École des Hautes Études Commerciales

HEC MONTRÉAL

Thesis submitted in partial fulfillment of the requirements for the Master of Science in Management

(Finance)

The Impact of the Environmental Score on Systemic Risk in the Banking Industry, post COP21

2024/2025

Written by : M. Norman Viron

Supervised by : M. Jakub Hajda

Résumé

Cette thèse étudie la relation entre la performance environnementale des banques et leur contribution au risque systémique dans l'industrie bancaire dans le contexte post-COP21, où les questions environnementales sont devenues une grande partie de la finance sur la scène mondiale. Pour cela, nous utilisons un panel de 113 banques de toutes les régions du monde, et nous utiliserons différents modèles et régressions pour évaluer l'impact que le score environnemental a sur les différentes mesures de risque systémique que nous utiliserons. Ce mémoire révélera des idées nuancées similaires à la littérature existante et empirique. Ces résultats soulignent que toute initiative environnementale peut réduire le risque systémique par rapport aux banques qui n'en ont aucun. Cependant, l'ampleur des efforts environnementaux joue un rôle crucial dans la formation du profil de risque des banques, révélant souvent une certaine pression réglementaire et actionnariale. Ce travail contribue à la littérature croissante sur la finance durable et l'ESG, fournissant des preuves aux régulateurs et aux investisseurs cherchant à aligner la performance environnementale avec la stabilité financière.

Mots-clés : Score environnemental (E score), ESG, Risque systémique, Effet de contagion, COP21

Abstract

This thesis investigates the relationship between the Environmental performance of banks and their systemic risk contribution within the banking industry in the post-COP21 context, where Environmental issues became a big part of finance on the global scene. For that, we use a panel of 113 banks from all regions in the world, and we will employ different financial econometrics models and regressions to assess the impact that the Environmental scores (E Score) have on the different systemic risk measures we will use. This thesis will reveal nuanced insights similar to existing and empirical literature. These results highlight that any environmental initiative can reduce systemic risk compared to banks with none. However, the magnitude of environmental efforts plays a crucial role in shaping the risk profile of banks, often revealing some regulatory and shareholder pressure. This work contributes to the growing literature on sustainable finance and ESG, providing evidence for regulators and investors seeking to align environmental performance with financial stability.

Keywords : E score, ESG, Systemic risk, Spillover, COP21

Acknowledgments

I would like to thank my supervisor Mr. Jakub Hajda for his support throughout this work, his help, and his patience. I would also like to emphasize his quick response and availability.

I want to thank my family, my sisters, and my friends for supporting me during these 5 years of study abroad, for pushing me to do it, and for believing in me.

And, I finally want to thank all the professors who accompanied me during my journey at HEC Montreal. I want to say thank you to HEC Montreal as well and on a broader spectrum, thank you to all the professors in France, in the USA, in Canada, who have brought me here.

Table of Contents

Introduction

Over the past decade, environmental issues have risen to the forefront of global finance, reshaping investment and risk management practices. This transformation has been particularly pronounced in the banking sector, where institutions are now evaluated not only on financial metrics but also on their environmental, social, and governance (ESG) performance, with growing pressure from investors. Among these three pillars, the environmental dimension has gained relevance in the post-COP21 era, with clear worldwide objectives set. The Paris Agreement, adopted in 2015, sets global targets for carbon neutrality and also calls for realigning financial flows to support climate-resilient and low-emission development pathways (UNFCCC, 2015). As a result, climate-related risks are now understood as long-term externalities and material financial risks that can manifest through market volatility and systemic disruptions.

Banks are at the center of this evolving landscape because, as capital intermediaries, they are exposed to environmental risks through their credit portfolios, investment allocations, and operational dependencies. These exposures create vulnerabilities both at the micro and macro levels. At the micro level, individual institutions may face losses from stranded assets or regulatory sanctions. At the macro level, correlated exposures and liquidity shortages can lead to system-wide stress. This dual exposure has led to growing interest in environmental scores, which attempt to quantify environmental performance across multiple dimensions such as emissions disclosure, energy usage, water management, and green financing (Gunarathne et al., 2023; OECD, 2020). These E scores are becoming a standard metric for ESG investment decisions and are also increasingly used by regulators and central banks in supervisory frameworks, stress testing, and macroprudential oversight.

In this context, this thesis's central question is whether higher E scores are associated with reduced systemic risk in the banking industry. The answer to this question is not straightforward. Some prior research (Aevoae et al., 2023; Gidage et al., 2024; Capotă et al., 2022; Cerqueti et al., 2021) suggests that stronger ESG practices can reduce systemic risk impact, but other recent studies (Dong [2023, 2025], Pistolesi and Teti [2024]) also warn that rapid and homogeneous adoption of ESG strategies may create new forms of systemic risk, including portfolio crowding, correlated exposures, and green asset bubbles. Furthermore, most empirical studies have focused on aggregated ESG scores or governance indicators. In contrast, the specific role of the E score remains underexplored, especially in a global banking sample.

This thesis seeks to fill that gap by investigating the link between Bloomberg Environmental Scores and the systemic risk contribution of banks, using a balanced panel of 113 publicly listed banks across seven global regions from 2012 to 2023. The analysis concentrates on the post-COP21 period, where regulatory pressure and market incentives to adopt environmental strategies significantly accelerated. Using five systemic risk metrics, namely Value at Risk (VaR), Expected Shortfall (ES), Marginal Expected Shortfall (MES), Conditional Value at Risk (CoVaR), and Delta CoVaR, this study evaluates whether higher E scores are associated with more stable and less systemically risky banks.

To examine this relationship, our research will mainly focus on two hypotheses, we will try to understand if banks with higher Environmental Scores exhibit lower systemic risk contributions, as measured by the systemic risk measures. This hypothesis builds on the assumption that environmental engagement strengthens long-term resilience. It is supported by empirical research showing that strong ESG performance is linked to lower systemic risk, as Aevoae et al. (2023) find that higher ESG scores significantly reduce banks' contribution to system-wide financial distress, as measured by ΔCoVaR. Although governance plays a leading role, the overall ESG composite, including environmental performance, contributes to resilience. Gidage et al. (2024) similarly report that firms with higher ESG scores have lower equity betas, indicating reduced sensitivity to market-wide shocks. At the fund level, Capotă et al. (2022) show that ESG funds experience more stable flows during crises, while Cerqueti et al. (2021) find that high-ESG funds are less vulnerable to fire-sale contagion. These findings support the idea that stronger environmental performance enhances institutional resilience and lowers systemic risk contributions, consistent with MES, CoVaR, and ΔCoVaR dynamics. Then we will investigate if the introduction of environmental practices following the 2015 COP21 agreement has a positive effect on reducing systemic risk, regardless of the absolute level of the E score. This hypothesis is supported by recent research showing that even limited or initial environmental commitments can serve as effective signals for reducing financial risk. Ilhan et al. (2021) and Krueger et al. (2020) emphasize the role of environmental disclosures and initiatives as threshold effects, where their presence, rather than their magnitude, significantly reduces perceived and actual financial risk. Similarly, in the banking context, Monasterolo and Raberto (2019) demonstrate that incorporating environmental risk awareness can offset systemic exposure, supporting the notion that the adoption of environmental practices matters more than their depth in certain contexts. The empirical analysis is grounded in a robust econometric framework where we conduct different panel regressions and Difference-in-Differences (DiD) analysis, but the results are also backed with a quantile regression at different percentiles. These models are conducted with daily log returns for each bank which, and control variables including macroeconomic banks and bank specific variables, to reduce omitted variable bias and capture some unexplained movements. E scores were obtained from Bloomberg and were lagged one year in the analysis to reflect the realistic market adjustments and data publication timeline.

