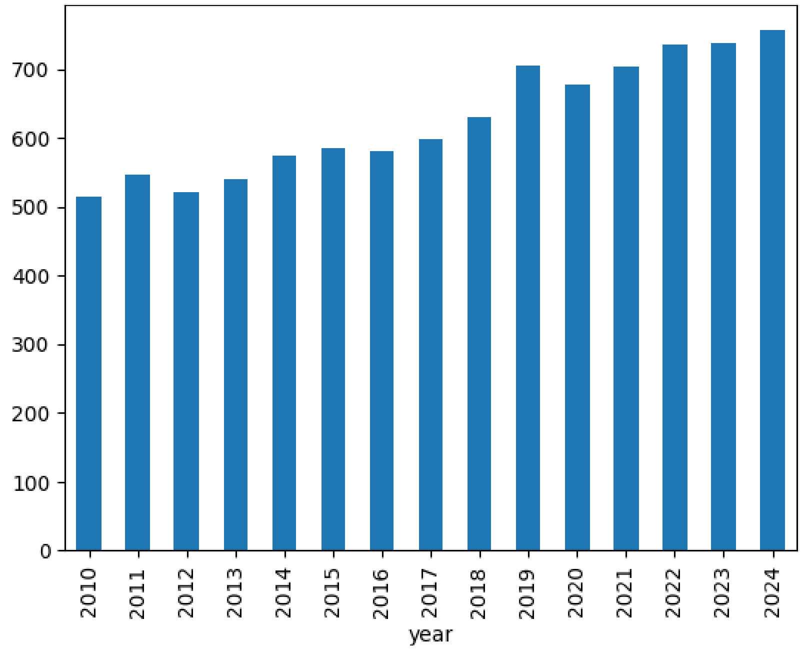


Figure 10 : Distribution of companies by year in the modeling sample

The above figure displays the distribution of the 9406 observations across different years. Number of observations evenly increase over time, reflecting an increasing availability of data. The steady rise in observations is driven by the fact that data collection or reporting has become more frequent and accessible over the years and that the number of total companies being rated is increasing over time.

Then we check the percentages of missing values of the remaining variables.

Variables	% Missing Values
Turnover of Employees	45.24
Inventories_f	33.97
Labor_intensity	28.86
Labor_cost_f	28.83
Quick Ratio_f	26.58
Total Current Assets_f	26.31
Total Current Liabilities_f	26.29

Variables	% Missing Values
Current Ratio_f	26.28
CO2 Equivalent Emissions Total_f	19.23
ESG Controversies Score	11.44
Human Rights Score	11.44
Workforce Score	11.44
Management Score	11.44
Independent Board Members_f	11.08
Board Gender Diversity, Percent_f	11.08

Table 2 : Percentage of missing values per shortlisted variables

Note that the variables ended by "_f" are after the above-mentioned missing value treatment. The missing value of employee turnover is too significant to be treated appropriately. Therefore, we have decided to drop this variable. Instead, we switch to ESG controversary score, human right score, workforce score, and management score to reflect the social aspects of a company.

The liquidity indicators (i.e current ratio and quick ratio) have an important percentage of missing values. We have further forward filled up to 2 missing years. However, there are still around 24% missing values for current ratio and quick ratio. In this case, the missing value is not technical but structural as it is not a data gap anymore but firm's selection and reporting structure. Therefore, we have decided to fill the rest of missing values by the company's median if there are less than 30% missing values at firm level, otherwise we fill them with the year median. We also added another column indicating the presence of missing value before filling the medians. As current ratio has slightly higher availability than quick ratio and these two ratios both represent liquidity, we use mainly the current ratio variable for the regression model.

Furthermore, we fill the missing "CO2 Equivalent Emissions Total_f", "Independent Board Members_f" and all the score-related variables by firm-level median if the firm has less than 30% missing value in this variable, otherwise we fill them by the year median to reflect reporting standards at that time and general ESG level shifts. Identically, missing value indicators are created for ESG variables.

Then we performed descriptive analysis for each variable. We have seen extreme values in the variable EBITDA margin sourced from Refinitiv. For example, the minimum value is -14266.7%, the maximum value is 1421.5%, and the standard deviation is 192%.

Parameters	EBITDA_margin_f
count	9.402
mean	20
std	192

Parameters	EBITDA_margin_f
min	-14.267
25%	12
50%	20
75%	34
max	1.421

Table 3 : Descriptive statistics of variable "EBITDA_margin_f"

After investigation, the two maximum values (197% and 1421%) are a result of data treatment (using the value of EBITDA divided by total revenue, which are both negative). In Refinitiv, these values are set to "NaN". The extreme negative value of EBITDA margin can happen when operating loss is too high compared to relatively smaller negative revenue. This is actually a very strong signal of creditworthiness deterioration. We have indeed observed some downgrades when EBITDA margin is highly negative. However, these extreme values may distort the regression model and techniques such as winsorization or cap and floors may cause a loss of important signal of creditworthiness deterioration. Therefore, instead of using EBITDA margin, we decided to switch to the ratio between EBITDA and Total Asset to represent profitability. This choice is also supported by literature. Shumway has included this variable in predicting bankruptcy risk (Shumway, 2001).

The same issue applies to labor intensity (labor cost over total revenues) and carbon intensity (carbon emissions over total revenues). Under normal situation, higher carbon intensity represents higher transition risk. With negative revenues, the ratio may become negative, but it does not necessarily represent lower transition risk. Therefore, we have decided not to include them in the regression. Instead, we use directly the log values of carbon emissions.

To reflect size effect, we take logs on total assets and total revenues. Note that when revenues are negative, its log value is set to 0. With the presence of negative revenues, we prefer using the log of total assets as a proxy of company's size effect.

Below is a descriptive analysis of the key explanatory variables. Each variable is analyzed in terms of its count, mean, standard deviation, minimum and maximum values, as well as different quantiles.

Parameters	flg_dg	log_asset	log_revenue	EBITDA_Asset	leverage
count	9.406	9.406	9.406	9.406	9.406
mean	0,07	24,28	22,99	8,92	0,73
std	0,26	2,14	2,01	7,60	0,20
min	0,00	16,41	0,00	-56,25	0,06
25%	0,00	22,83	21,76	2,04	0,58

50%	0,00	24,02	23,02	8,71	0,73
75%	0,00	25,48	24,20	13,07	0,89
max	1	33,32	30,70	62,75	2,72
Parameters	**Independent Board Members_f_filled**	**Board_missing**	**Board Gender Diversity, Percent_f_filled**	**UR_f**	**GDP_growth_f**
count	9.406	9.406	9.406	9.406	9.406
mean	66,94	0,12	27,07	6,73	1,87
std	22,74	0,32	13,27	3,82	3,26
min	0,00	0,00	0,00	0,10	-10,94
25%	53,85	0,00	18,18	4,11	0,98
50%	70,59	0,00	27,27	5,52	2,00
75%	84,62	0,00	36,36	8,07	2,86
max	100	100	75	28,47	24,62
Parameters	**log_CO2**	**Human Rights Score_filled**	**ESG Controversies Score_filled**	**Workforce Score_filled**	**Management Score_filled**
count	9.406	9.406	9.406	9.406	9.406
mean	12,65	61,51	80,55	75,45	64,36
std	2,31	30,86	31,30	20,63	25,39
min	1,95	0,00	0,31	0,45	0,10
25%	11,40	43,25	67,86	64,96	47,98
50%	12,80	70,42	100	80,58	67,87
75%	13,86	87,74	100	91,46	85,31
max	19,09	100	100	100	100

Table 4 : Descriptive statistics of all shortlisted and imputed variable

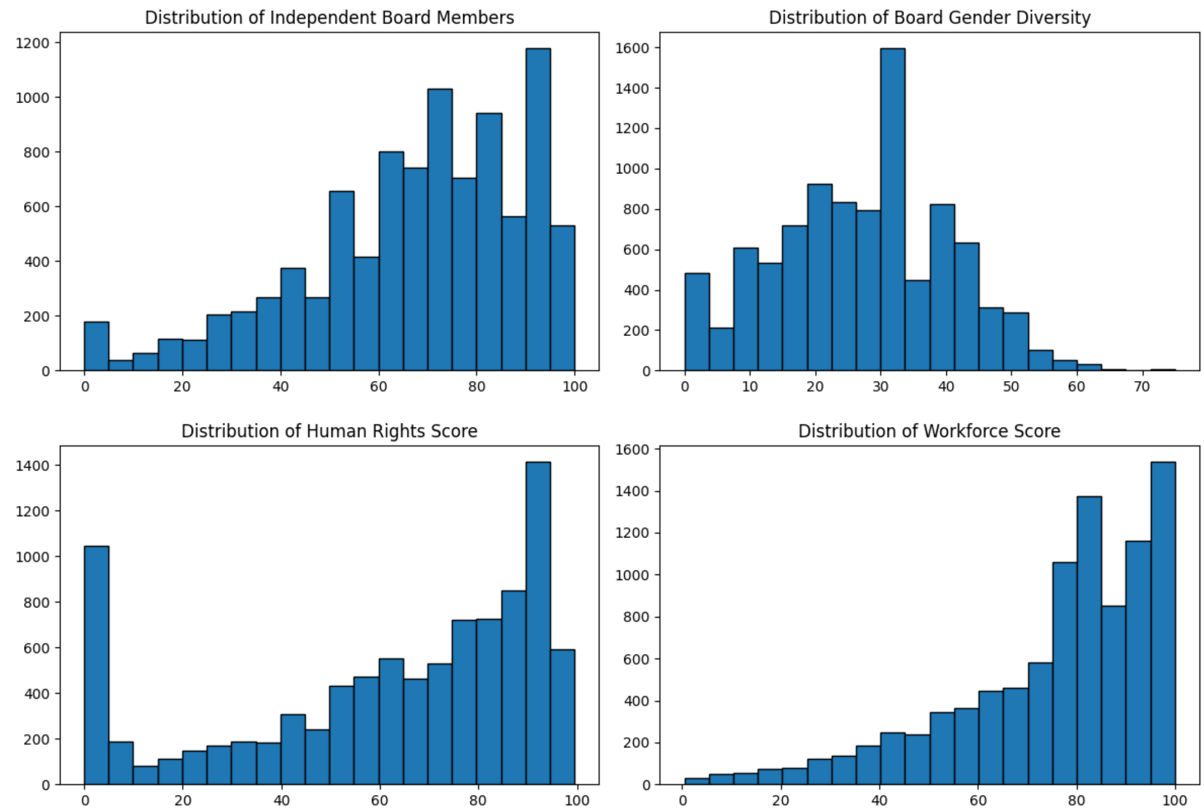
The dependent variable ("flg_dg") is a downgrade indicator capturing whether a firm experiences a credit rating downgrade in a given year. Its low mean (0.07) and standard deviation (0.26) confirm that downgrade events are relatively rare in the sample, which is consistent with the literature on credit risk. Firm size, proxied by "log_asset" and "log_revenue", shows very close means and medians, which displays symmetric distribution. The latter has a wider range than the former showing a more volatile nature in total revenues. Profitability is captured by the EBITDA-to-asset ratio, which shows a moderate variability (standard deviation of 7.60) and a wide range from strongly negative to positive values. This dispersion suggests significant differences in operating performance across firms. Leverage has a relatively low dispersion.

Turning to corporate governance variables, board independence has a mean of 66.94% and a median of 70.59%, indicating that most firms maintain a majority of independent board members. The standard deviation of 22.74% reveals essential variation among firms. Similarly, board gender diversity has a mean of 27% women on board and a moderate variation.

Most of the current ratio values cluster around small numbers with notable outliers at the higher end, suggesting a slightly right-skewed distribution. The outlier 17.13 is possibly driven by firms holding unusually high levels of current assets relative to liabilities. As for unemployment rate, the relatively high standard deviation (3.82) and wide range (0.10 to 28.47) highlight substantial cross-country and temporal variation. The extreme value of 28.47 probably corresponds to crisis periods. Then the GDP growth variable also shows significant economic fluctuations across countries and over time.

The environmental dimension is proxied by the log value of CO_2 emissions. The standard deviation of 2.31 reveals moderate dispersion across firms. The range from 1.95 to 19.09 suggests heterogeneity in emission levels, which is consistent with differences in industry exposure and production processes.

For the score-related variables, the descriptive statistics reported in the summary table may not be fully informative. Unlike continuous financial variables, these scores are often discrete, bounded, and may exhibit clustering or skewness driven by rating methodologies. As such, measures such as the mean or standard deviation may not adequately capture their underlying distribution. To better understand their patterns, histograms are drawn for each of the score-related variables.



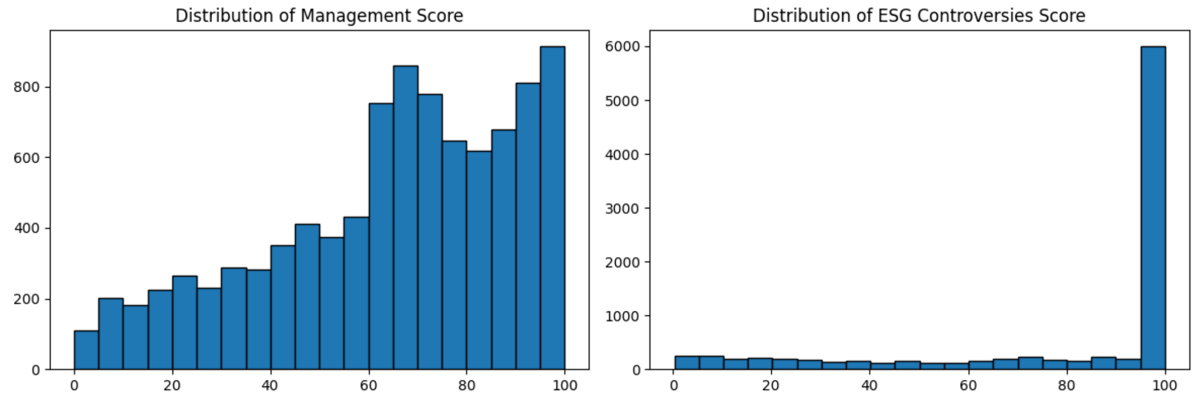


Figure 11 : Distribution analysis of score-like variables

We have observed that the maximum board independence could reach 100%. This is not deemed as data quality issue. According to Ben Levett, more than 30 companies in S&P 500 index have 100% board independence and the average board independence reaches 85% (Levett, 2025).

There is a large number of observation points in our sample having zero Human Right Score, which reflects a key limitation of Refinitiv ESG data: missing and zero values are indistinguishable (Ehlers, Elsenhuber, Jegarasasingam, & Jondeau, 2022). For other variables, the percentages of zero values are not material. For Board Independence and Board Gender Diversity, zero values are informative. Therefore, for the missing indicators of Human Right Score, Management score, and Workforce score will also consider the situation of zero scores and fill them with firm medians.

One may notice that the distribution of ESG Controversies scores highly skewed towards 100. This is because this variable is event-based. If there is no controversy, a 100 score is given. Thus, it is not deemed as a data quality issue.

The shortlisted variables used to fit in with the logit regression includes log of total assets, EBITDA asset ratio, leverage ratio, board independence, current ratio, unemployment rate, GDP growth, log of CO2 emissions, human right score, workforce scores, ESG controversies scores, and the missing value indicators on current ratio, carbon emissions, and score-related variables.

5.3 Variable selection

To understand the relationships between variables, we have computed the correlations between each combination of main variables (missing indicators are excluded and will be tested independently).