In addition to these regressions, several robustness checks were conducted to isolate the actual effect of E scores. These include excluding the year 2020, which was dominated by the COVID-19 crisis and may have introduced large market distortions and volatility, also limiting the sample to banks in North America and Europe, which are two of the regions exhibiting the highest average E scores and are more integrated into global financial markets, possibly reflecting the real effects of the E Score while some regions could bias results as they are less integrated into the financial world (see the preliminary analysis of systemic risk measures and average returns). The results provide a complex but revealing answer to the research question as when E scores are used as a continuous variable in panel regressions, the results suggest that higher E scores are associated with higher systemic risk, especially in measures like MES, CoVaR, and ΔCoVaR. However, the DiD analysis offers a different interpretation. By comparing banks that adopted environmental practices after COP21 to those that did not, the model finds that even minimal environmental engagement leads to a relative reduction in systemic risk. Based on these double faced results, we conducted a quantile regression, on three q levels (0.1, 0.25 and 0.75) and the results also show that E scores play a protective role particularly in the lower tail of the risk distribution, meaning during periods of financial stress, while the effect is negligible in more stable conditions. That is something we also observed when running another DiD regression, comparing banks in North America and Europe, with the treatment being the period post EU Taxonomy of 2020. In a tough period where systemic risk was high, banks receiving the treatment experienced a lower increase in systemic risk compared to the banks in

North America, at the bank level and also in the spillover, showing that the E Score reduces the systemic risk increase at significant levels.

Finally, by focusing again only on European and North American banks but for the panel regression, the results become clearer and more consistent. In this sub-sample, higher E scores are significantly associated with lower systemic risk for several measures, and this remains true even after excluding the year 2020 for robustness verification. These results suggest that the relationship between environmental performance and systemic stability is not linear and may depend on the regulatory context or region. Overall, the thesis contributes to academic and regulatory discussions on sustainable finance, by showing that environmental practices and E Score can be both a signal of resilience and also reduce the effect of systemic risk, especially on the risk transferred to other banks, mainly in the lower tails.

Chapter 1 : Literature Review

This first chapter introduces several key concepts that underpin our research, including the Environmental score, systemic risk, and the landmark COP21 agreement. We also review existing literature and related studies to provide a deeper understanding of our topic and to explain the rationale behind our choice of research focus.

1.1 Institutional Background

In this first chapter, dedicated to the literature review, we will first explore and highlight the key elements of the subject in order to provide a thorough understanding of the research and analysis that follow. The aim is to lay the groundwork for a comprehensive grasp of the subject as a whole.

1.1.1 Systemic Risk

Systemic risk refers to the potential for a breakdown of an entire financial system or market, as opposed to the failure of any single entity in isolation. In essence, it is the risk that localized problems such as the collapse of one major institution could trigger cascading failures across many institutions, ultimately threatening the stability of the broader financial system (International Monetary Fund [IMF], Financial Stability Board [FSB], & Bank for International Settlements [BIS], 2009). One authoritative definition describes systemic risk as "a risk of disruption to financial services that is caused by an impairment of all or parts of the financial system and has the potential to have serious negative consequences for the real economy" (IMF, FSB, & BIS, 2009, p. 2). This definition highlights that what makes a risk systemic is not merely large losses at one firm, but the negative externalities, the spillover effects, that such an event could impose on other institutions and on economic activity at large. Because of these spillovers, systemic risk is a primary concern for regulators and central banks, who aim to maintain financial stability and prevent scenarios in which confidence in the financial system erodes (Acharya, Pedersen, Philippon, & Richardson, 2017).

Mechanisms of systemic risk often involve strong interlinkages among financial institutions and markets. For example, banks and other intermediaries are connected through interbank lending, payment networks, and common asset holdings; if one key player fails, its counterparties may incur losses, liquidity can dry up, and a chain reaction of distress can occur (Brunnermeier, 2009). Also, the failure of a large firm (such as a major bank or insurer) can lead to a spreading failure that undermines trust and stability across financial markets. These contagion dynamics mean that systemic risk is fundamentally about network effects and collective behavior, not just individual losses (Allen & Gale, 2000). Policymakers often summarize this by saying certain institutions are "too big to fail" or "too interconnected to fail," since their distress could pull others down and endanger the entire system (Mishkin, 2011). The 2007–2009 Global Financial Crisis is a perfect example of systemic risk in action: people unable to repay their mortgage led to problems in the subprime mortgage market, which were spread to major global banks and markets, freezing credit flows and precipitating the worst recession in decades (Gorton, 2010).

1.1.2 Systemically Important Financial Institutions (SIFI)

In the wake of recurring financial crises, regulators have introduced the label "Systemically Important Financial Institution" (SIFI) to denote firms whose distress or failure could gravely threaten the entire financial system. These are institutions as mentioned earlier seen as too big to fail. According to the Financial Stability Board and other standard-setting bodies, a financial institution is considered systemically important if "its distress or disorderly failure would cause significant disruption to the wider financial system and economic activity, due to its size, complexity, and interconnectedness" (Financial Stability Board, 2010). In practical terms, SIFIs are those banks, insurance companies, or other financial firms whose collapse could initiate a systemic crisis. The negative externalities of their failure (spillovers to other institutions, markets, and ultimately the real economy) are so severe that authorities have a strong incentive to prevent such an event.

This was exemplified in 2008 when governments intervened to support large banks and insurers; letting them fail freely was judged too damaging for global finance (Gorton, 2010; Mishkin, 2011). However, the bailout of too-big-to-fail institutions creates a moral hazard, since investors might expect future rescues and SIFIs might take on excessive risks under the assumption of implicit guarantees (Acharya, Cooley, Richardson, & Walter, 2010).

Identifying and regulating SIFIs has become a cornerstone of post-2008 financial reforms. Internationally, the Basel Committee on Banking Supervision and the FSB now annually designate Global Systemically Important Banks (G-SIBs) and Global Systemically Important Insurers (G-SIIs), publishing lists of those entities (Basel Committee on Banking Supervision [BCBS], 2013). These designations use an indicator-based methodology: banks are scored on factors including their total assets (size), interconnectedness (e.g. interbank exposures), lack of substitutability (whether they provide critical services that would be hard for others to quickly replace), global (cross-jurisdictional) activity, and complexity (FSB, 2011). Institutions with the highest scores are placed into buckets reflecting their systemic importance, which in turn determines additional regulatory requirements. In particular, SIFIs are required to hold higher capital buffers than ordinary institutions to increase their resilience. They are also subject to more intensive supervision and must formulate “living wills” or resolution plans to ensure that, if they do fail, they can be wound down in an orderly way (BCBS, 2013). The overall goal of this framework is to reduce both the probability of failure of SIFIs and the systemic impact if one does fail, to avoid consequences of the 2008 where governments had to largely intervene. The goal is to mitigate the threat they pose to financial stability (FSB, 2011).

It is important to note that SIFI status is not limited to banks and insurers. In principle, any financial institution (including investment firms, large asset managers, etc.) could be deemed systemically important if its failure would impair the system (FSB, 2013). Debates continue about how to identify non-bank, non-insurer SIFIs, given the diversity of business models in areas like hedge funds or market infrastructure. Nonetheless, the consensus is that effective macroprudential policy must pay particular attention to a relatively small number of critical nodes in the financial network. The SIFI framework is a targeted post 2008 crisis approach to contain systemic risk, make the global system more stable and avoid similar events.

1.2 ESG Metrics

Sustainable finance has brought ESG metrics to the forefront of risk analysis in banking. These scoring systems distill a firm's performance on a broad range of sustainability criteria into quantitative ratings, and investors and regulators increasingly use these scores to gauge long-term risks.

1.2.1 The Environmental Component

Among the three pillars of ESG, the environmental component has gained particular prominence in banking risk discussions. Climate change and environmental degradation pose tangible financial risks for banks at the micro level, such as credit and market risks that threaten individual institutions, and at the macro level through broader systemic vulnerabilities. Financial authorities now recognize that climate-related events represent significant threats to financial stability. Climate risks can materialize in several ways : physical risks stem from direct losses linked to extreme weather events, such as floods or hurricanes, which can damage assets or disrupt business operations and transition risks arise when policies or technologies evolve toward a low-carbon economy, potentially stranding assets in carbon-intensive sectors. Banks also face liability risks if held legally accountable for environmental damages or poor stewardship (Carney, 2015). These climate risks typically flow through banks' lending and investment portfolios. As an illustrative example, a bank that has lent heavily to coal mining operations or invested in coastal real estate developments could see heightened credit risk if carbon regulations tighten or rising sea levels erode property values. In response to these evolving threats, regulators have begun to scrutinize how banks incorporate environmental risks as we can notice, in 2022, the European Central Bank conducted its first climate stress test to measure banks' readiness for climate-related shocks (European Central Bank, 2022).

Integrating environmental criteria into risk management has allowed banks to forward-think, it may involve assessing borrowers' carbon footprints or steering more financing to green projects, which can in the end strengthen a bank's long-term stability and reduce the overall risk. For example, Gangi et al. (2019) find that banks in 35 countries with higher environmental sensitivity, such as lending to sustainable initiatives or robust climate disclosure, exhibit lower risk exposures and more stable balance sheets. Similarly, Chiaramonte et al. (2021) show that including environmental and social factors in banks' stress testing models improves their ability to anticipate distress, implying that traditional risk models benefit from ESG integration. These findings are supported by evidence from loan pricing trends. Delis et al. (2019) report that after the 2015 Paris Agreement, banks started charging higher interest rates to carbon-intensive borrowers, reflecting an adjustment for transition risk in loan pricing. At the same time, research by Kacperczyk and Peydró (2022) shows that banks adopting net-zero targets actively reduce credit to high-emission sectors.

All these developments reinforce the view that environmental considerations are not simply ethical choices but financially relevant factors. This E factor is now an integrated part of traditional risk management and of how banks assess and control their exposures in an era of climate-driven change.

1.2.2 ESG Reputational Risks

ESG factors influence banks' reputational risk, an often intangible but critical aspect of risk management. Reputational risk is the possibility that a company will face negative public or investor perception, which can harm its operations or financial standing. In the context of ESG criteria, reputational risk can arise if a bank mismanages these factors and is perceived negatively. For example, this risk can be materialized through the financing of controversial projects (such as fossil fuels or deforestation), which may provoke public protests or increased regulatory scrutiny, ultimately impacting the bank’s profitability and funding costs. Modern media and civil society quickly spotlight ESG-related missteps, meaning banks face heightened scrutiny on climate action, diversity, and corporate ethics. Studies indeed show that adverse ESG incidents have measurable financial repercussions. Wong and Zhang (2022) find that media coverage of ESG-related bad news (scandals or negative disclosures) significantly negatively impacts a firm's market valuation. In other words, when banks (or firms generally) make headlines for ESG failures, their stock prices and reputations suffer. This media-driven reputational damage can increase a bank's cost of capital and even trigger liquidity pressures if stakeholders lose trust (Wong & Zhang, 2022).

In empirical terms, ESG controversies erode the risk benefits of otherwise strong ESG performance. A bank might have high ESG ratings on paper, but if it experiences significant controversies (e.g., fraud, environmental accidents, governance scandals), the anticipated risk reduction from "good" ESG can vanish. Galletta and Mazzù (2023) provide evidence of this dynamic: they found that banks with fewer ESG controversies tend to have significantly lower risk, for instance, such banks showed lower risk-weighted assets and higher Z-scores, indicating greater stability. Shakil (2021) observes a similar pattern in the oil and gas industry: ESG controversies (like environmental spills or social conflicts) materially weaken the inverse relationship between ESG performance and risk-taking. In essence, controversies introduce reputational and legal risk spikes that can cancel out the risk mitigation effect of otherwise strong ESG policies. This has led rating agencies to incorporate controversy tracking into their scoring. For instance, Refinitiv's ESG scoring methodology includes a separate ESG Controversies Score, which penalizes firms for adverse ESG events and is averaged with the base ESG score to produce an "ESG Combined Score" (Refinitiv, 2023). The logic is that a company's sustainability profile should reflect not only its policies and disclosures but also its actual behavior and any transgressions that could signal management failures. A severe controversy can quickly damage a bank's reputation and invite costly regulatory penalties, as seen in past money laundering cases or unethical sales practices in the banking sector. Managing ESG reputational risk requires banks to maintain transparency and adapt to ESG-related issues. By doing so, they can protect their brand value and avoid the accumulation of consequences (from customer distrust to regulatory sanctions) that can follow an ESG scandal.

1.3 COP21 - Paris Agreement

Before COP21, the Kyoto Protocol (adopted in 1997 and entered into force in 2005) represented the first binding international attempt to reduce greenhouse gas (GHG) emissions. It set quantitative targets for industrialized countries based on the principle of common but differentiated responsibilities. Despite some notable achievements, particularly in Europe, the protocol showed apparent limitations: it applied to only a small portion of global emissions. It failed to engage major emerging emitters (only 55 countries participated) (UNFCCC, 1998).

The 2015 Paris Conference marked a decisive turning point in this context as the agreements were unanimously adopted by the 196 Parties to the United Nations Framework Convention on Climate Change (UNFCCC, 2015), established a universal framework for all countries, without the former North-South dichotomy, and based on nationally determined contributions (NDCs). Its core objective is to limit the global average temperature rise to well below 2°C above pre-industrial levels while pursuing efforts to restrict the increase to 1.5°C (UNFCCC, 2015, Art. 2). One of its most crucial yet often underestimated elements is the ambition to realign global financial flows. Article 2.1(c) introduces an unprecedented goal: "making finance flows consistent with a pathway towards low greenhouse gas emissions and climate-resilient development" (UNFCCC, 2015, Art. 2.1(c)). This provision signals a pattern shift; climate change is now seen as a structural lever for transforming the global economy and financial system, which changes from the "simple" environmental issue view (Carney, 2015).

The Agreement calls on developed countries to mobilize "jointly USD 100 billion per year" by 2020 to support mitigation and adaptation efforts in developing countries, with this target intended to increase in the following years (UNFCCC, 2015). It also emphasizes the need for predictable, transparent, and sustained finance involving public and private actors. The challenge is to direct capital toward green infrastructure, energy transition, and technological innovation (European Commission, 2021).

As such, sustainable finance has emerged as a strategic pillar of the international climate response. The Agreement calls for enhanced transparency in climate finance flows (Articles 9 and 13), establishing technology support mechanisms, and increased cooperation among financial institutions, governments, civil society, and the private sector (UNFCCC, 2015). Initiatives such as the Net-Zero Banking Alliance (NZBA) and significant banking commitments to green finance reflect this momentum (UNEP FI, 2021).

Crucially, the COP21 agreements elevated finance to a central operational lever in the climate transition. They invited financial institutions to rethink their strategies, risk assessment tools, and investment policies. This has led to an increased focus on ESG considerations by companies and banks, particularly on the E score, as firms now seek to disclose their sustainability efforts more transparently (Carney, 2015). Initially, many of these commitments were voluntary. However, they are increasingly being institutionalized through stricter regulations and standardized reporting frameworks (such as CSRD and SFDR) but also because of new "rules" such as the EU taxonomy and structured non-financial indicators (European Commission, 2022).

As a brief summary, the Paris Agreement put finance back at the heart of the climate concerns, requiring a redirection of capital flows and a renewed focus on long-term risks. It has reshaped the ESG landscape, making E score central to evaluating resilience and responsibility in the banking sector if we want to reach the objectives set by these agreements.

1.4 Climate-Related Risks and Transition Dynamics

1.4.1 Physical Climate Risks

Physical climate risks refer to the tangible and diverse impacts of climate change that can directly affect a company's assets, operations, and financial stability (TCFD, 2017). These risks arise from acute events such as hurricanes, storms, heat waves, and droughts, as well as chronic processes like rising temperatures and sea-level increases. Banks with significant real estate exposures or loan portfolios tied to vulnerable sectors are particularly at risk of incurring losses when these physical disasters strike.

For example, the BNP Paribas Risk presentation (Guay & Sposito, 2024) emphasizes that assessing these risks demands forward-looking climate models that extend beyond historical weather data to incorporate future climate scenarios and intensities. Such models are crucial because extreme weather events can disrupt operations, cause business interruption losses, and degrade collateral asset values. In regions exposed to severe weather, insurance premiums can spike dramatically and become a pain for inhabitants of these regions. As an example, this is illustrated by the projected surge in extreme rainfall premiums in Texas, climbing from 1% in 2002 to as high as 18% by 2080. This trend underscores the need to move beyond relying solely on past data, as climate change is transforming the risk landscape.

Beyond asset-level impacts, banks must also monitor liquidity strains in the wake of climate-related disasters. Households and firms affected by physical climate shocks may be in need to withdraw funds or may face difficulties repaying loans, which can severely amplify both credit and liquidity risks for banks, showing a real negative impact. Moreover, the ripple effects of physical climate risks can spread to broader macroeconomic vulnerabilities, as disasters disrupt regional or national economic activity. External studies, including those by the Intergovernmental Panel on Climate Change (IPCC, 2021), confirm that the frequency and intensity of such events are expected to increase in the upcoming years, making the integration of physical risk assessments into banking risk management frameworks an urgent and strategic priority for the financial sector.