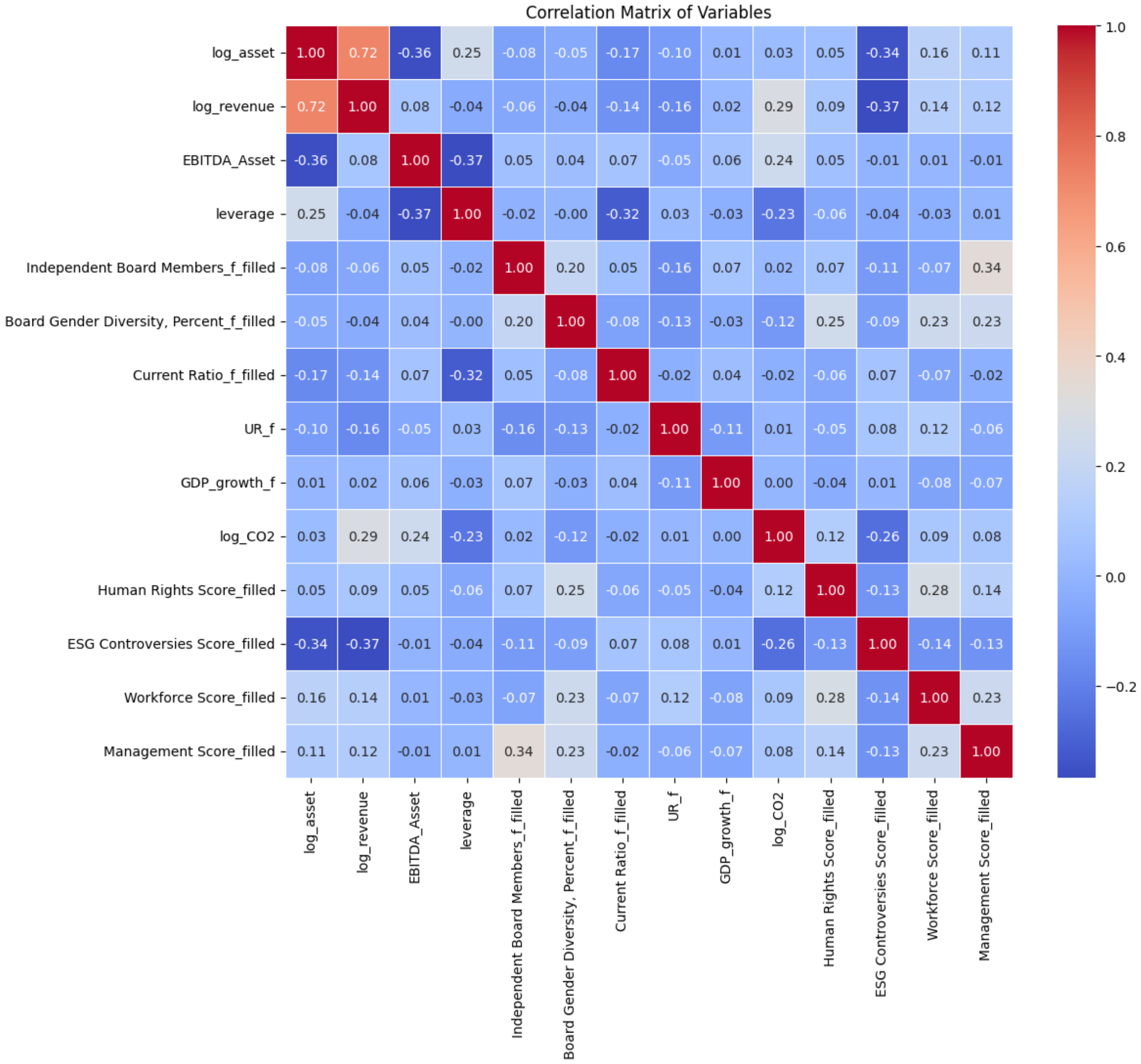


Figure 12 : Correlation matrix between key explanatory variables

From the above correlation matrix, it is not surprising to see a high correlation between the two size effects. We have also observed that moderate correlations of -0.36 between log of asset and EBITDA asset ratio and of -0.37 between leverage and EBITDA-to-asset ratio. The correlation between ESG Controversies score and size effect is around -0.35. To avoid multicollinearity, we have computed VIF for each control variable, to ensure validity of the baseline model presented in the section 7.

Variable	VIF
log_asset	35
EBITDA_Asset	2
leverage	18
Independent Board members_f_filled	10

Variable	VIF
Board_missing	1
Board Gender Diversity, Percent_f_filled	5
Current Ratio_f_filled	5
UR_f	4
GDP_growth_f	1

Table 5 : VIF of main variables

As a general practice, a VIF of less than 5 indicates a lack of multicollinearity among variables and a VIF greater than 10 indicates serious multicollinearity problems. With a value of 35, the log total asset shows extremely serious collinearity issues, which makes sense as EBITDA Asset ratio and leverage variables contain total assets in their computation. This is reasonable as both ratios are computed from total assets. We have also tested by including the variable log of total revenue, but the conclusion remains the same. To avoid spurious regression, we have decided to drop logs of total revenue and total assets.

In addition, we have created a new variable called “log_asset_resid” which is the residual after fitting an OLS model with EBITDA asset ratio and leverage ratio. Like this, we could still retain the size effect that is not captured by the other two ratios. The VIF results have been greatly improved even if leverage and board independent still show moderate collinearity.

Variable	VIF
log_asset_resid	1
EBITDA_Asset	2
leverage	8
Independent Board members_f_filled	8
Board_missing	1
Board Gender Diversity, Percent_f_filled	5
Current Ratio_f_filled	4
UR_f	4
GDP_growth_f	1

Table 6 : VIF of main variables after transforming the logarithm of total asset to residuals

We have also computed the correlation matrix with only missing values indicators.

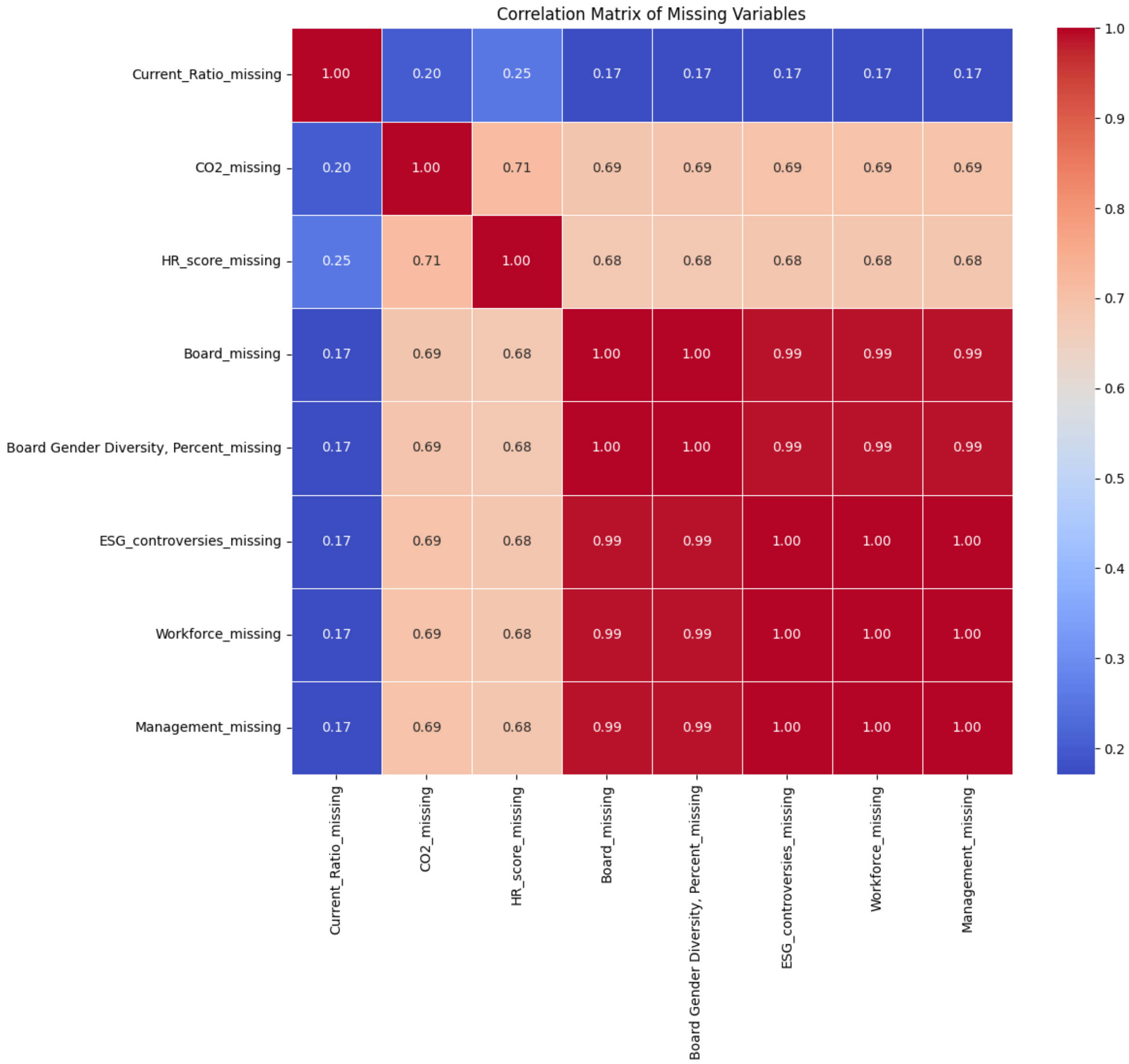


Figure 13 : Correlation matrix between missing value indicators

Among these missing indicators, those of human right score, board independence, ESG controversies, workforce score, and management score are highly correlated with a value of 100%. Therefore, in the regressions later, we will only include the missing indicator for board independence to avoid redundancy. Besides, there is no evidence of collinearity between the missing indicators of board independence and carbon emissions (VIF is around 2).

As a result, the final shortlisted variables include residual of log value of total assets, EBITDA asset ratio, leverage ratio, board independence, board gender diversity, current ratio, unemployment rate, GDP growth, log of CO2 emissions, human right scores, workforce scores, ESG controversies scores, management scores, and the missing value indicators on current ratio, carbon emissions, and board independence.

7. Results

This section presents a sequence of empirical models designed to assess the incremental explanatory power of the selected ESG variables in predicting the likelihood of downgrade events.

Before fitting models, we listed the expected signs of the coefficients of each explanatory variable and provided theoretical rationale underlying each hypothesis:

Variables	Expected sign	Remark
Residual of log of total assets	Negative	This is based on the hypothesis that larger companies may have lower downgrading risks.
EBITDA to Asset	Negative	This is based on the hypothesis that more financially profitable companies have lower downgrading risks.
leverage	Positive	This is based on the hypothesis that a highly leverage firm is more likely to face liquidity and refinancing issues, and thus higher downgrading risk.
Board independence	Unknown	Higher board independence percentage will alleviate the agency problem. However, too many independent board members may lead to a lack of knowledge and expertise in the business. Please refer to the hypothesis H3a.
Board gender diversity	Unknown	No clear empirical evidence shows a one-direction relationship between this variable and the target variable. Please refer to the hypothesis H3b.
Current ratio	Negative	This is based on the hypothesis that liquidity effects are negatively associated with the probability of downgrading events.
Unemployment rate	Positive	Rising unemployment rate can act as a signal of economic downturn and may be linked to higher downgrade risk.
GDP growth	Negative	Rising GDP growth reflects a healthy macroeconomic environment and may be linked to lower downgrade risk.

Variables	Expected sign	Remark
CO2 emission	Positive	Carbon intensive companies are more exposed to transition risk, which could lead to higher credit risk. Please refer to the hypothesis H1.
Human right score	Negative	Stronger labor practices, employee relations, and workplace safety are less likely to experience credit risk downgrades. Please refer to the hypothesis H2b.
Workforce score	Negative	Firms with better human right performance are less likely to face credit risk downgrades. Please refer to the hypothesis H2a.
Management score	Negative	Firms with better management quality are less likely to be downgraded. Please refer to the hypothesis H3c.
ESG controversies score	Positive	Firm is involved more in ESG controversies are more likely to be downgraded. Please refer to the hypothesis H2c.

Table 7 : Expected signs of final shortlisted explanatory variables

Overall, the expected directions are consistent with economic intuition and prior research. Firm-specific financial metrics such as leverage are anticipated to increase downgrade risk, while strong probability and liquidity indicators are expected to mitigate it. At the governance level, we cannot conclude a clear relationship between variables such as board independence and board diversity and downgrading risk as prior research contributes to various opinions. For ESG-related factors, better performance is generally expected to reduce credit risk. Finally, as stronger economic conditions reduce and adverse conditions increase downgrading risk, macroeconomic conditions are incorporated through variables such as GDP growth and unemployment rate, which capture the cyclical environment where firms operate.

7.1 Model 1: Baseline regression

The baseline model includes only firm-level financial controls and macro-economic variables, together with the industry and year fixed effects. The sample is an unbalanced panel of data as not all firms have the same number of observations. The model could be expressed as follows, with p equal to the probability of downgrade event:

$$\ln\left(\frac{p}{1-p}\right) = constant + \beta_1 * Control_{i,t-1} + \beta_2 * Macro_{i,t-1} +$$

$$\beta_3 * Industry\ FE + \beta_4 * Year\ FE + \varepsilon_{i,t}$$

As discussed in section 3, ESG risks could impact firm's creditworthiness directly and indirectly through influencing firms' financial fundamentals. Therefore, we start by fitting a model that contains only traditional financial measures together with other control variables. Then in other models, by adding ESG variables, we could capture the effects that these variables may have in addition to their indirect impact through the channels of financial fundamentals.

The result of the logit regression of model 1 is as follows.

Model1: Logit Regression Results

Dep. Variable:	flg_dg	No. Observations:	9406
Model:	Logit	Df Residuals:	9372
Method:	MLE	Df Model:	33
Date:		Pseudo R-squ.:	0.07893
Time:		Log-Likelihood:	-2290.6
converged:	True	LL-Null:	-2486.9
Covariance Type:	cluster	LLR p-value:	4,11E-63

Table 8 : Model summary of Model 1

As discussed in section 5, cluster correlation method is applied when training the logit regression model. The log-likelihood ratio p-value is close to 0 meaning that we could reject the underlying hypothesis that all coefficients are equal to 0. The model is thus jointly significant. Pseudo R-squared is a refinement to R-squared and is applicable to logit regressions. The baseline model exhibits a pseudo-R square of 7.89%, which is within a reasonable range observed in logit regression models of credit risk. McFadden has stated that a pseudo-R square between 0.2 to 0.4 represent an "excellent" model fit (McFadden, 1979). Hemmert et al. who performed a meta-analysis of 274 published logistic regression models with different types of pseudo-R squares, found that their average values range from 0.108 to 0.376 (Hemmert, Schons, & Schimmelpfennig, 2016). Thus, the value of 7.89% suggests modest explanatory power given the binary nature and rare occurrence of downgrading events.

The table below displays the coefficients and p-values of each explanatory variable included in the regression model.

	Coef.	Std.Err.	z	P>|z|	[0.025	0.975]
const	-2,419	0,354	-6,835	0,000	-3,112	-1,725
log_asset_resid	-0,035	0,024	-1,487	0,137	-0,082	0,011
EBITDA_Asset	-0,042	0,012	-3,587	0,000	-0,066	-0,019
leverage	1,612	0,238	6,769	0,000	1,145	2,079
Independent Board Members_f_filled	-0,003	0,002	-1,717	0,086	-0,006	0,000
Board Gender Diversity, Percent_f_filled	-0,009	0,004	-2,398	0,016	-0,016	-0,002
Board_missing	0,043	0,130	0,328	0,743	-0,211	0,297
Current Ratio_f_filled	0,018	0,072	0,252	0,801	-0,123	0,160
Current_Ratio_missing	-0,396	0,229	-1,735	0,083	-0,844	0,051
UR_f	0,034	0,009	3,609	0,000	0,016	0,053
GDP_growth_f	-0,089	0,033	-2,743	0,006	-0,153	-0,026

Table 9 : Coefficients and p-values of main control variables

From the above results, we see that the variables EBITDA-to-asset ratio, leverage, board independence, board gender diversity, unemployment rate and GDP growth are significantly correlated with the logit of downgrade events with a threshold of p values smaller than 10%. The signs of the coefficients are also what we expect. In addition, the coefficient of leverage is much stronger than that of the EBITDA-to-asset ratio, indicating that capital structure dominates profitability in explaining downgrading risk.

Board independence and board gender diversity are both negatively correlated to the likelihood of downgrading, suggesting that higher board independence and more women on board (Note that the maximum value is 75% in our sample, and the 75% quantile is only 36.36%) would somehow mitigate the downgrading risk.

In contrast, current ratio is not significant. This result might be due to the 27% missing values in the raw data and the fact that the liquidity effect is weaker compared to leverage and profitability. On the contrary, the indicator of missing current ratio is significant. This may indicate that missing liquidity data is not random. However, the negative coefficient seems against intuition as missing current ratio is often associated with smaller and less complex firms whose downgrading risks tend to be higher. Such negative relationships likely reflect differences in firm visibility rather than true economic effects. For example, firms with complete financials are more transparent and more actively scrutinized by rated rating agencies. In other words, firms with missing liquidity data may be less exposed to active credit monitoring and thus fewer rating actions are observed.

What's more, both unemployment rate and GDP growth are highly significant, and the signs align with our hypotheses, which further confirms the validity of our model setting.