1.4.2 Transitional Risks

Transitional climate risks, in contrast, arise from the economic, technological, and regulatory shifts required to transition toward a low-carbon economy. While physical risks relate to external, often sudden shocks, transitional risks are embedded in structural changes to how economies and industries operate. Banks face these risks through the potential for asset stranding, as investments in carbon-intensive sectors like oil and gas lose value rapidly when policies tighten and market preferences evolve.

The BNP Paribas presentation (Guay & Sposito, 2024) illustrates that these transitional risks can lead to lower collateral values and higher loss-given default (LGD) for vulnerable assets. Moreover, abrupt policy shifts, such as stricter emissions regulations or new carbon pricing mechanisms, can significantly alter market dynamics, directly affecting credit, market, and liquidity risks. Transition risks also extend beyond credit risk to broader strategic risks: banks heavily invested in high-carbon sectors may face reputational risks, regulatory penalties, and declining investor confidence if they fail to adapt swiftly.

The urgency of meeting net-zero targets has already prompted many financial institutions to recalibrate risk models and asset valuations. For example, tools like MSCI's Climate Value at Risk (Climate VaR) have emerged to quantify these transitional risks by combining scenario analysis with financial modeling to predict potential monetary losses through 2100. External sources, including the Network for Greening the Financial System (NGFS, 2020), emphasize that banks must incorporate these transition scenarios into stress testing exercises to fully capture the multifaceted impacts of the low-carbon transition.

Such comprehensive approaches are crucial for ensuring that banks maintain capital adequacy and operational resilience as the global economy shifts towards more sustainable practices. In addition, regulators increasingly expect banks to disclose how they integrate these transitional risks into their decision-making and to demonstrate how their lending practices and portfolios align with climate goals. This marks a fundamental shift in how climate-related transition risks are treated, not as peripheral concerns, but as core financial risks.

1.5 Empirical Literature

Over the past fifteen years ESG has moved from a niche concern to a mainstream research topic, drawing the attention of scholars, investors and regulators alike. This surge in interest has produced a diverse empirical record. Some studies concentrate on specific regions or individual E, S and G pillars, while others adopt global samples and composite scores. The resulting evidence is mixed. A sizable group of papers concludes that strong ESG performance reduces both idiosyncratic risk and a firm's contribution to system-wide distress, pointing to better governance, more patient capital and enhanced stakeholder trust as key channels. In contrast, another line of inquiry warns that the rapid, crowded shift into a narrow universe of highly rated ESG assets can increase portfolio overlap, amplify correlation and therefore raise systemic risk. Together these findings highlight that the risk implications of ESG integration depend heavily on context such as market structure, measurement quality and the degree to which investors pursue identical sustainability signals. To clarify these competing narratives the next sections review a series of representative studies that capture both sides of the debate and shape our hypothesis.

1.5.1 ESG as a stabilizing force in the financial industry?

A growing body of empirical research suggests that strong ESG performance can mitigate the contribution of financial institutions and corporations to systemic risk. Several recent studies document this stabilizing effect across various contexts, ranging from global banking systems to firms in emerging markets, as well as in the behavior of investment funds.

For instance, using a dynamic panel model, Aevoae et al. (2023) analyze a sample of 367 listed banks across 47 countries (developed and developing) between 2007 and 2020. They use DeltaCoVaR to measure each bank's systemic risk contribution. Their results show that a higher overall ESG score is associated with a statistically significant reduction in systemic risk contribution. Notably, it is the governance pillar that seems to drive this effect: the improvement of corporate governance practices appears to be a key lever for reducing interconnectedness between banks and maintaining financial stability. The authors observe that banks with higher governance scores show lower tail-risk interconnectedness, even after controlling for size, capital adequacy, and other bank specific characteristics.

These results are robust to alternative systemic risk measures (including contribution-based metrics such as Delta CoVaR and exposure-based metrics such as ES) and remain valid in static model specifications. The stabilizing effect of ESG also appears more pronounced in developed economies, which suggests that sustainable practices, particularly in governance, translate more easily into reduced systemic vulnerabilities when financial systems are more mature or subject to stricter regulatory oversight. Stronger governance likely improves the quality of risk management and transparency, thereby reducing the likelihood that an institution becomes a vector of contagion during periods of stress. More broadly, the results of Aevoae et al. support the view that integrating ESG criteria, especially governance standards, into banking management and supervision can complement traditional prudential tools in preserving financial stability.

Complementary results are observed in the non-financial corporate sector, including in emerging markets. Gidage et al. (2024) examine the relationship between ESG performance and risk sensitivity of Indian firms, using equity beta as a proxy for systemic risk (a CAPM-derived measure indicating the sensitivity of a stock to market fluctuations). The authors show that, on average, firms with better ESG scores exhibit lower sensitivity to systemic risk, with less volatile stock performance, indicating a more defensive risk profile. However, this effect is not uniform: it is more pronounced for younger, smaller firms or those with higher borrowing costs.

Conversely, larger, established firms or those benefiting from low-cost financing perceive smaller marginal gains in terms of risk reduction. ESG practices thus seem to play a particularly protective role for the most vulnerable actors. These findings align with stakeholder theory, which posits that responsible corporate behavior reduces firm-specific risks and acts as a form of insurance in uncertain environments.

Another channel of stabilization linked to ESG emerges in the behavior of investment funds. Capotă et al. (2022) show that ESG-oriented funds exhibit a weaker relationship between investor flows and financial performance. Investors are less likely to withdraw after poor performance and less inclined to reinvest aggressively after strong returns. This more patient behavior limits the risk of large-scale redemptions or fire sales during times of stress, thereby reducing contagion effects. In fact, ESG funds (equity and green bonds) experienced significantly smaller outflows during market shocks compared to traditional funds. This flow stability indicates that ESG investors may have longer investment horizons or stronger non-pecuniary commitments, which can reduce financing frictions for firms during crises.

Moreover, Cerqueti et al. (2021) explore, using a network-based approach, whether ESG strategies help limit contagion risk among investment funds. They simulate forced sale scenarios and find that funds with strong ESG compliance are less exposed to indirect losses resulting from the liquidation of neighboring portfolios. During periods of low volatility, the relative losses of ESG funds are significantly lower than those of low-rated ESG funds. This suggests that these funds either hold more diversified portfolios or assets less susceptible to fire-sale discounts. Even in times of heightened turbulence, the relative resilience of ESG funds persists.

These studies support the hypothesis that strong ESG performance enhances the resilience of the financial system. This relationship is verified at the level of banks and corporations (through the reduction of default risk and contribution to systemic risk) and at the level of investment funds and markets (through lower contagion and more stable investment behavior). In these researches, governance emerges as the key pillar: well-governed institutions better anticipate and manage risks, including ESG risks, thereby protecting both their own stability and that of the system as a whole. Even in emerging markets, where ESG frameworks are still developing, positive effects are observed, although modulated by institutional characteristics. Finally, the relative stability of ESG fund flows during recent crises, such as the COVID-19 episode, illustrates a “safe haven” effect attributable to the long-term engagement of sustainability-oriented investors. Overall, these findings fully justify our first two hypotheses: first, that banks with higher E scores present a lower contribution to systemic risk, as measured by our systemic risk measures, and second, that banks who adopted environmental practices post COP21 are associated with a general reduction in systemic risk.

1.5.2 ESG investments as a source of systemic risk

Counterbalancing the optimistic findings above, another strand of recent literature argues that the rapid integration of ESG into investing could itself become a source of systemic risk. These studies focus on the potentially destabilizing effects of widespread ESG-driven investment flows, such as asset crowding, valuation bubbles, and new forms of correlated exposures. The central concern is that if many investors and institutions pivot in unison toward a relatively small universe of “sustainable” assets, the financial system may become less diversified and more vulnerable to shocks in those assets. Empirical evidence is emerging to support this cautionary view. For example, Dong (2023, 2025) investigates U.S. financial institutions and finds that those chasing ESG investments more aggressively tend to contribute more to systemic risk, not less. Using panel data on institutional equity holdings from 2010 to 2020, Dong’s analysis reveals a positive association between the share of a financial institution’s portfolio allocated to high-ESG-rated firms and that institution’s systemic risk as measured by its distress contribution to the system. In practical terms, banks and asset managers that tilted heavily towards companies with top ESG credentials were also those with higher ΔCoVaR or similar systemic risk indicators. This somewhat counterintuitive result can be explained by asset crowding and correlation: when many institutions hold similar “good” assets, their fortunes become linked. A negative shock to one large ESG-friendly firm or sector (a clean-tech industry setback or a governance scandal at a prominent ESG leader) could simultaneously hit all the institutions holding those assets, leading to a correlated dip in capital across ostensibly diverse firms. Dong (2025) emphasizes that the very act of concentrating investments in firms with similar ESG profiles increases co-movement among financial institutions’ asset values. From a systemic perspective, this reduces the benefits of diversification and hedging that would normally protect the system if different banks held truly uncorrelated portfolios. The finding does not imply that

ESG causes poor management or direct risk-taking but that herding into ESG assets creates a new form of interconnectedness. This insight adds nuance to the ESG discourse: even as individual institutions might appear safer by holding “high-ESG” assets, collectively, they may undermine system-wide stability by becoming too alike in their exposures. In short, portfolio alignment on ESG criteria can lead to systemic concentration risk, echoing classical lessons that any monoculture in investing (even well intentioned) can foster fragility.

Reflecting this concern, global regulators have warned of the possibility of “green bubbles” forming in financial markets. The Bank for International Settlements (BIS 2021) noted unusual valuations in segments of the market heavily influenced by ESG mandates. In a 2021 quarterly review, the BIS highlighted a paradox: When aligning finance with sustainability is crucial, the rapid inflow of money into ESG assets was outpacing the development of robust market assessments, leading to signs of stretched valuations for ESG assets. According to the BIS, certain ESG-labeled stocks and bonds were trading at elevated price multiples that fundamentals could not justify, mainly because investors were piling in for their perceived future potential or to meet portfolio ESG targets. Such overvaluation is a classic precursor to asset price bubbles. If a “green bubble” were to inflate and then burst, it could pose systemic risks analogous to other financial bubbles, albeit with an ESG tint. Financial institutions heavily exposed to overpriced green assets would suffer sudden losses in a sharp correction, and if those assets had been widely used as collateral or considered low-risk, the re-pricing could induce liquidity strains. The BIS warning that “ESG assets’ valuations may be stretched” is a call to recognize that ESG investing does not eliminate market cycles, exuberance can also build in this domain. Moreover, the BIS pointed out that the investment management industry was still adapting its risk models to ESG; with many asset managers using traditional tools, there was a risk that new correlations introduced by ESG preferences were not fully accounted for. This underscores that in the transition phase, mispricing and model uncertainty around ESG assets can themselves become sources of instability.

Academic studies further show how incomplete or inconsistent integration of ESG factors might increase risk in the short run. Pistolesi and Teti (2024) provide a nuanced perspective by examining U.S. equity markets from 2005 to 2021 and discovering an inverted U-shaped relationship between firms' ESG ratings and their systematic risk (as measured by market beta). Their findings indicate that companies with either very low or very high ESG scores tend to have lower risk than those in an intermediate range of ESG performance. At first glance, this suggests a counterintuitive outcome: firms with minimal ESG engagement can appear less risky than those doing a moderate amount. The authors interpret the left side of the "inverted U" as possibly reflecting that firms investing very little in ESG may have more free cash flow and strategic flexibility in the short term, or they may be in stable traditional industries.

In contrast, firms with mid-level ESG engagement incur some costs of transitioning without reaping the full benefits, resulting in slightly elevated risk profiles. Crucially, beyond a certain threshold of ESG investment, the relationship turns – high ESG performers exhibit significantly lower risk, confirming the long-run risk mitigation thesis. This pattern implies a transitional vulnerability: as firms ramp up ESG efforts, there may be a period during which their risk does not monotonically decrease and might even uptick, perhaps due to restructuring, learning curves, or investor uncertainty about their new strategy. At the macro level, if the whole economy is in the midst of such a transition (with many firms moving from low to higher ESG practices), the system could experience heightened fragility until the transition is complete. For instance, if policymakers and investors pressure companies to improve ESG standards rapidly, many firms may find themselves in that transition phase simultaneously which could temporarily raise credit or market risk. Systemic risk subsides only after they attain robust ESG integration (the right side of the U-curve). Pistolesi and Teti's evidence thus serves as a caution: the path to sustainability is not risk-free, and the timing and extent of ESG adoption matter. Policymakers might need to guard against systemic tremors during the ESG uptake process, ensuring that support (or at least careful monitoring) is in place as firms make this pivot.

Another facet of ESG-related systemic risk arises from the shifting risk exposures of banks and insurers as they incorporate ESG considerations. Curcio et al. (2024) examine how the financial system's vulnerability changes when banks and insurance companies increase their exposure to high-ESG firms. Their study focuses on European markets (2016–2022). It uses a Delta CoVaR approach at the sector level, linking the systemic risk of the banking and insurance sectors to the market performance of various stock indices representing ESG-compliant (“green”) companies vs. traditional “brown” companies. In line with the stabilizing strand of literature, Curcio et al. find that when green firms perform well (rising ESG equity index values), the systemic risk of banks and insurers falls relative to when conventional market indices rise. This suggests that a general shift towards sustainability in the economy benefits financial stability, likely because greener companies may be more resilient or because banks’ loans to and investments in such companies become safer. However, the study also uncovers a critical risk: banks are more sensitive to the downside of green assets. Suppose the ESG equity index experiences a drop or green firms become more volatile (for example, due to an adverse event or earnings shock). In that case, banks’ systemic risk increases more than it would for an equivalent deterioration in a broad or brown index.

In contrast, insurers in their sample remained more exposed to the riskiness of the oil and gas sector firms (perhaps reflecting insurance companies’ legacy underwriting in those industries). The takeaway is that banks rebalance their portfolios toward sustainable sectors and concentrate their exposure to those sectors’ fortunes. While this yields systemic risk benefits in good times (when the sector prospers), it also creates a potential single point of failure that could harm the banking system if the green sector faces headwinds. This finding dovetails with Dong (2025) and others, it is essentially the flip side of the ESG coin, where correlation risk creeps in through well-meaning reallocation of capital. It underscores that the ESG diversification benefit might be limited if everyone diversifies in the same way. If all banks diversify away from brown industries and towards the same green industries, the banking system may end up less diversified than before. Curcio et al. conclude that risk transmission channels are changing, the sources of systemic shocks in a greener financial system might differ from those in the past, and they call for further research and vigilant risk management to understand these evolving channels.

Finally, broad oversight institutions like the IMF have stressed that systemic vulnerability could spike during the ESG transition phase if information and incentive problems are not addressed. The IMF’s Global Financial Stability Report (2019) noted that markets could face price discovery challenges and abrupt shifts as sustainable finance grows. A key point raised is that large-scale reallocation of portfolios based on new ESG guidelines or climate policies could lead to disorderly price corrections if investors suddenly reassess the risks. For instance, if new environmental regulations or a rise in carbon prices render certain brown assets unviable, there could be a rush to divest that causes a rapid price collapse, a dynamic similar to previous episodes of benchmark-driven investing where many institutions simultaneously trade in the same direction. Moreover, the IMF highlighted issues of data and disclosure: inconsistent ESG definitions and lack of transparent risk disclosure can lull investors into misjudging their accurate exposures. If firms are not fully reporting climate-related risks, or if ESG ratings are not comparable across providers, banks and investors may think they are diversified or low-risk when, in fact, they are exposed to hidden correlations. Such information gaps and "greenwashing" could mean that the buildup of ESG-related risks goes unnoticed until a trigger event causes a sudden revelation. At that point, the adjustment would be sharp and potentially systemic. The IMF report explicitly warns that significant, sudden shifts in investor sentiment or norms (e.g. a wave of institutional investors pledging to divest from fossil fuels or, conversely, a loss of confidence in an overhyped clean-tech theme) could amplify volatility and stress in the financial system. This perspective aligns with academic concerns that markets might underestimate certain ESG-tail risks. For example, investors could underestimate the likelihood of extreme climate events or policy shifts and thus not price them in, leading to a scenario where corrective re-pricing happens simultaneously. One analysis analogizes this to the pre-2008 underestimation of correlated mortgage defaults – here, the underappreciated factor could be an ESG-driven correlation. Indeed, analogous to how banks' simultaneous foray into complex securities before 2008 increased systemic fragility, the simultaneous pursuit of ESG assets now may increase co-movements among portfolios. Researchers note that this could lead to asset price bubbles in ESG areas and a higher risk of contagion when those bubbles deflate. Additionally, the physical risks of ESG (such as climate-related natural disasters) are hard to predict and could directly shock the financial system in unexpected ways. The combination of opacity (limited data on who holds what ESG exposures or how a supply-chain climate risk might transmit) and correlation (many holding similar assets) means the system could be caught off guard by an ESG-related shock, paving the way to instability.