	Coef.	Std.Err.	z	P>|z|	[0.025	0.975]
Sector_Consumer Cyclicals	0,013	0,180	0,071	0,943	-0,340	0,366
Sector_Consumer Non-Cyclicals	-0,356	0,226	-1,573	0,116	-0,800	0,088
Sector_Energy	0,552	0,202	2,733	0,006	0,156	0,947
Sector_Financials	-0,454	0,286	-1,585	0,113	-1,015	0,108
Sector_Healthcare	0,174	0,236	0,736	0,462	-0,289	0,637
Sector_Industrials	-0,328	0,176	-1,864	0,062	-0,674	0,017
Sector_Real Estate	-0,046	0,284	-0,161	0,872	-0,602	0,511
Sector_Technology	0,008	0,193	0,041	0,967	-0,370	0,386
Sector_Utilities	0,050	0,198	0,251	0,801	-0,338	0,437

Table 10 : Coefficients and p-values of industry FE variables

Note that it is normal for some industry and year fixed effect variables to be insignificant. Their presence is aimed at avoiding bias introduced by unobservable characteristics within an industry or a given year. The industry fixed-effect coefficients capture differences in the conditional log-odds of downgrade across industries relative to the reference industry (i.e. Basic materials). Significance means that the conditional baseline log-odds of downgrade is statistically different from the reference. From the above results, only energy and industrial economic sectors are statistically different from the reference industry in terms of credit risk profile. Energy sector has a high-downgrade risk profile as it is exposed to volatile commodity prices and is usually highly leveraged and capital intensive. This sector is also highly exposed to environmental transition risk as growing scrutiny on carbon emission can lead to higher financing costs or stranded asset risks. The positive coefficient (0.552) is consistent with this interpretation. In contrast, although firms in the Industrials sector (e.g. aerospace, machinery, and transport companies) are characterized by high fixed costs and sensitivity to shifts in demand, the negative coefficient (–0.328) indicates that, conditional on controls, Industrials are less exposed to downgrade risk than firms in the Basic Materials sector. One possible explanation is that Basic Materials firms face more direct exposure to global commodity price shocks, leading to higher earnings volatility, whereas Industrials benefit from more diversified revenues and longer-term contracts (especially for aerospace and transport companies) that help stabilize cash flows during downturns.

For the rest of sector indicators, their insignificance relative to the reference industry is also justifiable. Consumer Cyclicals companies such as auto, retail, and hospitality are sensitive to business cycles compared to Basic Materials sensitive to commodity cycles. While Basic Materials firms are more capital-intensive and can benefit from tangible assets as collateral, they are also subject to commodity price volatility. In contrast, Consumer Cyclical firms have greater operational and cost flexibility, which may partially offset their higher sensitivity to demand fluctuations over the business cycle. These opposing forces may partially cancel each other out and an aggregation of them finally leads to similar downgrade risk profiles across the two sectors.

In terms of Financials industry, since 2010 banks are much more regulated. Basic Materials are mainly exposed to commodity price risk just like Financials to credit risk. They may imply average out to similar baseline downgrade odds in our sample. For Consumer Non-cyclicals such as food and beverage companies, cash flows are more predictable thanks to stable demand across cycles. Besides, the downgrade risk of the Healthcare and Technology sectors, generally having lower risks, is not significantly lower than Basic Materials either. This suggests that Basic Materials companies in our sample are not very risky probably because they are large, diversified, and even investment-grade companies. Finally, the Utilities and Real Estate sectors are both capital intensive but have predictable cash flows. Basic Materials are more cyclical but may have similar average default risk over a long panel.

	Coef.	Std.Err.	z	P>|z|	[0.025	0.975]
year_2011	0,290	0,183	1,586	0,113	-0,068	0,649
year_2012	-0,614	0,237	-2,585	0,010	-1,079	-0,148
year_2013	-0,585	0,242	-2,418	0,016	-1,059	-0,111
year_2014	-0,205	0,217	-0,942	0,346	-0,630	0,221
year_2015	-0,474	0,238	-1,994	0,046	-0,940	-0,008
year_2016	-1,393	0,304	-4,587	0,000	-1,988	-0,798
year_2017	-0,935	0,260	-3,594	0,000	-1,445	-0,425
year_2018	-0,746	0,253	-2,944	0,003	-1,242	-0,249
year_2019	0,611	0,202	3,023	0,003	0,215	1,008
year_2020	-1,471	0,377	-3,907	0,000	-2,210	-0,733
year_2021	-0,679	0,310	-2,190	0,029	-1,287	-0,071
year_2022	-0,675	0,264	-2,554	0,011	-1,192	-0,157
year_2023	-0,623	0,263	-2,372	0,018	-1,137	-0,108
year_2024	-0,295	0,250	-1,180	0,238	-0,786	0,195

Table 11 : Coefficients and p-values of year FE variables

The year fixed effects in the logit model, with 2010 as the reference year, reveal meaningful economic patterns in credit-cycle dynamics and the environment in which rating agencies operate between 2010 and 2024. A significant coefficient indicates that the controlled baseline downgrade risk in that year is statistically different from the post-GFC baseline of 2010, with the sign indicates whether the risk is higher (positive) or lower (negative). The results show that only two years (2011 and 2019) have positive and significant coefficients, meaning they exhibited statistically higher downgrade risk than 2010. The positive coefficient for 2011 captures the European sovereign debt crisis. During 2011-2012, concerns over Greek, Irish, Portuguese, Spanish, and Italian sovereign debt spilled over into corporate credit markets. Rating agencies downgraded not only sovereigns but also corporates reliant on domestic banking systems. This interpretation is supported by Kladakis and Skouralis who document the distribution of credit rating changes between 2006 and 2020 from three rating agencies (Moody's, S&P, and Fitch). Their evidence shows an elevated occurrence of downgrades in 2011 and 2012 (Kladakis & Skouralis, 2024).

In our sample, there is a surge in the percentage of downgrading events in 2019, which is reflected in the positive coefficient of the 2019-year effect in the baseline model. Note that this result is not consistent with distribution drawn in the paper of Kladakis and Skouralis mainly because the scope is different. The latter focus on worldwide and include sovereign ratings. Another reason is because a minority of companies have experienced multiple downgrades during this year. Provided that firm-level financial and ESG controls are lagged relative to each rating action, downgrade events are modeled at the event level, reflecting the conditional likelihood of a downgrade given prior firm characteristic. We deem that each event corresponds to distinct rating decisions and decide to keep multiple downgrades within a year in the modeling sample.

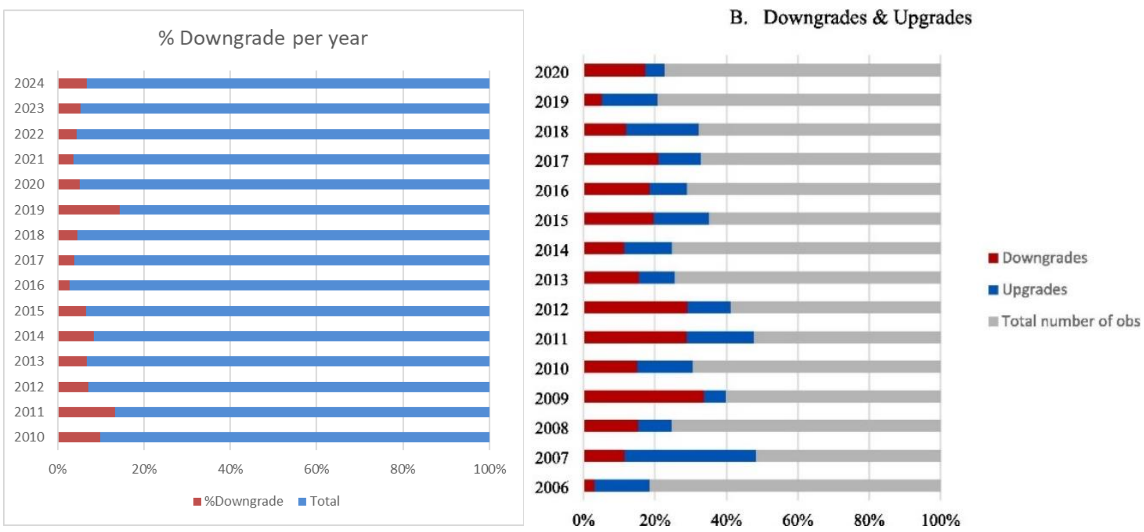


Figure 14 : Distributions of downgrades per year between our sample (left) and the sample from (Kladakis & Skouralis, 2023)

All the other years from 2011 to 2023, with the only exception of 2014, have significant but negative coefficients, indicating that downgrade risk was consistently lower than the 2010 baseline. This finding suggests that, for European public companies in our sample, the peak of the sovereign debt crisis was followed by a period of gradually improving economy with more stable credit conditions. Years such as 2012-2014 saw the gradual stabilization of the Eurozone through mechanisms such as the European Stability Mechanism and progressive bank recovery. During 2015 and 2016, the European Central Bank launched quantitative easing, which reduced credit spreads and supported corporate financial performance. The COVID-19 period from 2020 to 2022 produced negative coefficients, probably due to the massive policy response in Europe such as NextGenerationEU, a temporary recovery instrument of more than €800 billion to help repair the immediate economic and social damage (European Commission, s.d.). EU also widely used moratoria on loan repayments suspending debt obligations to prevent defaults that would have occurred under normal conditions. Such measures may temporarily suppress downgrades. The subsequent period of rising interest rates (2022-2023) also yielded negative coefficients, despite the sharp monetary tightening cycle. This may reflect a delayed effect of credit risk rather than an absence of underlying stress.

The years of 2014 and 2024 are statistically indistinguishable from the 2010 reference year. The non-significance of 2014 suggests that, following the peak of the Eurozone sovereign debt crisis, downgrade risk had returned to a level comparable to the post-GFC baseline level. On the other hand, the non-significance of 2024 suggests a return to normal level similar to the baseline year after the pandemic, the energy crisis following Russia’s invasion of Ukraine, and the monetary tightening cycle introduce by ECB.

In summary, these year fixed effects confirm that, for European developed markets, the 2010-2024 period was a low-downgrade-risk environment relative to the post-crisis baseline, with exception of 2011 and 2019. 2014 and 2024 represent neutral reference points within our sample.

The specification of the baseline model is theoretically supported by extensive literature on corporate credit risk (e.g.(Altman, 1968), (Shumway, 2001), etc.). The results confirm that the chosen financial metrics behave as expected. No coefficient in the baseline model violates economic norms. Signs are consistent across various measures, and p-values are generally below 10% for the core financial and macro variables.

Given that the baseline model is both statistically sound and economically coherent, it serves as a benchmark against which to assess the incremental contribution of ESG variables. Any improvement in model fit, coefficient stability, or predictive performance after adding ESG factors could be interpreted as these ESG metrics having a direct effect to the target variable. The next section proceeds with this analysis.

7.2 Model 2: Carbon emissions

Having established a statistically sound and economically coherent baseline model based on financial and macroeconomic determinants of downgrade risk, we now move to the main question of this study: whether environmental variables contribute incremental predictive power beyond traditional credit risk factors. We will introduce the logarithm value of firms' CO2 emissions as an additional explanatory variable in the logit regression framework. The logarithm value can reduce the scale and account for potential skewness in the raw data.

The table and the content below illustrate the results of this augmented model, together with a comparison with the baseline specification to assess changes in model fit, coefficient stability, and predictive performance.

Model2: Logit Regression Results

Dep. Variable:	flg_dg	No. Observations:	9406
Model:	Logit	Df Residuals:	9370
Method:	MLE	Df Model:	35
Date:		Pseudo R-squ.:	0.08253
Time:		Log-Likelihood:	-2281.6
converged:	True	LL-Null:	-2486.9
Covariance Type:	cluster	LLR p-value:	1,33E-65

Table 12 : model summary adding CO2 related variables

Adding CO2 emissions to the baseline model produces a small but consistent improvement in model fit. The pseudo-R2 increases from 7.89% to 8.25%, while the log-likelihood improves from -2290.6 to -2281.6. Hence, the likelihood ratio test, calculated as the difference between the log-likelihood of the baseline and the augmented models times minus two, is equal to 18. Although the absolute increase in pseudo-R2 is modest, this magnitude is typical for the addition of a single variable in logit models with rare events, where an increase of 0.5 – 1.0 percentage points in pseudo-R2 values is considered as meaningful.

The following tables display the coefficients of all the variables in the baseline logit model as well as the newly added log value of CO2 emissions.

	Coef.	Std.Err.	z	P>|z|	[0.025	0.975]
const	-3,609	0,532	-6,777	0,000	-4,652	-2,565
log_asset_resid	-0,086	0,028	-3,054	0,002	-0,141	-0,031
EBITDA_Asset	-0,040	0,012	-3,270	0,001	-0,064	-0,016
leverage	1,528	0,244	6,259	0,000	1,050	2,007
Independent Board Members_f_filled	-0,004	0,002	-2,032	0,042	-0,007	0,000
Board Gender Diversity, Percent_f_filled	-0,009	0,004	-2,363	0,018	-0,016	-0,002
Board_missing	0,334	0,212	1,574	0,115	-0,082	0,750
Current Ratio_f_filled	0,041	0,072	0,571	0,568	-0,099	0,181
Current_Ratio_missing	-0,340	0,252	-1,348	0,178	-0,834	0,154
UR_f	0,032	0,010	3,347	0,001	0,013	0,051
GDP_growth_f	-0,088	0,033	-2,661	0,008	-0,153	-0,023
log_CO2	0,080	0,025	3,222	0,001	0,031	0,129
CO2_missing	-0,484	0,197	-2,449	0,014	-0,871	-0,097

Table 13 : Coefficients and p-values of main control variables and CO2 related variables

	dy/dx	std err	z	P>|z|	[0,025	0,975]
log_asset_resid	-0,0056	0,002	-3,032	0,002	-0,009	-0,002
EBITDA_Asset	-0,0026	0,001	-3,282	0,001	-0,004	-0,001
leverage	0,0992	0,016	6,233	0,000	0,068	0,130
Independent Board Members_f_filled	-0,0002	0,000	-2,037	0,042	-0,000	-8,61e-06
Board Gender Diversity, Percent_f_filled	-0,0006	0,000	-2,359	0,018	-0,001	-9,9e-05
Board_missing	0,0217	0,014	1,582	0,114	-0,005	0,049
Current Ratio_f_filled	0,0027	0,005	0,571	0,568	-0,006	0,012
Current_Ratio_missing	-0,0220	0,016	-1,350	0,177	-0,054	0,010
UR_f	0,0021	0,001	3,339	0,001	0,001	0,003
GDP_growth_f	-0,0057	0,002	-2,661	0,008	-0,010	-0,002
log_CO2	0,0052	0,002	3,221	0,001	0,002	0,008
CO2_missing	-0,0314	0,013	-2,458	0,014	-0,056	-0,006

Table 14 : Marginal effects of main control variables and CO2 related variables

By adding variables related to CO2 emissions, the coefficients and significance of EBITDA-to-asset ratio, leverage, and board independence, and board gender diversity as well as macro-economic variables remain stable. Both log value of CO2 emissions, and the missing value indicator of CO2 emissions are significant. The sign of the former is as expected. This result suggests that CO_2 emissions are not simply proxying for profitability, leverage, and governance quality, but instead contain incremental information relevant to downgrade risk. Additionally, the positive coefficient and significance on emissions leads to a rejection of our first hypothesis (H1) and is consistent with the view that carbon-intensive firms could face higher exposure to regulatory costs, transition risk, and reputational concerns. Carbon emissions may affect credit ratings not only through expected higher costs, but also by increasing uncertainty around future cash flows. Overall, this result provides direct evidence that environmental risk is priced into corporate credit ratings in European developed markets.

Since the coefficients of a logit regression model is hard to interpret, we have calculated the marginal effect of each variable chosen. As shown by Table 14, the marginal effect of the logarithm of CO2 emission is 0.52 percentage points. This means that if CO2 emission double (i.e. 100% increase), the likelihood of downgrades would increase by around 0.36% (equal to ln(2) times 0.52%) if everything else remains constant. Considering that the average predicted probability of downgrade of the baseline model is 7.4%, doubling CO2 emission would lead to a 4.9% relative increase from the baseline.

Next, the negative correlation between the missing value indicator of CO2 emissions and the target variable shows that firms with unreported CO2 emissions have a lower downgrade risk than similar firms who report emissions (after controlling for size, profitability, and leverage). This effect should not be interpreted as evidence of lower emissions but instead it reflects differences in regulatory exposure and reporting incentives. The negative coefficient on the missing indicator may reflect the situation that firms that do not report CO2 emissions tend to be smaller and/or operate in less environmentally sensitive sectors and hence face lower regulatory and investor scrutiny. Consequently, they may experience fewer downgrades for reasons unrelated to their true environmental performance. To address this concern, we re-estimate the model by keeping only the logarithm of CO2 emissions. The positive and significant coefficient on reported CO2 emissions variable remains, confirming that the main result is not driven by imputation (see the table below). Also, in the robustness analysis (section 8), we have tested the model by omitting all the observation points with missing ESG values. The logarithm of CO2 emissions still remains significant.

Note that this finding does not contradict evidence that voluntary emissions disclosure and credible transition commitments are associated with lower credit risk (Carbone, et al., 2021). Rather, while proactive disclosure may signal effective climate risk management, the absence of disclosure in our sample primarily reflects lower regulatory and investor visibility among relatively smaller and/or less environmentally exposed firms.