The literature on ESG and systemic risk paints a two-sided picture. On one side, strong ESG performance, especially robust governance, has been empirically linked to lower risk at the firm and bank-level and to more stable financial flows, indicating a potential stabilizing influence on the financial system. This supports the optimistic view that aligning finance with sustainability can yield a safer, more prudent financial sector as better-managed companies and patient capital become the norm. Conversely, research highlights that poorly coordinated or overly enthusiastic ESG integration can introduce new systemic risks. During the transition to a sustainable economy, markets may experience pockets of overvaluation, herd behavior into similar assets, and novel interconnections that regulators and investors are not yet fully equipped to monitor. In the extreme, a rush into ESG assets could create the same conditions that foster systemic crises: concentrated exposures, unknown correlations, and the potential for abrupt collective reversals. Therefore, the current empirical evidence suggests a nuanced synthesis: ESG investing per se is neither an unalloyed safeguard nor a unilateral threat to financial stability; it can be both, depending on how it is implemented and in what market context. The challenge for policymakers and financial institutions is to maximize the stabilizing benefits of ESG while minimizing the destabilizing byproducts (through careful monitoring of valuation and concentration risks).

As the thesis progresses, our analysis will test the hypothesis that E score has a positive impact on banks and their systemic risk, the objective is to understand if the E score is a good investment that must be continued for banks, in order to reduce their systemic risk and contribute to the global industry's systemic risk.

1.5.3 Environmental Engagement as a Threshold Effect

The second hypothesis of this thesis posits that the introduction of environmental practices and of the environmental score alone following the 2015 COP21 agreement contributes to a statistically significant reduction in systemic risk, regardless of the absolute value of the E score. This hypothesis is based on the idea that environmental engagement acts as a threshold effect: the adoption or integration, even minimal, of environmental considerations already produces stabilizing effects within financial institutions. This approach departs from a linear logic in which risk reduction would be proportional to the intensity of ESG performance. It rather emphasizes the very presence of environmental engagement as a qualitative shift in institutional strategic behavior and in the signals sent to the market.

A body of existing empirical research supports this interpretation, such as Ilhan et al. (2021), who analyze the implications of disclosures related to climate risk and show that firms that communicate on their environmental commitments exhibit lower stock market volatility. Their study, based on a large sample of firms and institutional investors, reveals that the simple existence of environmental disclosures modifies investor risk perception. In this sense, financial markets interpret these disclosures as a signal of forward-looking governance and strategic alignment with climate resilience requirements, which reduces exposure to shocks and the required risk premium. In other words, it is the act of disclosing, and not necessarily its content, that generates a signal effect sufficient to reduce volatility.

In their research, Krueger et al. (2020) reinforce this reading by showing how institutional investors internalize climate risk by demonstrating that the simple recognition of environmental risks, even without a sophisticated mitigation plan, influences market expectations and risk evaluation mechanisms. This dynamic echoes stakeholder theory, according to which the recognition of non-financial issues strengthens the credibility and legitimacy of institutions. Thus, in the wake of COP21, the initial implementation of environmental strategies (even limited ones) can be perceived as a turning point for market-based risk perception.

In the banking sector, the signal effects linked to the consideration of environmental risks appear even more directly connected to the transmission of systemic risk. Monasterolo and Raberto (2019) use a macro-financial agent-based model (the EIRIN model) to simulate how the integration of climate risks influences bank behavior. Their results show that financial institutions that begin to integrate environmental risks are less vulnerable to transition shocks and less likely to amplify financial instability through their loan portfolios. The determining factor does not lie in the absolute level of ESG performance, but in the timing and nature of the integration. Early and even basic incorporation of environmental risks already leads to a substantial reduction in contagion channels. This supports the idea that the very existence of environmental engagement is more decisive than its intensity in mitigating systemic fragility. This leads to a clear conclusion, that environmental performance, especially in its early stages, constitutes a qualitative signal that modifies the behavior of financial stakeholders. In the bank sample analyzed in this thesis, the post-COP21 introduction of E score reflects the same "threshold effect". Banks that begin to publish an E score send a credible signal to regulators, investors and counterparties, showing their alignment with the structural transition toward sustainable finance. This change is not purely symbolic, it influences how banks are treated in capital markets, impacts investment flows, and modifies how interconnected institutions assess their counterparties.

As a result, this hypothesis is based both on empirical evidence and on systemic logic: environmental practices modify the strategic behavior of banks as well as their position in financial markets, even before their E score reaches a high level. This implies that the introduction of E scores after COP21 constitutes a structural shift in institutional behavior, initiating a process of systemic risk reduction without requiring a high level of ESG performance.

In sum, the literature suggests that environmental engagement functions as a threshold effect in the reduction of systemic risk. Ilhan et al. (2021) and Krueger et al. (2020) show that early-stage environmental policies, even if limited, influence risk perception and market volatility. Monasterolo and Raberto (2019) extend these conclusions to the macro-financial scale, demonstrating that even moderate integration of climate risks into banking behavior produces tangible stabilizing effects. These results provide strong theoretical and empirical support for the hypothesis that the introduction of E scores in banks post-COP21 is associated with a measurable reduction in systemic risk, regardless of the absolute level of the E score.

This thesis builds directly on this insight by testing whether the initial appearance of an eE score as a post-COP21 institutional shift acts as a structural break in risk perception and behavior. While previous studies have highlighted the effects of environmental disclosures and early engagement, this research leverages the COP21 timeline to isolate the effect of environmental score introduction itself. In doing so, it contributes new empirical evidence on whether the mere presence of E score reporting, regardless of its level, corresponds with lower systemic risk contributions. This "threshold" hypothesis offers a novel angle in the literature by anchoring environmental engagement within a defined global policy event, allowing for a clearer identification strategy grounded in regulatory transition.

The empirical literature reviewed in this section offers a nuanced and multi-layered understanding of the relationship between ESG performance and systemic risk. This body of research is essential in grounding the two hypotheses formulated in this thesis, both from a theoretical and an empirical standpoint.

First, the studies supporting a stabilizing role for ESG performance (Aevoae et al., 2023; Gidage et al., 2024; Capotă et al., 2022; Cerqueti et al., 2021) inform the first hypothesis of this work, which suggests that banks with higher E Scores contribute less to systemic risk, as measured by risk metrics such as MES, CoVaR, and Delta CoVaR. These studies provide strong empirical evidence across institutional types and geographic regions, suggesting that high ESG performance enhances risk management practices, reduces volatility, and contributes to more stable capital flows, all affecting systemic risk. The mechanisms highlighted directly support the notion that better ESG performance is correlated with reduced downside risk and contagion potential. By applying these insights specifically to E scores in the banking sector post-COP21, this thesis contributes to an ongoing line of inquiry while refining it toward a more targeted empirical context, focusing solely on the Environmental practices.

Second, the literature that highlights the limits of ESG integration (Dong, 2023, 2025; BIS, 2021; Pistolesi & Teti, 2024; Curcio et al., 2024; IMF, 2019) grounds the relevance of the second hypothesis: that the simple adoption of environmental practices or the publication of E scores post-COP21 leads to a structural reduction in systemic risk, even if the E score itself is not high. In particular, the work of Ilhan et al. (2021), Krueger et al. (2020), and Monasterolo and Raberto (2019) underscores the notion that environmental engagement functions as a threshold effect, that is, a qualitative shift rather than a linear improvement. These findings imply that the act of integrating climate considerations into decision-making already changes institutional behavior and investor perception, reducing systemic fragility regardless of intensity. The empirical frameworks adopted in these studies support the idea that perception, timing, and engagement strategy matter as much as (if not more than) absolute ESG scores in shaping systemic outcomes.