	Coef.	Std.Err.	z	P>|z|	[0.025	0.975]
const	-3,546	0,531	-6,682	0,000	-4,586	-2,506
log_asset_resid	-0,064	0,028	-2,325	0,020	-0,118	-0,010
EBITDA_Asset	-0,040	0,012	-3,344	0,001	-0,064	-0,017
leverage	1,577	0,243	6,495	0,000	1,101	2,052
Independent Board Members_f_filled	-0,004	0,002	-2,041	0,041	-0,007	0,000
Board Gender Diversity, Percent_f_filled	-0,008	0,004	-2,144	0,032	-0,015	-0,001
Board_missing	-0,030	0,137	-0,217	0,828	-0,298	0,239
Current Ratio_f_filled	0,034	0,072	0,473	0,636	-0,107	0,175
Current_Ratio_missing	-0,349	0,231	-1,514	0,130	-0,801	0,103
UR_f	0,035	0,010	3,666	0,000	0,016	0,054
GDP_growth_f	-0,091	0,033	-2,778	0,005	-0,154	-0,027
log_CO2	0,071	0,025	2,892	0,004	0,023	0,119

Table 15 : Coefficients and p-values of main control variables and only "log_CO2"

Back to the regression results with both "log_CO2" and "CO2_missing", compared to the baseline model, the residual of the log of total assets turns significant in this augmented model with an expected negative sign in its coefficient. Eskandari concluded that a positive relationship exists between firm size and the level of risk disclosure and that firm size can influence the risk disclosure level (Zadeh & Eskandari, 2012). In the baseline model, firm size may have absorbed variation related to both firms' likelihood of disclosing emissions and their underlying credit strength, as larger firms are more likely to report CO_2 emissions while also benefiting from lower downgrade risk. Once we explicitly control whether emissions are reported and how high emissions are, the remaining variation in size becomes significant. To put it differently, size contains both economic scale effect and tendency of ESG disclosure before adding CO2 variables, while the tendency of ESG disclosure part is now captured by the missing indicator and the emissions level. What remains of size, the pure economic scale effect (or other size-related credit advantages), now appears as significant.

Overall, this result is consistent with the view that larger firms benefit from economies of scale and greater market access, which buffer against downgrades.

	Coef.	Std.Err.	z	P>|z|	[0.025	0.975]
Sector_Consumer Cyclicals	0,239	0,191	1,255	0,210	-0,134	0,612
Sector_Consumer Non-Cyclicals	-0,191	0,231	-0,827	0,409	-0,645	0,262
Sector_Energy	0,649	0,209	3,107	0,002	0,240	1,059
Sector_Financials	-0,005	0,327	-0,014	0,989	-0,645	0,636
Sector_Healthcare	0,426	0,249	1,713	0,087	-0,062	0,914
Sector_Industrials	-0,108	0,185	-0,581	0,561	-0,471	0,256
Sector_Real Estate	0,297	0,312	0,954	0,340	-0,313	0,908
Sector_Technology	0,292	0,210	1,394	0,163	-0,119	0,703
Sector_Utilities	0,085	0,198	0,427	0,670	-0,304	0,473

Table 16 : Coefficients and p-values of industry FE variables and only "log_CO2"

	Coef.	Std.Err.	z	P>|z|	[0.025	0.975]
year_2011	0,290	0,184	1,577	0,115	-0,071	0,651
year_2012	-0,612	0,238	-2,570	0,010	-1,079	-0,145
year_2013	-0,564	0,242	-2,327	0,020	-1,040	-0,089
year_2014	-0,192	0,216	-0,892	0,373	-0,615	0,231
year_2015	-0,461	0,238	-1,941	0,052	-0,927	0,004
year_2016	-1,375	0,304	-4,529	0,000	-1,971	-0,780
year_2017	-0,927	0,260	-3,562	0,000	-1,437	-0,417
year_2018	-0,744	0,253	-2,935	0,003	-1,240	-0,247
year_2019	0,611	0,204	2,997	0,003	0,212	1,011
year_2020	-1,459	0,381	-3,831	0,000	-2,205	-0,712
year_2021	-0,661	0,313	-2,113	0,035	-1,274	-0,048
year_2022	-0,653	0,267	-2,450	0,014	-1,176	-0,131
year_2023	-0,592	0,262	-2,262	0,024	-1,106	-0,079
year_2024	-0,252	0,250	-1,007	0,314	-0,742	0,238

Table 17 : Coefficients and p-values of year FE variables and only "log_CO2"

The inclusion of CO_2 emissions also leads to economically meaningful changes in the industry fixed effects for Healthcare and Industrials. In the baseline model, the Healthcare sector exhibits a positive but statistically insignificant coefficient, which indicates downgrade risk broadly comparable to that of the Basic Materials sector. However, after considering emission levels and disclosure status, the Healthcare coefficient increases and becomes statistically significant. Such shift implies that Healthcare's relatively low emission profile correlated with lower observed downgrade risk. Also, once this environmental dimension is kept constant, the sector exhibits a higher baseline downgrade risk relative to Basic Materials to reflect potential regulatory uncertainty, patent cliffs, and litigation exposure.

On the other hand, Industrials display a negative and statistically significant coefficient in the baseline model, indicating lower downgrade risk than Basic Materials. After adding the CO_2-related variables, the magnitude of this negative coefficient declines greatly and becomes statistically insignificant. This may indicate that part of Industrials' low downgrade risk in the baseline model was associated with their emission profile. Hence, we deem that these changes underscore the importance of including environmental variables to avoid bias when comparing industry credit risks.

In contrast to the industry fixed effects, we have observed stability in the parameters on year fixed effects relative to the baseline specification. Such stability shows that the effect of CO_2 emissions does not capture cyclical or time-varying factors but primarily persistent firm-level characteristics independent from these factors.

To summarize this specification, we have observed that on top of the traditional control variables, firm-level carbon exposure and emissions reporting status provide incremental information about downgrade risk.

7.3 Model 3: Other ESG scores

In this section, we consider ESG score–adjusted variables and investigate their relevance for credit rating downgrade. We examine whether each of these variable shows independent association with downgrade risk respectively, and whether these effects are maintained after correcting for firm-level CO_2 emissions and emissions disclosure status. Such approach allows us to evaluate whether different ESG scores provide additional information beyond carbon exposure, or whether their effects are mostly absorbed by environmental risk captured through CO_2-related variables.

Before illustrating the result of each variable, it is important to recognize that ESG score components differ in their underlying measurement nature. Workforce, Human Right, and Management scores are mainly based on firms' self-reported disclosures in annual reports, sustainability reports, and other publicly available documents. As a result, these scores primarily capture the result of a qualitative assessment of a company's stated policies, commitments, and governance frameworks more than a number based and objective facts.

Differently, CO_2 emissions data and ESG controversy indicators are more closely tied to observable and externally verifiable information, such as reported (and audited) emissions figures, regulatory filings, and public incidents reported by media and regulatory sources. These measures therefore reflect more concrete aspects of firms' environmental exposure and realized events. OECD has classified ESG metrics into input-based, output-based, and business environment metrics (OECD, 2025). According to OECD's definitions, the input-based ones could be further divided into activity-based and policy-based metrics. The former relates to companies' measures and practices to manage ESG factors (relevant topics include corporate governance, business ethics, CSR, etc.) and the latter usually encompasses aspirational outlook of ESG performance (relevant topics include human rights, biodiversity, corruption, etc). On the contrary, output-based ESG data are associated with outputs of business activities such as employee salaries, injury rate, and GHG emissions. Apparently, Workforce and Human Right scores could be classified as input-based metrics while CO_2 emissions data and ESG Controversies score as output-based.

- Workforce score

We first analyze the effect of the variable Workforce score, one of the proxies of the social dimension of ESG.

The following tables display the coefficients of all the variables in the baseline logit model as well as the Workforce score without CO2-related variables.

	Coef.	Std.Err.	z	P>|z|	[0.025	0.975]
const	-2,421	0,377	-6,429	0,000	-3,159	-1,683
log_asset_resid	-0,035	0,024	-1,471	0,141	-0,083	0,012
EBITDA_Asset	-0,042	0,012	-3,592	0,000	-0,066	-0,019
leverage	1,612	0,239	6,740	0,000	1,143	2,081
Independent Board Members_f_filled	-0,003	0,002	-1,714	0,087	-0,006	0,000
Board Gender Diversity, Percent_f_filled	-0,009	0,004	-2,284	0,022	-0,016	-0,001
Board_missing	0,042	0,131	0,323	0,747	-0,214	0,298
Current Ratio_f_filled	0,018	0,072	0,253	0,800	-0,124	0,160
Current_Ratio_missing	-0,396	0,228	-1,739	0,082	-0,843	0,050
UR_f	0,034	0,010	3,514	0,000	0,015	0,053
GDP_growth_f	-0,089	0,033	-2,747	0,006	-0,153	-0,026
Workforce Score_filled	0,000	0,002	0,013	0,989	-0,005	0,005

Table 18 : Coefficients and p-values of main control variables and Workforce scores

We observed that the coefficient of this variable is statistically indistinguishable from zero and the p-value is close to one indicating that the variable has no detectable effect on the target variable. Consistent with this result, the inclusion of the variable leads to virtually no change in the pseudo-R squared.

Indeed, European labor markets benefit from strong regulatory protections, collective bargaining frameworks, and social protection systems. All of these tend to stabilize workforce outcomes regardless of firm-level policies. Besides, credit rating agencies do not systematically incorporate workforce metrics into their rating methodologies. Any financial consequences of poor workforce practices, such as reduced productivity or increased turnover, are likely already reflected in the firm's profitability and leverage ratios. Eventually, the relationship might be non-linear, with only extreme workforce disruptions such as mass strikes affecting downgrade risk.

Therefore, we fail to reject the null hypothesis H2a.

- Human Right scores

Another proxy of social dimension is the Human Right score.

	Coef.	Std.Err.	z	P>|z|	[0.025	0.975]
const	-2,318	0,374	-6,204	0,000	-3,050	-1,586
log_asset_resid	-0,032	0,024	-1,331	0,183	-0,078	0,015
EBITDA_Asset	-0,043	0,012	-3,601	0,000	-0,066	-0,019
leverage	1,611	0,239	6,738	0,000	1,142	2,079
Independent Board Members_f_filled	-0,003	0,002	-1,677	0,094	-0,006	0,000
Board Gender Diversity, Percent_f_filled	-0,008	0,004	-2,236	0,025	-0,015	-0,001
Board_missing	0,061	0,132	0,460	0,645	-0,197	0,318
Current Ratio_f_filled	0,016	0,073	0,217	0,828	-0,127	0,158
Current_Ratio_missing	-0,407	0,227	-1,795	0,073	-0,852	0,038
UR_f	0,035	0,009	3,645	0,000	0,016	0,053
GDP_growth_f	-0,090	0,033	-2,754	0,006	-0,154	-0,026
Human Rights Score_filled	-0,002	0,002	-0,879	0,379	-0,006	0,002

Table 19 : Coefficients and p-values of main control variables and Human Right scores

The above table shows that all control variables remain stable. The negative coefficient of Human Right score is consistent with the intuition that better human rights practices should reduce operational, legal, and reputational risks, and thereby lowering downgrade probability. However, the lack of statistical significance means that human rights performance in our sample does not contain incremental predictive power for credit downgrades beyond traditional financial and macroeconomic variables. Again, this could be explained by the fact that European developed markets are characterized by strong labor and human rights regulations (e.g., EU Charter of Fundamental Rights, EU Corporate Sustainability Reporting Directive (CSRD)) on large corporates. Another reason is that rating agencies simply do not involve Human Right scores in their regular assessment processes. However, once there are any controversies related to human rights, a credit assessment might be triggered.

In the current study, the human rights scores appear to be relatively high and uniform across firms, with limited variation for the model to exploit. Therefore, we fail to reject the null hypothesis H2a.

- Management score

As for the Governance pillar, on top of Board Independence and Board Gender Diversity, we further examine whether adding the variable Management score would contribute to explaining the target variable.

	Coef.	Std.Err.	z	P>|z|	[0.025	0.975]
const	-2,498	0,363	-6,876	0,000	-3,210	-1,786
log_asset_resid	-0,041	0,025	-1,680	0,093	-0,089	0,007
EBITDA_Asset	-0,042	0,012	-3,550	0,000	-0,065	-0,019
leverage	1,600	0,236	6,776	0,000	1,138	2,063
Independent Board Members_f_filled	-0,004	0,002	-1,956	0,051	-0,008	0,000
Board Gender Diversity, Percent_f_filled	-0,009	0,004	-2,515	0,012	-0,017	-0,002
Board_missing	0,030	0,131	0,231	0,817	-0,226	0,286
Current Ratio_f_filled	0,020	0,072	0,282	0,778	-0,121	0,161
Current_Ratio_missing	-0,385	0,230	-1,675	0,094	-0,836	0,065
UR_f	0,033	0,009	3,548	0,000	0,015	0,052
GDP_growth_f	-0,087	0,033	-2,651	0,008	-0,151	-0,023
Management Score_filled	0,002	0,002	1,056	0,291	-0,002	0,006

Table 20 : Coefficients and p-values of main control variables and Management scores

The Management score yields a small positive coefficient (0.002) that is statistically insignificant (p-value = 0.291), indicating no meaningful association with downgrade risk. Given the lack of statistical significance, the sign of the coefficient is not economically informative and likely reflects sampling variation rather than a systematic relationship. Similar to the other score variables, rating agencies may place little weight on generic management scores, preferring instead to evaluate more specific governance attributes such as Board Independence and Board Gender Diversity. The effect of management quality on credit risk may be largely passed through financial outcomes. For example, good management produces higher profitability and prudent leverage, which are already included as direct predictors.

Lohmann et al. noticed a U-shape relation between financial distress and ESG scores. Both financially distressed and well-performed companies can exhibit high ESG scores (Lohmann, Möllenhoff, & Lehner, 2025). The most plausible explanation is that “companies anticipate their upcoming financial distress and intensify ESG- supporting disclosures to manage their ESG scores upward”. In our model, such reverse causality cannot be ruled out, as firms experiencing financial pressure may strengthen management or governance practices in response. The Management score thus adds no significant incremental explanatory power beyond the existing financial and governance variables.

Hence, based on our empirical result, the hypothesis H3C is not rejected.

- ESG Controversies score

Similar as CO2 emissions, this variable is a result-based indicator. Higher scores represent fewer controversies. We have also tested the incremental explanatory power of this variable.

	Coef.	Std.Err.	z	P>|z|	[0.025	0.975]
const	-1,901	0,379	-5,014	0,000	-2,645	-1,158
log_asset_resid	-0,075	0,027	-2,729	0,006	-0,128	-0,021
EBITDA_Asset	-0,038	0,012	-3,211	0,001	-0,061	-0,015
leverage	1,472	0,238	6,185	0,000	1,006	1,939
Independent Board Members_f_filled	-0,004	0,002	-2,360	0,018	-0,008	-0,001
Board Gender Diversity, Percent_f_filled	-0,010	0,004	-2,786	0,005	-0,018	-0,003
Board_missing	0,105	0,132	0,796	0,426	-0,154	0,365
Current Ratio_f_filled	0,019	0,072	0,263	0,792	-0,122	0,160
Current_Ratio_missing	-0,372	0,237	-1,570	0,116	-0,837	0,093
UR_f	0,035	0,009	3,726	0,000	0,017	0,053
GDP_growth_f	-0,090	0,033	-2,687	0,007	-0,155	-0,024
ESG Controversies Score_filled	-0,006	0,001	-3,962	0,000	-0,009	-0,003

Table 21 : Coefficients and p-values of main control variables and ESG Controversies scores

	dy/dx	std err	z	P>|z|	[0,025	0,975]
log_asset_resid	-0,0048	0,002	-2,709	0,007	-0,008	-0,001
EBITDA_Asset	-0,0024	0,001	-3,221	0,001	-0,004	-0,001
leverage	0,0955	0,015	6,164	0,000	0,065	0,126
Independent Board Members_f_filled	-0,0003	0,000	-2,363	0,018	-0,000	-4,57e-05
Board Gender Diversity, Percent_f_filled	-0,0007	0,000	-2,781	0,005	-0,001	-0,000
Board_missing	0,0068	0,009	0,798	0,425	-0,010	0,024
Current Ratio_f_filled	0,0012	0,005	0,263	0,792	-0,008	0,010
Current_Ratio_missing	-0,0241	0,015	-1,572	0,116	-0,054	0,006
UR_f	0,0023	0,001	3,729	0,000	0,001	0,003
GDP_growth_f	-0,0058	0,002	-2,690	0,007	-0,010	-0,002
ESG Controversies Score_filled	-0,0004	9,5e-05	-3,977	0,000	-0,001	-0,000

Table 22 : Marginal effects of main control variables and ESG Controversies scores

The ESG Controversies score is statistically significant with a negative coefficient, indicating that firms with fewer and less severe controversies face a lower probability of downgrade. Therefore, we could reject the null hypothesis H2c. Compared with the baseline model, the log-likelihood improves from -2290.6 to -2282.2. The likelihood ratio test is thus 16.8, which is slightly lower than the impact of CO2 related variables.