Taken together, this literature review thus plays a double role. It validates the empirical relevance of this thesis by demonstrating that environmental performance is not a marginal consideration in risk analysis, but a determinant factor with proven links to financial stability. And it also provides the conceptual architecture for hypothesis formation. Rather than testing generic claims about ESG and risk, the thesis builds on these studies to assess two concrete and complementary questions: **(1)** whether better E scores correspond to lower systemic risk in banking, and **(2)** whether the post-COP21 adoption of E scores marks a significant inflection point in systemic risk dynamics. These hypotheses are not formulated in isolation but emerge directly from the tensions and patterns identified in the existing literature. In this sense, the empirical review is not merely descriptive, but formative in guiding the structure and objectives of this research.

Chapter 2 : Data Collection & Preliminary Analysis

2.1 Data Description

In this chapter, we present the data collected to support our analysis, providing a detailed overview of the variables and datasets that will inform our subsequent models. We also outline, step by step, how we processed and prepared this data, setting the foundation for the analysis in the following chapter. As part of this preliminary exploration, we grouped the banks by region to conduct an initial analysis of mean returns across regions and to examine how these returns relate to their Environmental scores.

2.1.1 Banks' Stock Returns

This log return analysis is based on a cleaned and filtered panel of 113 publicly listed banks selected from an original, much larger cross-border dataset. To get to this dataset, we first filtered banks based on returns and E scores, so the banks for which we did not have returns from 2012 to 2023, which were important for the analysis, were excluded. This data exclusion included high-profile banks that either collapsed or merged during the sample period such as including Credit Suisse, which was forcibly merged with UBS in 2023 after a long deterioration of investor trust, and Silicon Valley Bank (SVB), which failed abruptly in March 2023 due to interest rate risk mismanagement and deposit flight. Banks without E scores (from Bloomberg) were excluded from the original sample. We observed that many "smaller" banks did not show this score, so we excluded them from the panel. We then looked at the control variables we will use in our models, and we removed banks for which these variables were unavailable. The panel includes institutions across seven global regions: North America, Europe, Asia, the Middle East, Oceania, South America, and Africa. Each bank's daily stock returns were converted into logarithmic returns. The final sample period runs from January 2012 to December 2023, spanning major global and regional crises, including the post-GFC regulatory era, COP21 and the climate finance mobilization effort, the COVID-19 shock, inflationary tightening cycles, and the banking stress episodes of 2023.

Several notable patterns emerge when examining the distribution of daily log returns by region. Across all regions, returns cluster heavily around zero, with a strong central peak consistent with high-frequency market efficiency. However, the tails of the distributions, which represent extreme daily moves, vary significantly across geographies. North America, home to the sample's most liquid and deep financial markets, shows a relatively tight distribution with moderate negative skewness. However, its tail risk is non-negligible: In March 2020, the region's worst-performing banks observed daily returns as low as -16.3%, representing the fear when the Covid-19 crisis was announced. This was followed by a strong rebound, with some institutions posting daily log returns of over +12% within the same month, a pattern repeated during the U.S. banking panic of 2023 (See Annex 1 for Banks Log Returns).

Europe exhibits a similarly sharp central peak but with a slightly wider spread, and in certain years (notably 2016 and 2020), return extremes surpassed even those in North America. These reflect vulnerabilities tied to systemic uncertainty (e.g., Brexit, sovereign debt, the Greek crisis, and energy supply shocks) and stock-specific exposures. Notably, the average daily return in Europe across the sample was close to zero, suggesting stagnation in long-term capital performance. However, this conceals high dispersion across individual banks and years. For example, 2020 saw both the worst and some of the best daily returns in the region depending on the bank and timing of market interventions.

Asian banks presented more symmetrical return distributions, albeit with lighter tails than Europe and North America. However, this does not mean less risk. The return histogram for Asia reveals frequent low-amplitude fluctuations and occasional sharp corrections, often linked to Chinese regulatory actions, currency interventions, or regional geopolitical stressors. The average daily return in Asia was positive but minimal over the full sample, suggesting a modest but persistent upward trend, though less dramatic than in North America.

South America, by contrast, shows the widest and most leptokurtic return distribution. Several banks in this region experienced daily log returns as low as –14.5% and as high as +9.3%, reflecting the extreme sensitivity of these markets to macroeconomic instability, capital controls, and inflation shocks. Moreover, volatility in South America was both high in magnitude and persistent over time, making it one of the most volatile regions in the panel, which is consistent with weak monetary anchors and low investor confidence.

The Middle East and Africa offer a contrasting picture: both regions feature highly peaked return distributions with relatively thin tails. This suggests fewer trading days with significant returns and reflects lower market liquidity, fewer listed banks per country, or dominant state ownership. However, even in these more insulated markets, systemic events like COVID-19 still triggered return crashes, particularly in March 2020, where several African and Middle Eastern banks experienced daily drawdowns exceeding –11%.

Although smaller banks in that region are mapping, Oceania behaves more like North America: stable, centralized return distributions punctuated by crisis-related shocks. Its average daily log return hovered around +0.009%, suggesting healthy but modest growth. Still, Oceania was not immune to market-wide drawdowns, as shown by sharp spikes in volatility around the COVID-19 crash and global tightening cycles post-2021.

When we zoom out to consider the aggregate monthly and yearly log return patterns, it becomes clear that while average returns remained close to zero for most regions, the variability around those means was anything but trivial. Monthly averages oscillated sharply in reaction to global crises and localized financial stress, with dramatic collapses observed in March 2020 across all regions, followed by swift rebounds, particularly in countries where monetary and fiscal policy were strongly coordinated. North America again led the post-crisis rally regarding return amplitude and speed, whereas Europe and Asia recovered more slowly. Africa and the Middle East exhibited the most minor return swings, likely due to lower daily volatility and the presence of long-term investors who are less prone to panic trading.

This log return analysis provides strong evidence of non-Gaussian behavior, asymmetric downside risk, and region-specific volatility regimes. These findings matter deeply for any systemic risk assessment, highlighting the differential fragility of banking systems under stress. They also set the stage for deeper investigation into how E scores may interact with these risk structures. E score improvements might offer stronger downside protection or resilience signaling in markets with high volatility, such as South America. In contrast, for mature markets like Europe or North America, E scores may be more tightly priced, making them meaningful primarily in tail event mitigation rather than return enhancement. Either way, the regional distribution of returns and their extreme behavior must be acknowledged and integrated into any ESG–systemic risk framework. The patterns observed here are not mere noise, they are structural, persistent, and indicative of how finance, geography, and sustainability intertwine.

2.1.2 Log Return Transformation and Stationarity Testing

Before conducting any risk modeling or econometric analysis, a crucial preprocessing step involved converting all raw price series into logarithmic returns. Indeed, the raw prices are not consistent, some banks have prices in different currencies, but also some stock prices are higher than others. Therefore, a $1 increase in a stock could be smaller and less significant than a $1 increase for another bank. This transformation is a common practice in financial econometrics and in similar existing research. From a risk management perspective, the use of log returns allows for a more realistic and stable modeling of volatility clustering and tail behavior, which will be necessary for our calculations and analysis. It also simplifies the calculation of continuous compounding, which is widely used in theoretical finance and stress testing frameworks. We assume that log series are more stationary and that they are better normally distributed, characteristics that will be required for our analysis later on.

To test the stationarity of each bank's log return series, we apply the Augmented Dickey-Fuller (ADF) test. The test is based on a regression that includes a drift term (constant) but here we exclude any deterministic time trend, as log returns do not exhibit a systematic directional trend over time. Lag length is selected automatically using the Akaike Information Criterion (AIC), allowing the model to correct for autocorrelation without overfitting. The null hypothesis H0 assumes the presence of a unit root (i.e., the series is non-stationary), while the alternative H1 implies stationarity and therefore the absence of a unit root. A series is deemed stationary if the ADF test statistic is lower than the critical value at the threshold, here we use the 1%. This setup is consistent with standard practice in financial econometrics and ensures a robust assessment of the statistical properties of return data.

The results showed that the null hypothesis of a unit root was firmly rejected at the 1% significance level for all banks in the sample, with ADF Stats ranging from -60 to -9 and p-value nearly 0. This implies that the return series are stationary and thus suitable for econometric analysis.

From an intuitive standpoint, this outcome aligns with established financial theory and empirical evidence: asset prices are typically non-stationary due to their trending nature, but returns are widely accepted to be stationary, especially at higher frequencies like daily or weekly. By ensuring stationarity in returns, we avoid the risk of overestimating persistence or volatility in our risk models, which would bias systemic risk estimates. Transforming raw prices into log returns and validating their stationarity through ADF tests was essential in preparing the dataset for the econometric analyses that follow. This step adheres to best practices in financial econometrics and provides strong statistical foundations for modeling systemic risk across global banks using high-frequency market data.