This result is economically intuitive, as controversies represent realized negative events that can directly impair creditworthiness through multiple channels. For example, such events are often followed by immediate financial costs, including regulatory fines, legal fees, and remediation expenses, as well as higher financing costs. These mechanisms are particularly relevant in European developed markets, where regulatory enforcement is stringent and where investor sensitivity to ESG incidents is high. Consistent with this view, LaRosa demonstrated that environmental controversies are likely to increase the cost of equity in countries where the market regulation is stronger (Rosa & Bernini, 2022). Then on the revenue side, controversies could cause immediate reputational damage, loss of customers, contract terminations, or even restrictions on operating licenses. As controversies reveal weaknesses in internal management, it may also signal future credit deteriorations.

Again, we have calculated the marginal effect of the ESG Controversies score. One unit decrease in the score will lead to 0.04% increase in the probability of downgrades. Such marginal effect appears lower than that of the logarithm of CO2 emissions.

The significance of the ESG controversies score stands in contrast to the results for Workforce and Management scores. This reveals that rating agencies respond primarily to what a firm has actually done but may not account for in their regular credit assessment processes the qualitative metrics that reflect what a company states or the opinion of the ESG rating agencies. A firm may have excellent workforce policies and a high management score, but a single major controversy could override these attributes and trigger a downgrade. The ESG controversies score therefore captures a specific dimension of credit risk that is not already reflected in financial ratios, macroeconomic conditions, or other governance proxies.

7.4 Model 4: Combined model

Having established that CO2 emissions and the ESG controversies score are individually significant predictors of downgrade risk, while Management, Workforce, and Human Rights scores are not, we now proceed to a further specification: establish a model that includes both CO2 emissions related variables and ESG Controversies score simultaneously. This specification tests whether these two output-based ESG metrics capture independent dimensions of downgrade risk or whether one is more impactful than the other.

The following table reports the results of the augmented logit model that includes these additional variables compared to the baseline model.

	Coef.	Std.Err.	z	P>|z|	[0.025	0.975]
const	-2,864	0,584	-4,901	0,000	-4,009	-1,718
log_asset_resid	-0,108	0,030	-3,576	0,000	-0,168	-0,049
EBITDA_Asset	-0,037	0,012	-3,040	0,002	-0,060	-0,013
leverage	1,425	0,244	5,842	0,000	0,947	1,903
Independent Board Members_f_filled	-0,004	0,002	-2,465	0,014	-0,008	-0,001
Board Gender Diversity, Percent_f_filled	-0,010	0,004	-2,701	0,007	-0,018	-0,003
Board_missing	0,366	0,213	1,722	0,085	-0,051	0,783
Current Ratio_f_filled	0,036	0,072	0,505	0,613	-0,104	0,176
Current_Ratio_missing	-0,334	0,257	-1,300	0,194	-0,838	0,170
UR_f	0,033	0,010	3,460	0,001	0,014	0,052
GDP_growth_f	-0,089	0,034	-2,624	0,009	-0,155	-0,022
log_CO2	0,059	0,026	2,318	0,020	0,009	0,110
CO2_missing	-0,434	0,199	-2,180	0,029	-0,823	-0,044
ESG Controversies Score_filled	-0,005	0,002	-3,029	0,002	-0,008	-0,002

Table 23 : Coefficients and p-values of main control variables, CO2-related and ESG Controversies scores

	dy/dx	std err	z	P>|z|	[0,025	0,975]
log_asset_resid	-0,0070	0,002	-3,550	0,000	-0,011	-0,003
EBITDA_Asset	-0,0024	0,001	-3,052	0,002	-0,004	-0,001
leverage	0,0923	0,016	5,820	0,000	0,061	0,123
Independent Board Members_f_filled	-0,0003	0,000	-2,471	0,013	-0,000	-5,72e-05
Board Gender Diversity, Percent_f_filled	-0,0007	0,000	-2,698	0,007	-0,001	-0,000
Board_missing	0,0237	0,014	1,731	0,083	-0,003	0,051
Current Ratio_f_filled	0,0023	0,005	0,505	0,614	-0,007	0,011
Current_Ratio_missing	-0,0217	0,017	-1,301	0,193	-0,054	0,011
UR_f	0,0021	0,001	3,453	0,001	0,001	0,003
GDP_growth_f	-0,0057	0,002	-2,624	0,009	-0,010	-0,001
log_CO2	0,0038	0,002	2,317	0,020	0,001	0,007
CO2_missing	-0,0281	0,013	-2,186	0,029	-0,053	-0,003
ESG Controversies Score_filled	-0,0003	0,000	-3,036	0,002	-0,001	-0,000

Table 24 : Marginal effects of main control variables, CO2-related and ESG Controversies scores

The model fit improves substantially: the log-likelihood increases from -2290.6 in the baseline model to -2276.6, and the pseudo-R2 rises from 7.89% to 8.46%. The results appear to be better when considering only environmental and controversies variables respectively. A likelihood-ratio test rejects that the null that the added variables are jointly insignificant.

The coefficient on the logarithm of CO2 emission is positive and significant, confirming that higher CO2 emissions is associated with higher downgrade risk if holding all else constant. As shown by the marginal effect, one-unit increase in the logarithm of CO2 raises the probability of downgrade by 0.38 percentage points. Compared to Model 2, part of the marginal effect is absorbed by the newly introduced ESG Controversies score.

Besides, firms with missing CO2 data tend to have lower downgrade risk. As discussed earlier, this reflects non-random missingness. Non-disclosing firms tend to be less scrutinized by rating agencies in terms of ESG performance. They are not necessarily cleaner, but simply less transparent.

As for ESG Controversies score, a one-point increase in the controversies score (i.e. fewer controversies) reduces the probability of downgrade by 0.03 percentage points, a lower marginal effect than the logarithm of CO2 emissions.

All core financial, governance, and macroeconomic variables remain stable in sign, magnitude, and significance relative to the baseline model, except for the current ratio, which continues to be non-significant.

To conclude, these results provide strong evidence that outcome-based ESG metrics – carbon emissions and realized controversies – contain incremental predictive power for credit downgrades beyond traditional credit risk factors and input-based ESG scores.

7.5 Model 5: Change in CO2 emissions

In addition to the level of CO2 emissions, which captures a firm's static carbon exposure, we also consider the annual change in the logarithm of CO2 emissions as a potential predictor of downgrade risk. This is driven by the fact that rating agencies may respond not only to how carbon-intensive a firm is at a given point in time, but also firms' emissions trajectories over time. A firm that is rapidly increasing its emissions, even from a low starting point, may be seen as heading toward future regulatory or reputational risk. Whereas a firm that is aggressively reducing emissions may be viewed as proactively managing transition risk. Moreover, the change variable captures a different economic channel than the level because it reflects corporate actions and strategic direction rather than inherited or structural carbon intensity.

However, a reliable measure of the change of log CO2 emissions requires two consecutive years of non-missing CO2 data. Given that we have applied forward-filling (maximum two years) to address missing values in the level variable, the computed change for observations that rely on imputed data would be artificially smoothed to zero, introducing measurement error and potentially spurious correlations. Another important issue is that the change in emissions of each firm is mechanically missing in the first year observed. Because the sample begins in 2010 and emission levels are forward-filled when missing, a dummy indicating missing changes would be strongly correlated with the first sample year and the imputation procedure. As a result, this indicator captures a mechanical flaw in our data rather than firms' actual emission dynamics and thus generates spurious correlations. To avoid this issue, we estimate the change specification only on the subsample of firm-year observations without missing value of the change of log CO2 emission.

Consequently, we obtain the following statistics for the newly created variable.

Parameters	log_CO2_change
count	8.605
mean	-0,06
std	0,57
min	-9,63
25%	-0,10
50%	-0,02
75%	0,02
max	7,30

Table 25 : Descriptive statistics of changes in logarithm of CO2

The mean and median of this variable are both negative, indicating a slight downward tendency in emissions over the sample period. This is consistent with a gradual decarbonization trend among European developed public companies. The standard deviation is 0.57, reflecting a considerable cross-sectional and temporal variation in emissions changes. The 25th and 75th percentiles are -0.10 and 0.02, respectively, meaning that the majority of observations lie within a narrow interval around zero. However, the minimum of -9.63 and the maximum of 7.30 reveal the presence of extreme outliers, possibly due to non-disclosure status.

To prove this hypothesis, we have traced the underlying CO2 reporting status for observations with a change of log CO2 emissions greater than 2. Among the 131 outliers identified, 96 (73%) occur when the lagged CO2 value is missing while the current value is actually reported. The rest of cases cannot be directly attributed to imputation. They may reflect genuine but rare events such as major acquisitions, divestitures, or plant closures, or they may stem from other data reporting inconsistencies.

As a result, we deem that by adding such variable, the logit model may be "fitting" these outliers even though they do not represent real economic variation. Consequently, we do not rely on this variable for the main analysis. The level of CO2 emissions, which is more robust in data quality, remains the primary environmental predictor in the final specification.

7.6 Final model

The Current Ratio and its associated missing indicator are not statistically significant in the combined model (Model 4). This finding is consistent with prior credit risk research, which often finds that profitability and leverage dominate liquidity in predicting rating downgrades, particularly for longer forecast horizons or for firms in developed markets with access to credit lines and capital markets. For example, Gopalan et al. found that rating agencies may underestimate liquidity risk (Gopalan, Song, & Yerramilli, 2010). Therefore, we have decided to drop the insignificant control variables. We start by dropping Current Ratio and its missing indicator. It turned out that the missing indicator of Board Independence also becomes non-significant. Indeed, the board missing indicator is currently marginally significant at a p value of 0.085. Marginal significance of a missing indicator is inherently difficult to interpret. Unlike a core variable such as leverage or profitability, a missing indicator may not capture a sure economic construct. Hence, we also drop this variable.

The following table presents the full results of the final model.

Final Model: Logit Regression Results

Dep. Variable:	flg_dg	No. Observations:	9406
Model:	Logit	Df Residuals:	9372
Method:	MLE	Df Model:	33
Date:		Pseudo R-squ.:	0.0835
Time:		Log-Likelihood:	-2279.2
converged:	True	LL-Null:	-2486.9
Covariance Type:	cluster	LLR p-value:	1,13E-67

Table 26 : Summary of final model

	Coef.	Std.Err.	z	P>|z|	[0.025	0.975]
const	-2,871	0,529	-5,428	0,000	-3,908	-1,835
log_asset_resid	-0,122	0,031	-3,986	0,000	-0,182	-0,062
EBITDA_Asset	-0,036	0,012	-2,947	0,003	-0,059	-0,012
leverage	1,347	0,241	5,601	0,000	0,876	1,819
Independent Board Members_f_filled	-0,004	0,002	-2,410	0,016	-0,007	-0,001
Board Gender Diversity, Percent_f_filled	-0,010	0,004	-2,650	0,008	-0,018	-0,003
UR_f	0,033	0,010	3,373	0,001	0,014	0,052
GDP_growth_f	-0,089	0,033	-2,659	0,008	-0,154	-0,023
log_CO2	0,064	0,025	2,555	0,011	0,015	0,113
CO2_missing	-0,242	0,134	-1,798	0,072	-0,505	0,022
ESG Controversies Score_filled	-0,005	0,002	-2,946	0,003	-0,008	-0,002

	Coef.	Std.Err.	z	P>|z|	[0.025	0.975]
Sector_Consumer Cyclicals	0,197	0,185	1,063	0,288	-0,166	0,559
Sector_Consumer Non-Cyclicals	-0,272	0,227	-1,197	0,231	-0,716	0,173
Sector_Energy	0,624	0,205	3,049	0,002	0,223	1,025
Sector_Financials	-0,273	0,246	-1,112	0,266	-0,755	0,208
Sector_Healthcare	0,358	0,251	1,427	0,154	-0,134	0,850
Sector_Industrials	-0,160	0,181	-0,884	0,377	-0,514	0,194
Sector_Real Estate	0,263	0,318	0,829	0,407	-0,359	0,886
Sector_Technology	0,218	0,205	1,066	0,286	-0,183	0,620
Sector_Utilities	0,111	0,195	0,568	0,570	-0,271	0,492
year_2011	0,308	0,183	1,688	0,091	-0,050	0,666
year_2012	-0,572	0,236	-2,422	0,015	-1,035	-0,109
year_2013	-0,525	0,241	-2,173	0,030	-0,998	-0,051
year_2014	-0,145	0,215	-0,673	0,501	-0,566	0,276
year_2015	-0,379	0,237	-1,598	0,110	-0,844	0,086
year_2016	-1,313	0,303	-4,329	0,000	-1,908	-0,719
year_2017	-0,846	0,261	-3,239	0,001	-1,359	-0,334
year_2018	-0,682	0,254	-2,686	0,007	-1,180	-0,184
year_2019	0,660	0,204	3,235	0,001	0,260	1,060
year_2020	-1,404	0,382	-3,671	0,000	-2,153	-0,654
year_2021	-0,592	0,316	-1,877	0,060	-1,211	0,026
year_2022	-0,590	0,267	-2,207	0,027	-1,114	-0,066
year_2023	-0,537	0,262	-2,048	0,041	-1,052	-0,023
year_2024	-0,219	0,248	-0,880	0,379	-0,706	0,268

Table 27 : Coefficients and p values of final model

After removing the Current Ratio and the board missing indicator, the final logit model achieves a pseudo-R-squared of 8.35% with all retained variables statistically significant and displaying economically intuitive signs.

The final model could be expressed as:

$$\ln\left(\frac{p}{1-p}\right) = constant + \beta_1 * Control_{i,t-1} + \beta_2 * Macro_{i,t-1} +$$
$$\beta_3 * Industry\ FE + \beta_4 * Year\ FE +$$
$$\beta_5 * CO2_{i,t-1} + \beta_5 * Controversies_{i,t-1} + \varepsilon_{i,t}$$

From the final model, we conclude that, as evidence from S&P Global ratings for public companies in the European developed countries, credit rating downgrades are predicted by a combination of firm-level carbon emissions (transition risk) and ESG controversies. Input-based ESG scores (management, workforce, human rights) do not add predictive power.

Although ESG data is accessible to rating agencies (some even have their own ESG ratings), they may not react immediately when such data becomes available. The findings of the study by Baldan et al. shows a much weaker statistical significance of ESG scores compared to traditional financial variables on firms' credit stability (Baldan, Zen, & Targhetta, 2026). Hence, as the link between ESG performance to credit risk is not as prominent as traditional financial metrics, rating agencies often wait until material events happen. Also as mentioned in section 4, S&P integrates ESG information mainly in a qualitative way and only since recent years. In this way, ESG variables act as a leading indicator to downgrades despite their accessibility.

8. Robustness analysis

To assess the reliability and stability of the main results, we have conducted a series of robustness checks. The necessity of such tests arises from several features of the main model: the presence of missing data in ESG and governance variables, the use of forward-filling imputation and missing indicator dummies, the unbalanced panel structure that varying firm observation lengths, the potential bias raised from omitted variables (particularly country-level physical risk), the possibility that relationships may differ over time, the potential bias introduced by divergence of CO2 emissions across industries, and the different ESG effect for different initial ratings of downgrade events. A finding that survives multiple alternative specifications is less likely to be spurious due to data handling, sample composition, or model specification.

8.1 Test 1: Removing missing indicators

As a first robustness check, we re-estimate the main model by excluding the missing indicators. This specification retains only the non-imputed actual reported values for each variable, dropping observations with missing data rather than including them via indicator variables. The goal is to assess whether the significant coefficients on logarithm CO2 emissions and ESG Controversies score are driven by systematic differences between firms that report ESG data and those that do not.

The list of variables includes board independence, board gender diversity, profitability, leverage, firm size, macroeconomic controls, log CO2 emissions, and ESG Controversies score. This subsample comprises 7402 observations, compared to the larger imputed sample used in the main analysis.

	Coef.	Std.Err.	z	P>|z|	[0.025	0.975]
const	-3,295	0,637	-5,173	0,000	-4,543	-2,047
log_asset_resid	-0,126	0,035	-3,627	0,000	-0,193	-0,058
EBITDA_Asset	-0,044	0,011	-3,983	0,000	-0,066	-0,022
leverage	1,353	0,300	4,506	0,000	0,765	1,942
Independent Board Members_f_filled	-0,005	0,002	-2,577	0,010	-0,008	-0,001
Board Gender Diversity, Percent_f_filled	-0,014	0,004	-3,313	0,001	-0,022	-0,006
UR_f	0,026	0,011	2,442	0,015	0,005	0,046
GDP_growth_f	-0,060	0,036	-1,697	0,090	-0,130	0,009
log_CO2	0,068	0,029	2,329	0,020	0,011	0,126
ESG Controversies Score_filled	-0,005	0,002	-2,993	0,003	-0,008	-0,002

Table 28 : Coefficients and p values of the main variables from the model without missing indicators

The pseudo-R squared increases to 8.86%, slightly higher than the 8.35% achieved in the main model, suggesting a cleaner signal when only originally reported data are used. Critically, all key variables retain their signs and statistical significance.

This result confirms that the predictive power of carbon emissions and ESG controversies for downgrade risk is not a coincidence of missing data imputation or the inclusion of missing dummies. Even when restricting the analysis to observations with complete, originally reported data for every variable, the core conclusions remain unchanged.

Apart from that, the significances of industry fixed effects also remain stable, which is not the case for the year fixed effects. As shown below, we have observed a great shift in the significance of year FE.

	Coef.	Std.Err.	z	P>|z|	[0.025	0.975]
year_2011	0,823	0,282	2,923	0,003	0,271	1,375
year_2012	0,150	0,314	0,479	0,632	-0,465	0,765
year_2013	0,011	0,348	0,033	0,974	-0,671	0,693
year_2014	0,781	0,285	2,740	0,006	0,222	1,340
year_2015	0,397	0,317	1,255	0,210	-0,223	1,018
year_2016	-0,578	0,389	-1,486	0,137	-1,341	0,184
year_2017	-0,436	0,365	-1,195	0,232	-1,151	0,279
year_2018	-0,093	0,330	-0,281	0,779	-0,739	0,554
year_2019	1,311	0,288	4,554	0,000	0,747	1,875
year_2020	-0,497	0,441	-1,129	0,259	-1,361	0,366
year_2021	-0,096	0,389	-0,248	0,804	-0,858	0,665
year_2022	-0,137	0,345	-0,397	0,691	-0,813	0,539
year_2023	0,170	0,335	0,506	0,613	-0,488	0,827
year_2024	0,560	0,318	1,759	0,079	-0,064	1,184

Table 29 : Coefficients and p values of the year FE variables from the model without missing indicators

In the complete-case robustness test, the year fixed effects exhibit a very different pattern from the main model. Relative to the reference year 2010, the years 2011, 2014, and 2019 show significant higher baseline downgrade risk, with 2024 being marginally significant. Indeed, when missing observations are dropped, the sample loses many firms for early years (especially around 2010–2012). The remaining sample is thus less balanced across years. This shift likely reflects changes in sample composition rather than a substantive shift in credit-cycle dynamics. For instance, our final model considers also the relatively smaller firms with discontinuous ESG reporting with different downgrade dynamics. Nevertheless, in our current subsample, we only include firms that consistently report ESG information and thus are more scrutinized by rating agencies.

Most importantly, the core variables remain significant and stable across different specifications. The year fixed effects are control variables, and their shift in significance does not undermine the main conclusion that output-based ESG metrics are relevant to downgrade risk.

From this robustness analysis, we conclude that the main results are robust to sample selection.

8.2 Test 2: Shorter observation period

Second, we restrict the sample to the period from 2015 onwards for mainly two reasons. The Paris Agreement marked a turning point in global climate policy and investors' attention towards climate risk, potentially increasing the importance of ESG factors in credit assessment. Additionally, ESG data coverage and quality improved notably after 2015, particularly following the EU Non-Financial Reporting Directive (2014, effective in 2017).

The coefficients and p-values of the key variables are summarized as below.

	Coef.	Std.Err.	z	P>|z|	[0.025	0.975]
const	-2,219	0,643	-3,452	0,001	-3,479	-0,959
log_asset_resid	-0,123	0,043	-2,884	0,004	-0,206	-0,039
EBITDA_Asset	-0,033	0,013	-2,614	0,009	-0,058	-0,008
leverage	1,282	0,298	4,294	0,000	0,697	1,867
Independent Board Members_f_filled	-0,005	0,002	-1,909	0,056	-0,009	0,000
Board Gender Diversity, Percent_f_filled	-0,013	0,005	-2,831	0,005	-0,022	-0,004
UR_f	0,004	0,016	0,252	0,801	-0,027	0,035
GDP_growth_f	-0,044	0,039	-1,131	0,258	-0,120	0,032
log_CO2	0,028	0,031	0,902	0,367	-0,033	0,090
CO2_missing	-0,314	0,185	-1,693	0,090	-0,677	0,050
ESG Controversies Score_filled	-0,005	0,002	-2,519	0,012	-0,009	-0,001

Table 30 : Coefficients and p values of the main variables from the model with shorter window

This subsample comprises 6710 observations, and the model specification is identical to the final main model (including all controls, log CO2, CO2 missing indicator, and ESG Controversies score). The pseudo-R squared increases to 9.842%, slightly higher than in the full-sample main model (8.35%). However, several variables lose statistical significance. The macroeconomic controls – unemployment rate and GDP growth – are no longer significant, though their coefficients retain the expected signs (positive for unemployment, negative for GDP growth). More importantly, the coefficient on log CO2 also becomes non-significant but remains positive.

Such results are normal and reasonable as recent evidence suggests that the valuation of ESG factors in credit risk assessment is time-varying rather than stable across periods (Ferriani & Pericoli, 2026). A key factor underlying these changes is the sharp reduction in downgrade events following the exclusion of the 2011–2012 crisis period. Rating downgrades are relatively rare and tend to be concentrated during downturn periods, and the European sovereign debt crisis accounts for a substantial share of observed downgrade activity in the full sample. Removing this period would therefore reduce variation in the dependent variable and limit the statistical power of certain risk factors.

Beyond this mechanical effect, the loss of significance of the CO_2 coefficient is also consistent with a regime shift after the Paris Agreement. As companies start to be more transparent in ESG disclosure, ESG related losses become more predictable than pre-2015 period. Besides, EU-wide CO_2 emissions declined by 26.4% during 2000-2022 (Tutar, Štreimikienė, & Kyriakopoulos, 2026). Therefore, cross-sectional dispersion in CO_2 emissions may have narrowed as many European firms initiated decarbonization efforts.

As for ESG controversies score, it remains statistically significant in this subsample.

Despite the loss of significance for log value of CO2 emissions, the overall model fit improves, and the signs of all coefficients remain consistent with the economic intuition. The result for CO2 in the post-2015 period does not invalidate the full-sample finding but suggests that the predictive power of carbon emissions for downgrades is stronger when including the pre-Paris-Agreement era. The robustness of the ESG Controversies score across both samples confirms its role as an important indicator of downgrade risk.

8.3 Test 3: Removing companies with fewer observations

From our sample used for the final model, we have drawn the frequency of companies with different number of observation years.

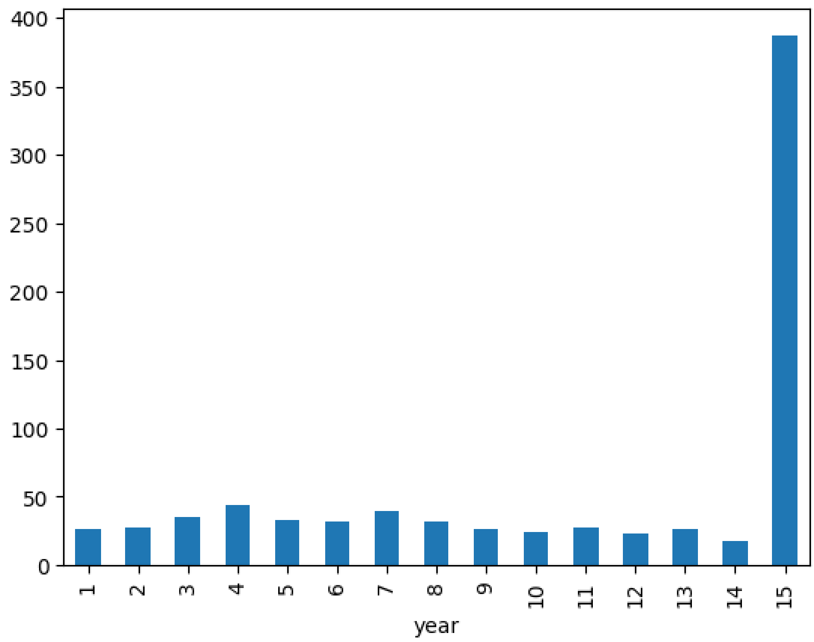


Figure 15 : Distribution of companies with different total number of year observations

As the figure shows, despite that most companies in our sample have more than 10 years of observations, there are still a minority of companies with less than 3 observation points. These companies contribute little information. They potentially introduce bias when estimating within-firm changes. Their downgrade risk may be thus poorly estimated. To account for composition bias, we have performed another robustness analysis by removing the companies with fewer than 3 observation points.

	Coef.	Std.Err.	z	P>|z|	[0.025	0.975]
const	-2,926	0,532	-5,497	0,000	-3,969	-1,883
log_asset_resid	-0,119	0,030	-3,914	0,000	-0,179	-0,059
EBITDA_Asset	-0,034	0,012	-2,857	0,004	-0,058	-0,011
leverage	1,360	0,242	5,623	0,000	0,886	1,834
Independent Board Members_f_filled	-0,004	0,002	-2,310	0,021	-0,007	-0,001
Board Gender Diversity, Percent_f_filled	-0,010	0,004	-2,730	0,006	-0,018	-0,003
UR_f	0,033	0,010	3,424	0,001	0,014	0,052
GDP_growth_f	-0,085	0,033	-2,557	0,011	-0,151	-0,020
log_CO2	0,063	0,025	2,472	0,013	0,013	0,112
CO2_missing	-0,263	0,135	-1,945	0,052	-0,528	0,002
ESG Controversies Score_filled	-0,004	0,002	-2,796	0,005	-0,007	-0,001

Table 31 : Coefficients and p values of the main variables from the model without short panels

The model specification is identical to the main final model with 9318 observations. The results are nearly identical to those of the main model. This finding confirms that the main results are not driven by firms with very short panel histories, which might have different risk characteristics or measurement error.

The conclusions of this study are therefore robust to the exclusion of firms with limited time-series information.

8.4 Test 4: Including country vulnerability index

As another robustness check, we added the country-level ND-GAIN Country Vulnerability Index (Notre Dame Global Adaptation Initiative) to the main model. The index provides a composite score for each country, ranging from 0 (least vulnerable) to 1 (most vulnerable). It measures a country's exposure to climate-related hazards (e.g. floods, storms, heatwaves), sensitivity of the population and infrastructure to those hazards and capacity to adapt to the resulting negative effects (e.g. early warning systems, governance, infrastructure resilience). A lower score indicates lower physical risk.

The ND-GAIN index is widely validated in the literature as a robust proxy for physical climate risk since it aggregates both geophysical and socio-economic dimensions. Garschagen et al. compared three indices assessing country-level climate and disaster risk: the World Risk Index, the INFORM Risk Index, ND-GAIN Index, and the Climate Risk Index (Garschagen, Doshi, Reith, & Hagenlocher, 2021). They found that Indices like ND-GAIN are especially reliable at identifying which country is vulnerable and lacks capacity, even if hazard exposure is hard to be measured precisely. Moreover, Fitch Ratings has stated a consideration of Climate Vulnerability Signals in its credit assessment (Fitch Ratings, 2023).

In the context of this study, physical risk affects corporate creditworthiness through channels distinct from the transition risk captured by firm-level CO2 emissions in the form of direct damage to assets, supply chain interruptions, higher insurance costs, and potential loss of market access after extreme weather events.

Hence, we merged the index with each firm based on its country of headquarters. However, some firm-year observations are lost due to missing matching values (882 observations). The main reason for these missing values is because the ND-GAIN data is only available until 2023. This is one of the reasons why we include this analysis in the section of robustness analysis as loss of approximately 900 observations, or nearly 10% of the sample would reduce the sample size and lead to a loss of valuable information for the most recent year.

The rest of small number of missing values is because the countries of headquarters are not on the list of the ND-GAIN index. Nevertheless, despite the loss of 882 observations, the reduced sample remains sufficiently large to conduct a meaningful robustness analysis.

As explained in section 3, we assume that the countries of headquarters are where firms primarily operate. This may reveal a potential difference between our “country of headquarter” assumption and S&P Global’s markets classification methodologies. Another point of attention is that many public companies in our sample operate globally, with production facilities, supply chains, and customer bases spread across multiple countries with different physical risk profiles. A firm headquartered in a low-vulnerability country may have significant operations in high-vulnerability country, and vice versa. Using the headquarters country may eventually misallocate physical risk exposure. As a consequence, the coarseness and potential measurement error of this variable make it less suitable as a core variable in the main specification.

Despite all these potential issues, the country level index is still a useful complement to test the broad hypothesis that physical climate risk matters.

The table below presents the coefficients and p-values of each explanatory variable included in the regression model, where both CO2 related variables and ESG Controversies scores are kept.

	Coef.	Std.Err.	z	P>|z|	[0.025	0.975]
const	-4,524	0,804	-5,629	0,000	-6,100	-2,949
log_asset_resid	-0,142	0,033	-4,269	0,000	-0,208	-0,077
EBITDA_Asset	-0,033	0,012	-2,653	0,008	-0,057	-0,009
leverage	1,283	0,252	5,097	0,000	0,790	1,776
Independent Board Members_f_filled	-0,005	0,002	-2,614	0,009	-0,008	-0,001
Board Gender Diversity, Percent_f_filled	-0,006	0,004	-1,585	0,113	-0,014	0,002
UR_f	0,030	0,010	2,989	0,003	0,010	0,050
GDP_growth_f	-0,092	0,035	-2,603	0,009	-0,161	-0,023
log_CO2	0,090	0,027	3,301	0,001	0,036	0,143
CO2_missing	-0,249	0,143	-1,734	0,083	-0,530	0,032
ESG Controversies Score_filled	-0,005	0,002	-2,826	0,005	-0,008	-0,001
Country Vulnerability Index	3,563	1,662	2,144	0,032	0,306	6,820

Table 32 : Coefficients and p values of the main variables from the model with country vulnerability index

A positive coefficient would indicate that firms headquartered in countries with higher physical vulnerability face a higher downgrade probability, holding constant firm-level transition risk and other controls. The result enriches the overall study in that physical risk acts as an independent channel, which completements the other variables. In other words, while earlier results showed that firm-level CO2 emissions and ESG controversies predict downgrades, the significance of the country-level vulnerability index demonstrates that external, geography-based climate hazards also directly affect creditworthiness.

8.5 Test 5: Demeaned CO2 emission by industry

Even though we have included the industry fixed effects, they do not adjust the scale of the CO2 emissions. If the variances of the logarithm of CO2 emission in some industries are much higher than the others, then the coefficients we obtained may be largely driven by the former. To account for this potential bias, we have performed another robustness test, which is to demean the log CO2 by the average value within the industry and re-estimate the model. We have firstly simply replaced the log CO2 with the demeaned value (i.e. the difference between the log CO2 and the industry-average). However, the model does not differ from the final model at all, meaning that the industry FE dummies fully absorb the industry averages of the logarithm CO2.

	Coef.	Std.Err.	z	P>|z|	[0.025	0.975]
const	-1.907	0.316	-6.037	0.000	-2.526	-1.288
log_asset_resid	-0.173	0.026	-6.588	0.000	-0.224	-0.121
EBITDA_Asset	-0.020	0.009	-2.242	0.025	-0.037	-0.002
leverage	0.976	0.213	4.586	0.000	0.559	1.394
Independent Board Members_f_filled	-0.004	0.002	-2.219	0.026	-0.007	0.000
Board Gender Diversity, Percent_f_filled	-0.011	0.004	-2.770	0.006	-0.019	-0.003
UR_f	0.033	0.010	3.508	0.000	0.015	0.052
GDP_growth_f	-0.093	0.033	-2.801	0.005	-0.158	-0.028
log_CO2_demeaned	0.079	0.024	3.234	0.001	0.031	0.127
CO2_missing	-0.380	0.133	-2.858	0.004	-0.640	-0.119
ESG Controversies Score_filled	-0.005	0.002	-3.404	0.001	-0.008	-0.002

Table 33 : Coefficients and p values of the main variables from the demeaned model

Then, we dropped the industry dummies to account for such multicollinearity. As the above table shows, all of the selected variables are significantly correlated with the target variables with signs as expected. Since the demeaned logarithm of CO2 emissions represents the variation within industry, its significance implies that the relationship between the relative carbon performance and the target variable is robust within industries. In other words, within an industry, a higher-than-average CO2 emission could lead to higher downgrade risk. However, without industry FE dummies, we are not capturing other unobservable differences in downgrade dynamics across industries. This is reflected by the fact that the pseudo-R square (7.58%) and the log-likelihood (-2298.4) are both lower than the baseline model.

8.6 Test 6: Initial rating before downgrades

As the final robustness test, we take into account the possibility that the downgrade dynamics of firms with worse initial ratings are different from those with better initial ratings. The following graph helps us to understand better the distributions of the initial ratings of firms who have experienced downgrades.

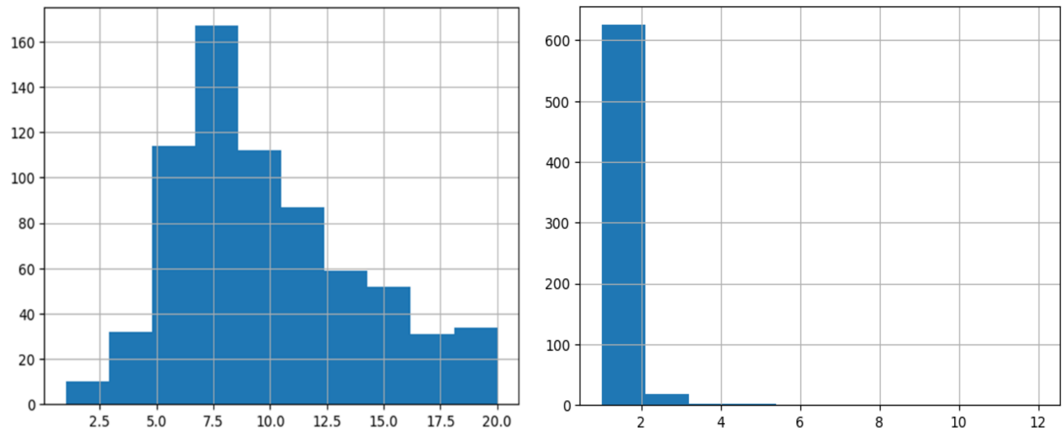


Figure 16 : Distribution of initial ratings before downgrades and notches of downgrades

We have mapped the S&P rating scales to numerical values with AAA as 1 and D as 20. The mean value of the initial rating number is 10, representing the rating BBB-. Indeed, this is the worst rating within the investment grade category. The majority of the downgrades are limited to one notch.

By filtering only the downgrades initiated from ratings no better than BBB-, we have obtained the following model specification:

	Coef.	Std.Err.	z	P>|z|	[0.025	0.975]
const	-4.574	0.833	-5.490	0.000	-6.207	-2.941
log_asset_resid	-0.400	0.072	-5.571	0.000	-0.541	-0.259
leverage	0.967	0.401	2.409	0.016	0.180	1.753
Independent Board Members_f_filled	-0.006	0.004	-1.755	0.079	-0.013	0.001
Board Gender Diversity, Percent_f_filled	-0.030	0.007	-4.187	0.000	-0.043	-0.016
UR_f	0.071	0.015	4.775	0.000	0.042	0.101
GDP_growth_f	-0.118	0.046	-2.575	0.010	-0.208	-0.028
log_CO2	0.091	0.045	2.032	0.042	0.003	0.180
ESG Controversies Score_filled	-0.005	0.003	-1.697	0.090	-0.012	0.001

Table 34 : Coefficients and p values of the main variables from the rating-reduced model

We have dropped the insignificant variables including the EBITDA-to-asset ratio, the current ratio and its missing indicator, as well as the missing indicators of board and CO2 information. The remaining variables all have the expected signs. A remarkable finding from this specification is that the model fit has improved substantially with a pseudo-R square of 17.76% and log-likelihood of -1095.8. This further supports the divergence of downgrade dynamics between good and bad initial ratings. The explanatory variables are far more relevant to explain the downgrades among lower-rated firms. Such results are consistent with economic intuition in that lower-rated firms are more sensitive to risks so their ratings tend to be more responsive to negative signals. It is also worth noting that the non-disclosure status is no longer relevant for this subsample. The benefit from being less scrutinized due to less disclosing offsets with the closer scrutiny of firms with higher static credit risks.

Overall, the results further support the relevance of ESG variables in predicting downgrade risks, with lower-rated downgrade more likely explained by these factors.

To summarize this section, the robustness analysis confirms that the main findings are stable across alternative sample choices and data treatments. Output-based ESG metrics, particularly ESG controversies, consistently predict downgrade likelihood; while the relevance of firm-level carbon emissions is stronger in earlier periods with higher dispersion and more frequent downgrade events. Moreover, the significance of the climate vulnerability proxy indicates that country-level exposure to physical climate risk, provides incremental information for downgrade risk beyond firm-level ESG characteristics. Finally, the test on solely lower-than-average-rated firms shows a higher explanatory power of ESG variables on this subgroup.

9. Implications for banks

This study provides robust evidence that credit rating downgrades in European developed markets are predicted by outcome-based ESG metrics: higher firm-level CO2 emissions (transition risk), more frequent and severe ESG controversies, and higher country-level physical climate vulnerability (physical risk). Input-based scores (management, workforce, human rights) do not appear to contribute incremental predictive power. These findings have practical relevance for banks that incorporate external credit ratings into their SICR frameworks. While large banks typically maintain sophisticated internal rating systems, smaller banks and those with limited resources may reply more on external ratings as a reference point for assessing SICR, especially for corporate and sovereign counterparties.

Before turning to the practical implications of this study, an important reminder is necessary. The dependent variable in all the above models is derived from credit rating downgrades as determined by S&P Global. A downgrade represents a rating agency's opinion, not a direct measure of a firm's true credit quality. As evidenced in this study by the significant negative coefficient on the CO2 missing indicator (i.e. non-disclosing firms tend to receive systematically more favorable rating treatment regardless of their underlying risk), rating agencies may be slow, inconsistent, or influenced by factors unrelated to fundamental credit risk. Consequently, the finding that certain ESG variables predict downgrades does not necessarily mean that these variables predict genuine increases in credit risk. Although rating agencies perform regularly calibration to improve model accuracy, these variables could only predict what rating agencies would do under our models. For banks using external ratings as an input into IFRS9 SICR assessment, the following implications should be understood in this regard, i.e. these findings help banks to better understand and to anticipate rating agency behavior. The prominent factors found in this study may also inspire banks when conducting their own independent assessment of whether credit risk has truly increased.

- Early warning signals

The finding that both CO2 emissions and ESG controversies are significant predictors of downgrade risk confirms that rating agencies do incorporate these factors into their rating decisions beyond financial fundamentals and macroeconomic conditions. However, it does not represent that these factors are calibrated in a timely sense so as to reflect true credit deterioration. Rating agencies may react slowly to emerging ESG risks, waiting for confirmed financial deterioration before taking actions. For banks using external ratings as an input into IFRS9 SICR assessments, the practical implication is that CO2 emissions and ESG controversies can serve as indicators of when a downgrade is likely to occur. Concretely, a counterparty with CO2 emissions substantially above its sector's average may be flagged for enhanced monitoring, even if its current external rating remains stable. Note that a bank should not mechanically treat a predicted downgrade as equivalent to a real increase in credit risk. Instead, these variables may be integrated into banks' internal risk assessment framework. Similarly, a recent ESG controversy such as a regulatory fine, a major operating incident, or a governance failure, should prompt an immediate credit review. Controversies are discrete and observable events that often precede formal downgrades by several months. When a counterparty exhibits both high emissions and a recent controversy, this combination should be treated as high-informative early warning signal, triggering a full credit assessment within a short timeframe. A final remark is that banks shall keep in mind the stronger impacts of these variables on high-yield firms. Non-disclosure status does not matter anymore in face of this subgroup.

- Non-disclosure bias

A key empirical finding of this study is that the CO2 missing indicator enters the downgrade model with a negative and statistically significant coefficient. This means that, holding all other factors constant, firms that do not report their CO2 emissions have a systematically lower observed downgrade risk than similar firms that do report. This effect should not be interpreted as reflecting lower underlying credit risk among non-disclosing firms but instead is consistent with a disclosure-driven selection mechanism affecting observed rating outcomes. For banks that rely on external ratings when assessing significant increases in credit risk under IFRS 9, this finding has clear practical implications. Banks may need to flag counterparties with incomplete ESG disclosures and treat the absence of reported emissions or related metrics as an indicator for enhanced due diligence. Where appropriate, the application of temporary rating overlays may be justified until adequate information is obtained.

- Input-based vs output-based ESG data

One of the central findings of this study is that input-based ESG scores – specially the management score, workforce score, and human rights score – do not emerge as statistically significant predictors of downgrade risk in the estimated models. By contrast, output-based metrics, namely CO2 emissions and the ESG controversies score, are consistently significant across multiple specifications. This divergence provides a direct and practical reminder for banks that relies on commercial ESG data for SICR assessment. Many ESG data providers aggregate activity or policy indicators such as whether a firm has a human rights policy or a workforce training programme into overall scores. These scores, while easy to use, do not appear to contain information that predicts credit downgrades. For rating agencies, what matters more for credit risk is what a firm actually does based on factual data, not what it states it is doing or will do. A counterparty may have a comprehensive human rights policy and a perfect workforce score, yet if it emits high level of CO2 or has recently been involved in any controversies, it remains a downgrade risk. Conversely, a firm with moderate or missing input-based scores but low emissions and no controversies may be more creditworthy. For SICR practitioners, the implication is not to ignore input-based factors completely but to recognize that external ratings alone do not fully capture them. Input-based ESG scores are valuable once validated against real factual outcomes. More generally, the results highlight the importance of examining individual ESG dimensions separately, as reliance on an aggregated score may dilute the contribution of specific risk-relevant factors, such as emissions intensity or recent controversies.

- Country vulnerability

The robustness test that adds the ND-GAIN Country Vulnerability Index reveals that physical climate risk at the country level is a positive and significant predictor of downgrade risk, even after controlling for firm-level CO2 emissions and ESG controversies. This suggests that country-level physical vulnerability carries incremental explanatory power for downgrade risk, over and above other firm-level ESG characteristics. For SICR assessment, banks may therefore integrate country-level physical risk data as a separate factor. One practical approach is to apply a country overlay to all exposures from high-vulnerability countries, such as those in southern Europe with elevated exposure to heatwaves, droughts, or floods. For long-term loans in these jurisdictions, banks may require counterparties to provide climate adaptation plans or evidence of adequate insurance coverage. What's more, special attention should be paid to firms with cross-border operations. For instance, a counterparty headquartered in a low-vulnerability country may still hold substantial production infrastructure or supply chain facilities in high-vulnerability regions. In such cases, relying solely on headquarters location may understate physical risk exposure. Banks should request information on the geographic distribution of material assets and adjust SICR assessment accordingly. The overarching principle is that physical risk is location-specific and cannot be fully captured by firm-level ESG data alone. Incorporating country-level vulnerability indices into SICR frameworks hence provides an independent, complementary layer of risk assessment that improves the timeliness and accuracy of identifying significant increases in credit risk.

To conclude this section, this study helps banks to understand how rating agencies behave with respect to ESG factors. CO2 emissions, ESG controversies, and country-level physical vulnerability are all associated with future downgrades. However, a downgrade is an opinion and may not reflect the true credit quality. As rating agencies may be slow, inconsistent, or biased, banks that rely on external ratings for SICR could hence use the findings of this study to anticipate rating agency actions without mechanically equating a predicted downgrade with a real increase in credit risk. For carbon exposure and physical risk, banks should conduct their own independent assessments regardless of what rating agencies do. Eventually, for input-based ESG scores, banks should recognize that external ratings may not capture them.

10.Limitation and future development

While this study provides robust evidence on the relationship between ESG factors and credit rating downgrades, several limitations, particularly data related, should be acknowledged. These limitations do not invalidate the findings of this study but rather qualify their scope and suggest directions for future research.

- Sample restricted to S&P-rated companies

The sample is restricted to publicly traded companies in European developed markets that are rated by S&P Global meaning that the study focuses on firms that are sufficiently large, transparent, and capital-market-oriented. This introduces a size-based selection bias as smaller, privately held, and unrated corporates are excluded by construction. However, these entities represent an important share of the real economy and banks' loan portfolios. Credit risk dynamics for such firms may differ materially from listed corporates, both in terms of financial structure and the role of ESG-related risks. Consequently, the findings cannot be generalized to unrated or smaller firms. Future research could extend the analysis to private firms using internal bank ratings or alternative data sources.

- Reliance on a single rating agency

Due to licensing restrictions, this study uses only S&P Global credit ratings. While S&P is one of the three major global rating agencies (alongside Moody's and Fitch) and rating outcomes across agencies are highly correlated with each other, different agencies may nevertheless apply different methodologies to measure and weight ESG factors, and different timing of rating actions. Empirical studies have already proven this point. Early evidence by Ederington suggests that the bond ratings from Moody's and S&P are basically equivalent (Ederington, 1986). More recent studies, however, highlight systematic differences across agencies. For example, Livingston et al. showed that Moody's is more conservative in terms of bond ratings (Livingston, Wei, & Zhou, 2010). The results of our study may therefore be specific to S&P's rating practices and may not fully generalize to ratings from other agencies. Because the study does not capture potential divergences across agencies (e.g. split ratings), future research with access to multiple rating sources could compare whether the predictive power of CO2 emissions and ESG controversies varies across agencies.

- Reliance on Refinitiv ESG data

The ESG variables in this study are sourced from LSEG Refinitiv workplace, a widely used commercial ESG database. Several limitations inherent to this data should be addressed. First of all, ESG scores rely heavily on company self-disclosures. Under the EU's Non-Financial Reporting Directive (NFRD), non-financial disclosures were generally not required to be audited, raising concerns about reporting bias, selective disclosure, and potential greenwashing. According to the survey conducted by Accountancy Europe in 2020, 14 out of 26 European countries indicated to have companies seeking only voluntary independent assurance under NFRD (Accounting Europe, 2020). The European Parliament has also explicitly stated the shortcomings related to reliability in non-financial data during the implementation of the NFRD (Hahnkamper-Vandenbulcke, 2021). Additionally, the significant negative coefficient on the CO2 missing indicator in this study is direct evidence of such disclosure bias.

Another limitation is that Refinitiv applies a proprietary methodology for measuring, weighting, and aggregating ESG metrics, which could diverge from those used by other ESG data providers or credit rating agencies. Prior research documents evident divergence across major ESG rating providers. For example, Berg et al. has compared 6 ESG data providers (i.e. Sustainalytics, S&P Global, Moody's ESG, KLD, Refinitiv, and MSCI) and found that their correlations at the ESG level range from 0.38 to 0.71 with measurement and scope contributing the most divergence (Berg, Kölbel, & Rigobon, 2022). Finally, CO2 emissions data, while more objective than input-based scores may not be accurate depending on the measurement precision. We have seen companies such as BNP Paribas updating measurement methodologies from time to time.

Despite these limitations, Refinitiv is one of the most comprehensive and widely used ESG databases in academic research, and the robustness checks including the complete-case subsample and the exclusion of short-panel firms could mitigate concerns about data quality and missingness. Future research should cross-validate findings using alternative ESG data providers and, where possible, independently verified emissions data from mandatory disclosures under the EU's CSRD.

- Annual observation frequency

The analysis is constrained to yearly observations due to the availability of ESG data (typically reported annually) and the frequency of rating actions (most rating changes occur at discrete points). This limitation means that the study cannot capture companies' point-in-time financial and ESG performance within a single year, nor can it identify the precise timing of the controversy relative to a rating action. High-frequency (e.g. quarterly) data might reveal more immediate effects of ESG controversies or changes in emissions, but such data are not readily available for the moment. The findings should therefore be interpreted as reflecting relationships observed over yearly reporting intervals rather than short-term or intra-year dynamics.

- Missing data and imputation

Despite using a comprehensive panel, missing data remains as a common limitation in ESG and governance variables. To ensure sufficient sample size while addressing non-random missingness, the analysis employs forward-filling with a maximum window of two years, complemented by missing-indicator dummies. As demonstrated in section 7 model 5, the extreme outliers in the change of log value of CO_2 emissions are mostly driven by imputation effects rather than genuine economic changes. While robustness tests on the complete-case subsample without imputation confirm the main results, imputation inevitably introduces measurement error, which can bias coefficient estimates.

Furthermore, the CO_2 missing indicator is significant and negative, indicating that non-disclosing firms systematically differ from disclosing ones. Although this effect is explicitly modelled and interpreted as evidence of disclosure bias, the presence of residual confounding related to unobserved firm characteristics cannot be fully excluded. Future research with completer and more standardized ESG disclosures, particularly following the full implementation of the Corporate Sustainability Reporting Directive (CSRD) in Europe, could reduce reliance on imputation.

Please also note that in the last robustness analysis, the subsample potentially contains non-downgraded and investment grade firms. Ideally, these firms should be removed from the dataset to generate more concrete results. However, due to data unavailability and time constraints, we retain this potential improvement in future research.

- Country risk not fully controlled

As discussed in section 4, the main model does not include proprietary country risk measures used by S&P Global. Whereas macroeconomic controls (GDP growth, unemployment rate) and year fixed effects capture partially cross-country and common time-specific shocks, time-invariant country characteristics (e.g. legal environment and sovereign risk) are not directly controlled. The robustness test with the ND-GAIN vulnerability index addresses physical climate risk at the country level, but it is not a comprehensive substitute for broader country risk. Future research with access to such data could test whether the firm-level ESG effects remain robust after including country risk measures.

Despite the limitations discussed above, this study remains valid and contributes to meaningful evidence on the relationship between ESG factors and credit rating downgrades. The core findings that higher CO_2 emissions, more ESG controversies each independently predict higher downgrade risk survive multiple robustness checks. These results align with economic theory and together point to a consistent empirical relationship. Note that the findings from this study do not claim generalizability beyond European developed markets or to unrated firms, nor do it assert causality. Limitations are transparently acknowledged, and the methodology follows standard practices in empirical credit risk research.

11.Conclusion

This study examines the relationship between ESG factors and corporate credit rating downgrades, after controlling for firm-level financial metrics, governance attributes, industry and year fixed effects, and macroeconomic conditions. By using an unbalanced panel of publicly listed firms in European developed markets rated by S&P Global, we found empirical evidence that realized ESG outcomes, specifically CO_2 emissions, ESG controversies, and country-level physical climate vulnerability, are systematically associated with higher downgrade risk. Such findings remain robust across multiple model specifications and robustness tests.

One of the key findings of our study lies in highlighting the different predictory power of ESG measures of different natures. Commonly used input-based ESG scores related to management practices, workforce policies, and human rights do not exhibit predictive power for downgrades, whereas output-based indicators consistently do. This finding suggests that credit rating actions are more closely linked to realized and factual environmental and social results than to stated activities, policies or commitments by companies. Furthermore, the significance of missing CO_2 emission disclosure addresses the potential disclosure bias, probably driven by less scrutiny on non-disclosing firms. Meanwhile, the analysis also emphasizes the relevance of physical climate risk at the country level. The inclusion of the ND-GAIN Country Vulnerability Index shows that country-level exposure to physical climate risks has incremental explanatory power to downgrade risk. Eventually, by limiting the downgrades from lower-rated firms, we have observed a substantial improvement in model fit, which indicates a higher explanatory power of ESG factors for this subgroup.

Nevertheless, the study acknowledges several limitations. The reliance on publicly listed, rated firms in European developed markets and the exclusive use of S&P Global ratings constrain generalizability to a broader scope of corporates. Then the use of annually frequent explanatory data limits the ability to capture short-term dynamics. The study is also subject to data unavailability. Missing data and imputation may potentially introduce measurement errors. These limitations are transparently discussed, and the robustness analyses help mitigate their potential impact.

Overall, the study does not assert any causal relationships or claim universal generalizability. Specifically, it contributes robust associational evidence that output-based ESG metrics contain incremental predictive information for credit rating downgrades beyond traditional financial and macroeconomic controls. By clarifying which ESG dimensions appear most relevant for credit risk assessment, the findings have implications for banks, rating agencies, and researchers, particularly in the context of IFRS 9 SICR assessments and the ongoing evolution of ESG disclosure frameworks. Future research could build on this work by incorporating multiple rating agencies, higher-frequency data, downgrade severity measures, and more complete ESG disclosures emerging under the EU's CSRD.

12.References

Accounting Europe. (2020, February 19). *Towards reliable non-financial Information across Europe.* Retrieved from Accounting Europe: https://accountancyeurope.eu/publications/towards-reliable-non-financial-information-across-europe/

Adusei, M., Akomea, S. Y., & Poku, K. (2017). Board and management gender diversity and financial performance of microfinance institutions. *Cogent Business & Management*.

Ahold Delhaize. (2026). *thriving people*. Retrieved from https://aholddelhaize.com/about/thriving-people/

Altman, E. I. (1968). Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. *The Journal of Finance*, 589-609.

Amstad, M., Packer, F., & Shek, J. (2018). *Does sovereign risk in local and foreign currency differ?* BIS.

Atif, M., & Ali, S. (2021). Environmental, social and governance disclosure and default risk. *Business Strategy and the Environment*, 3937–3959.

Baldan, C., Zen, F., & Targhetta, M. (2026). The Impact of ESG Factors on Corporate Credit Risk: An Empirical Analysis of European Firms Using the Altman Z-Score. *Accounting and Auditing*.

Ballester, L., González-Urteaga, A., & B. M. (2020). The role of internal corporate governance mechanisms on default risk: A systematic review for different institutional settings. *Research in International Business and Finance*.

Bauer, R., & Hann, D. (2010). Corporate Environmental Management and Credit Risk.

BCBS. (2021). *Climate-related risk drivers and their transmission channels.* BIS.

Beaver, W. H. (1968). Market Prices, Financial Ratios, and the Prediction of Failure. *Journal of Accounting Research*, 179-192.

Berg, F., Kölbel, J. F., & Rigobon, R. (2022). Aggregate Confusion: The Divergence of ESG Ratings. *European Finance Review*.

Bessis, J. (2010). *Risk Management in Banking.* New York: John Wiley & Sons Inc.

BIS. (2017). *IFRS 9 and expected loss provisioning – Executive Summary .*

BNP Paribas Group. (2025). *Universal Registration Document and Annual Financial Report.*

Bo, H., & Driver, C. (2012). Agency Theory, Corporate Governance and Finance. *Handbook on the Economics and Theory of the Firm*.

Breuer, M., & deHaan, E. (2024). Using and Interpreting Fixed Effects Models. *Journal of Accounting Research*.

Campbell, J. Y., Hilscher, J., & Szilagyi, J. (2008). In Search of Distress Risk. *The Journal of Finance*.

Capasso, G., Gianfrate, G., & Spinelli, M. (2020). Climate change and credit risk. *Journal of Cleaner Production*.

Carbone, S., Giuzio, M., Kapadia, S., Krämer, J. S., Nyholm, K., & Vozian, K. (2021). The low-carbon transition, climate commitments and firm credit risk. *Statistics for Sustainable Finance.* Paris: BIS.

Castro, V. (2013). Macroeconomic determinants of the credit risk in the banking system: The case of the GIPSI. *Economic Modelling*.

Chava, S. (2011). *Environmental Externalities and Cost of Capital.* Georgia Institute of Technology.

Chen, D. (2014). The Non-monotonic Effect of Board Independence on Credit Ratings. *Journal of Financial Services Research*, 145-171.

Dang, C., Li, Z. (., & Yang, C. (2018). Measuring firm size in empirical corporate finance. *Journal of Banking & Finance*, 159-176.

Dang, M., Puwanenthiren, P., Mazur, M., Hoang, V. A., & Nadarajah, S. (2025). Firm complexity and credit ratings. *International Review of Financial Analysis*.

Demirhan, H. (2024). Financial Anomalies and Creditworthiness: A Python-Driven Machine Learning Approach Using Mahalanobis Distance for ISE-Listed Companies in the Production and Manufacturing Sector. *Journal of Financial Risk Management*.

Dodd, O., Kalimipalli, M., & Chan, W. (2021). Evaluating corporate credit risks in emerging markets. *International Review of Financial Analysis*.

Dumrose, M., & Höck, A. (2023). Corporate Carbon-Risk and Credit-Risk: The Impact of Carbon-Risk Exposure and Management on Credit Spreads in Different Regulatory Environments. *Finance Research Letters*.

ECB. (2020). *Guide on climate-related and environmental risks: Supervisory expectations relating to risk management and disclosure.*

ECB. (2024). *IFRS 9 overlays and model improvements for novel risks.*

Ederington, L. H. (1986). Why Split Ratings Occur. *Financial Management*, 37–47.

Edmans, A. (2011). Does the stock market fully value intangibles? Employee satisfaction and equity prices. *Journal of Financial Economics*.

Ehlers, T., Elsenhuber, U., Jegarasasingam, A., & Jondeau, E. (2022). *Deconstructing ESG scores: How to invest with your own criteria.* BIS.

European Commission. (n.d.). *The European Commission's response to the coronavirus crisis*. Retrieved from European Commission: https://commission.europa.eu/strategy-and-policy/coronavirus-response_en

Faisal, M., & Wenno, M. (2023). Analysis of the impact ownership structure and firm size on credit risk of private banks period 2018-2022. *Jurnal Ekonomi*.

Ferriani, F., & Pericoli, M. (2024). *ESG risks and Corporate Viability: Insights from Default Probability Term Structure Analysis.*

Ferriani, F., & Pericoli, M. (2026). ESG risks and corporate viability: Insights from default probability term structure analysis. *International Review of Financial Analysis*.

Figlewski, S., Frydman, H., & Liang, W. (2012). Modeling the Effect of Macroeconomic Factors on Corporate Default and Credit Rating Transitions. *International Review of Economics & Finance*.

Fitch Ratings. (2023, October 24). *Over Half of Corporates Facing Climate-Related Downgrades by 2035 Are Investment Grade*. Retrieved from Fitch Ratings: https://www.fitchratings.com/research/corporate-finance/over-half-of-corporates-facing-climate-related-downgrades-by-2035-are-investment-grade-24-10-2023

García, C. J., Herrero, B., & Morillas, F. (2022). Corporate board and default risk of financial firms. *Economic Research-Ekonomska Istraživanja*, 511–528.

Garschagen, M., Doshi, D., Reith, J., & Hagenlocher, M. (2021). Global patterns of disaster and climate risk—an analysis of the consistency of leading index-based assessments and their results. *Climatic Change*.

Gopalan, R., Song, F., & Yerramilli, V. (2010). Debt Maturity Structure and Credit Quality. *The Journal of Financial and Quantitative Analysis*.

Hahnkamper-Vandenbulcke, N. (2021). *Briefing: Non-financial Reporting Directive.* European Parliamentary Research Service .

Hasnaoui, J. A., & Hasnaoui, A. (2022). How does human capital efficiency impact credit risk?: the case of commercial banks in the GCC. *The Journal of Risk Finance*, 639-651.

Hemmert, G. A., Schons, L. M., & Schimmelpfennig, H. (2016). Log-likelihood-based Pseudo-R2 in Logistic Regression: Deriving Sample-sensitive Benchmarks. *Sage Journals*.

Horrigan, J. O. (1966). The Determination of Long-Term Credit Standing with Financial Ratios. *Journal of Accounting Research*, 44-62.

Hossain, M. I., Mia, M. A., & Hooy, C.-W. (2023). Employee turnover and the credit risk of microfinance institutions (MFIs): International evidence. *Borsa _ Istanbul Review*, 936–952.

Huang, H. H., Kerstein, J., & Wang, C. (2018). The impact of climate risk on firm performance and financing choices: An international study. *Journal of International Business Studies*, 633-656.

Hugon, A., & Law, K. (2018). Impact of Climate Change on Firm Earnings. *SSRN Electronic Journal*.

Isaev, M., & Masih, M. (2017). Macroeconomic and bank-specific determinants of different categories of non-performing financing in Islamic banks: Evidence from Malaysia. *MPRA*.

King, G., & Zeng, L. (2003). Logistic Regression in Rare Events Data. *Journal of Statistical Software*.

Kladakis, G., & Skouralis, A. (2023). Credit rating downgrades and systemic risk. *Journal of International Financial Markets, Institutions and Money*.

Kladakis, G., & Skouralis, A. (2024). Credit rating downgrades and systemic risk. *Journal of International Financial Markets, Institutions and Money*.

Levett, B. (2025, February 12). *Does Board Independence Matter for Performance? It's Complicated.* Retrieved from Neutral Alpha: https://www.neuralalpha.com/research-and-insights/does-board-independence-matter-for-performance-its-complicated

Livingston, M., Wei, J., & Zhou, L. (2010). Moody's and S&P Ratings: Are They Equivalent? Conservative Ratings and Split Rated Bond Yields. *Journal of Money, Credit and Banking*, 1267 - 1293.

Lohmann, C., Möllenhoff, S., & Lehner, S. (2025). On the Relationship Between Financial Distress and ESG Scores. *Wiley*.

LSEG. (2026). *DataItemBrowser*. Retrieved from Workspace Refinitiv: https://workspace.refinitiv.com/web/Apps/DataItemBrowser/

Lui, A. (2021, November 20). *Does Size Matter? Incorporating company size in credit risk assessment.* Retrieved from S&P Global: https://www.spglobal.com/market-intelligence/en/news-insights/research/does-size-matter-incorporating-company-size-in-credit-risk-assessment

Ma, Y., & Hu, W. (2026). Does ESG performance mitigate corporate debt default risk? Evidence from emerging markets. *Finance Research Letters*.

March, J. G., & Simon, H. A. (1958). *Organizations.* New York: Wiley.

McFadden, D. (1979). Quantitative Methods for Analysing Travel Behaviour of Individuals: Some Recent Developments. *Behavioural Travel Modelling*, 279–318.

Mcwayizeni Thwala, C. (2016). The sensitivity of bank credit risk indicators to macroeconomic variables. *University of the Witwatersrand Business* .

Mehmood, A., & Luca, F. D. (2023). How does non-interest income affect bank credit risk? Evidence before and during the COVID-19 pandemic. *Finance Research Letters*.

Nguyen, J. H., & Phan, H. V. (2020). Carbon risk and corporate capital structure. *Journal of Corporate Finance*.

OECD. (2025). *Behind ESG ratings: unpacking sustainability metrics.* OECD.

Oliveira, N. A., & Basso, L. F. (2023). The Impact of Credit Ratings on Financial Performance (ROA) and Value Creation (Tobin's Q). *Chinese Business Review* .

Petersen, M. A. (2008). Estimating standard errors in finance panel data sets: comparing approaches. *The Review of Financial Studies*, 435–480.

Pirson, M. A., & Turnbull, C. S. (2018). Decentralized Governance Structures Are Able to Handle CSR-Induced Complexity Better. *Business & Society*, 929-961.

Rapposelli, A., Birindelli, G., & Modina, M. (2023). The relationship between firm size and efficiency: why does default on bank loans matter? *Quality & Quantity*, 3379–3401.

Rosa, F. L., & Bernini, F. (2022). ESG controversies and the cost of equity capital of European listed companies: the moderating effects of ESG performance and market securities regulation. *International Journal of Accounting and Information Management*.

S&P Global. (2024). *Criteria | Corporates | General: Corporate Methodology.* Retrieved from Criteria | Corporates | General: Corporate Methodology: https://www.spglobal.com/ratings/en/regulatory/article/-/view/type/HTML/id/3415681

S&P Global. (2026). *Understanding Credit Ratings*. Retrieved from Understanding Credit Ratings: https://www.spglobal.com/ratings/en/credit-ratings/about/understanding-credit-ratings

Samaniego-Medina, R., & Giraldez-Puig, P. (2022). Do Sustainability Risks Affect Credit Ratings? Evidence from European Banks. *The AMFITEATRU ECONOMIC journal*, 720-720.

Scandizzo, S., & Hughes, T. (2025). *Climate-Related Financial Risk: Arguments, Analytics and Regulations.* Risk Books.

Schneider, T. E. (2010). *Is environmental performance a determinant of bond pricing? Evidence from the U.S. pulp and paper and chemical industries.* University of Alberta.

Schwaab, B., Koopman, S. J., & Lucas, A. (2016). *Global credit risk: world, country and industry factors.* ECB.

Sharfman, M. P., & Fernando, C. S. (2008). Environmental Risk Management and the Cost of Capital. *Strategic Management Journal*, 569-592.

Shumway, T. (2001). Forecasting Bankruptcy More Accurately: A Simple Hazard Model. *The Journal of Business*, 101-124.

Sindaco, M. (2021, March 31). *The Evolution of ESG Factors in Credit Risk Assessment: Social Issues.* Retrieved from S&P Global: https://www.spglobal.com/market-intelligence/en/news-insights/research/the-evolution-of-esg-factors-in-credit-risk-assessment-social-issues

Smith, A. (1776). *An Inquiry into the Nature and Causes of the Wealth of Nations.* London: W. Strahan and T. Cadell.

Spicer, B. (1978). Investors, Corporate Social Performance and Information Disclosure: An Empirical Study. *Accounting Review* .

Supiyadi, D., & Novita, I. (2023). The Effect of Firm Size, Credit Risk, Interest Rates, and Liquidity on Bank Profitability: Study on State-Owned Banks in Indonesia. *Jurnal Ilmu Keuangan dan Perbankan*.

Switzer, L. N., Wang, J., & Tu, Q. (2016). Corporate Governance and Default Risk in Financial Firms Over the Post Financial Crisis Period: International Evidence. *SSRN Electronic Journal*.

Ting, W. (2011). Top management turnover and firm default risk: Evidence from the Chinese securities market. *China Journal of Accounting Research*.

Tutar, H., Štreimikienė, D., & Kyriakopoulos, G. L. (2026). Decarbonization Pathways in the European Union: Sectoral Contributions to CO2 Emissions Reductions (2000–2022). *Environments*.

Wang, C.-J., & Lin, J.-R. (2010). Corporate Governance and Risk of Default. *International Review of Accounting, Banking and Finance*, 1-27.

Wang, L. (2020). Does The Enhancement of Profitability Necessarily Reduce Bank Credit Risk? *International Journal of Business and Management*.

Wattanatorn, W. (2025). The role of climate exposure and ESG in forward-looking default risk: A global perspective. *Journal of Multinational Financial Management*.

Yeboah, A., Mogre, S. a., Menzo, D. a., & Prince, B. (2024). Beyond the Numbers: Social Factors in Credit Risk. *Munich Personal RePEc Archive*.

Zadeh, F. O., & Eskandari, A. (2012). Firm Size As Company's Characteristic and Level of Risk Disclosure: Review on Theories and Literatures. *International Journal of Business and Social Science*.

Zhang, J., Andreeva, G., & Dong, Y. (2025). *Are ESG Responsible Companies Loss Responsible? Modelling LGD with ESG Information.* University of Edinburgh Business School.

13.Tables and Figures