2.1.3 Macroeconomic and Bank-related Control Variables

For our model, we choose some macroeconomic variables that affect all banks and some banks specific variables with economic data specific to the very same banks. The objective of having control variables is to make the model and regressions less biased and increase the variance explained in the model. We will therefore try to capture some movements that can be explained by these controls, in order to isolate the effect of the E Score on the systemic risk measures.

Therefore, we choose the following macroeconomic variables: The inflation with a t+1 lag, that is the inflation in 2021 captures the effect in 2020 for example, which is common as usually the inflation appears after a crisis or an effect, even more because we are using yearly data, so the inflation in 2021 could explain events in 2020, that's why we use it for 2020 then. Then, we use the TBill rate, as we think end of year returns could explain some unexpected events during a year, as this rate is known as the "risk free" and supposed to be stable, so an event could increase this variable returns. Also, we use the Vix Index, as this index measures volatility and therefore could explain large movements during a year due to some events, crisis, to capture this effect.

Then, we have the following bank specific variables: leverage, market capitalization, price-earnings ratio, return on assets (ROA), trading volume, debt, debt-to-equity ratio, and annual revenue. All these are annual values with values for each bank in the panel data. They are essential in the thesis to account for bank-specific financial characteristics that could also influence systemic risk. Leverage and debt ratios may capture the financial fragility of banks, while market capitalization and volume provide insights into market liquidity and investor perceptions, which can be different in crisis time for example. The P/E ratio and ROA indicate profitability and efficiency, and dividends reflect payout policies. Revenue acts as a proxy for overall size and business scope. Together, these controls, once again, help isolate the unique impact of the E score on systemic risk, ensuring that our analysis accounts for fundamental financial drivers beyond environmental performance.

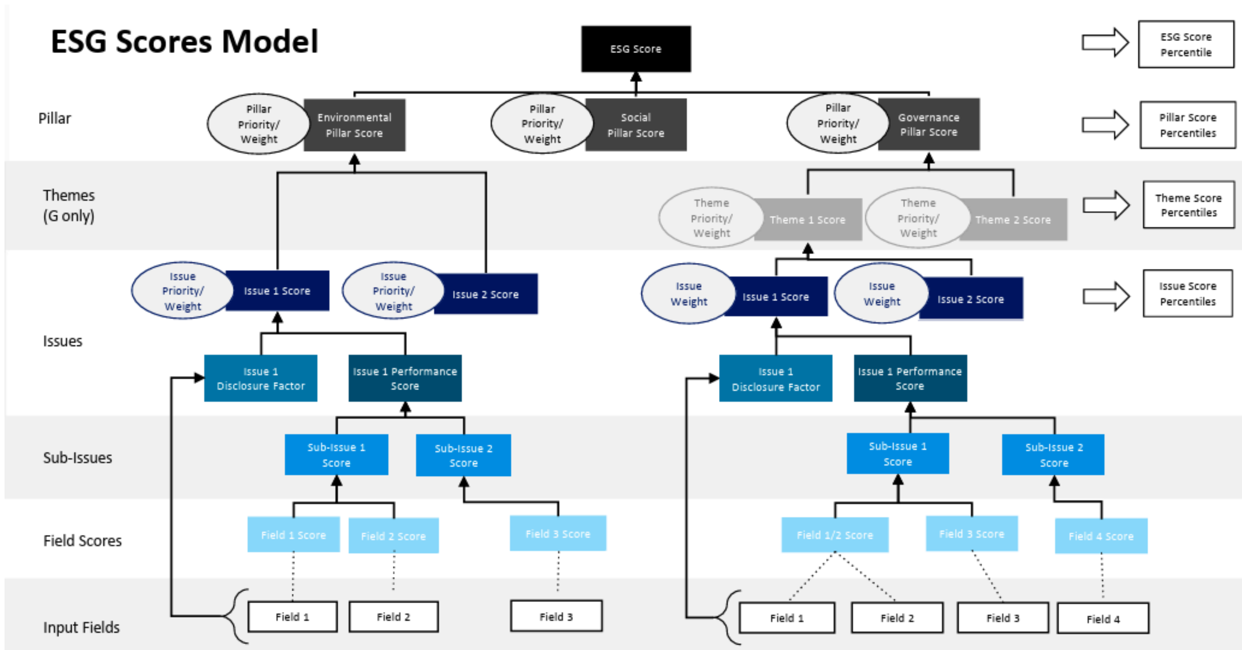
In order to make the control variables in our regression models more comparable and interpretable, we calculate the log returns of all of them before integrating them in the estimation process. The objective once again is to avoid all varying scales, and make the results more robust with less bias, because some could have higher impact on the regression (e.g. market capitalization values are much bigger than the ones for PE ratio and therefore could dominate the estimation simply due to scale, even if its economic impact on systemic risk is not necessarily larger.)

2.2 Bloomberg E Scores: Methodology and Relevance

Bloomberg's methodology for calculating E scores provides a standardized, quantitative, and transparent framework for assessing a company's management of financially material environmental issues. Introduced in the aftermath of the 2015 Paris Agreement, this component scoring approach is aligned with the post-COP21 regulatory and market environment, which places growing emphasis on sustainability, climate-related risk, and responsible corporate behavior. The E score is part of a broader ESG scoring system, which evaluates how effectively a company identifies, discloses, and manages environmental risks and opportunities compared to its peers.

Bloomberg's E scores are published annually at the end of the year, and range from 0 to 10, 10 reflecting best-in-class performance in managing environmental risks and aligning with sustainable practices. The highest score in our panel is 7. These scores measure how well a company manages environmental issues compared to firms with similar business models and exposures. This makes them particularly useful in financial research, where sector-adjusted performance metrics are often required to ensure meaningful cross-sectional comparisons.

The scoring process consists of five distinct phases. First, Bloomberg conducts a research phase by collecting publicly disclosed data from corporate sustainability reports, integrated annual reports, regulatory filings (e.g., 10-Ks), and other non-financial documents. Unlike some ESG providers, Bloomberg does not incorporate subjective analyst assessments or forward-looking opinions, ensuring a purely data-driven approach. Second, a quality assurance process ensures the accuracy and consistency of all collected data through multiple validation layers. Third, each field and environmental sub-indicator is scored using a quantitative model based on the quality of disclosure, the financial materiality of the issue, and its relevance to the company's operations. Fourth, the scoring outputs undergo a statistical and heuristic validation before being published to ensure reliability. Finally, the results are available on Bloomberg terminals under the section ESG as part of the Financial Analysis section of a company.



Source*: ESG Scores Model, Bloomberg (2023, p. 12), Bloomberg ESG Scores: Overview & FAQ.*

At the core of Bloomberg's E scoring framework are environmental key performance indicators (KPIs) such as carbon emissions intensity, energy efficiency, water usage, waste management, and compliance with environmental standards. These are aggregated through a structured hierarchy: individual fields are grouped into sub-issues, which are then combined into broader issues, and subsequently into themes and the final E pillar score. Each step of aggregation uses weighted averages and includes a Disclosure Factor, a metric from 0 to 1 that reflects the comprehensiveness and transparency of company reporting. This layered system ensures that the final E score is not only a measure of environmental performance, but also of the quality of a company's disclosures and its engagement with sustainability issues.

Significantly, pillar weights within each peer group are determined using Bloomberg Intelligence's fundamental sector research. Each pillar (Environmental, Social, Governance) is assigned a relevance ranking from 1 to 5, which is then translated into a percentage weight. This process guarantees that the E score is sensitive to the specific material risks and opportunities faced by institutions in different industries, including the banking sector.

Bloomberg's annual updates to E scores allow for longitudinal performance tracking, enabling analysts to observe how institutions respond to increasing regulatory pressure, investor expectations, and climate-related risks year after year, with the detail of water usage, E investments and gas emissions for largest companies. For this study, the use of Bloomberg E scores offers a consistent, peer-adjusted, and regulatorily aligned environmental performance measure that can be directly linked to risk dynamics in the banking sector. By converting a complex set of environmental data into a single interpretable metric, Bloomberg empowers investors, regulators, and researchers to assess environmental risk exposures in a standardized and financially relevant way.

2.2.1 E Score Geographical overview

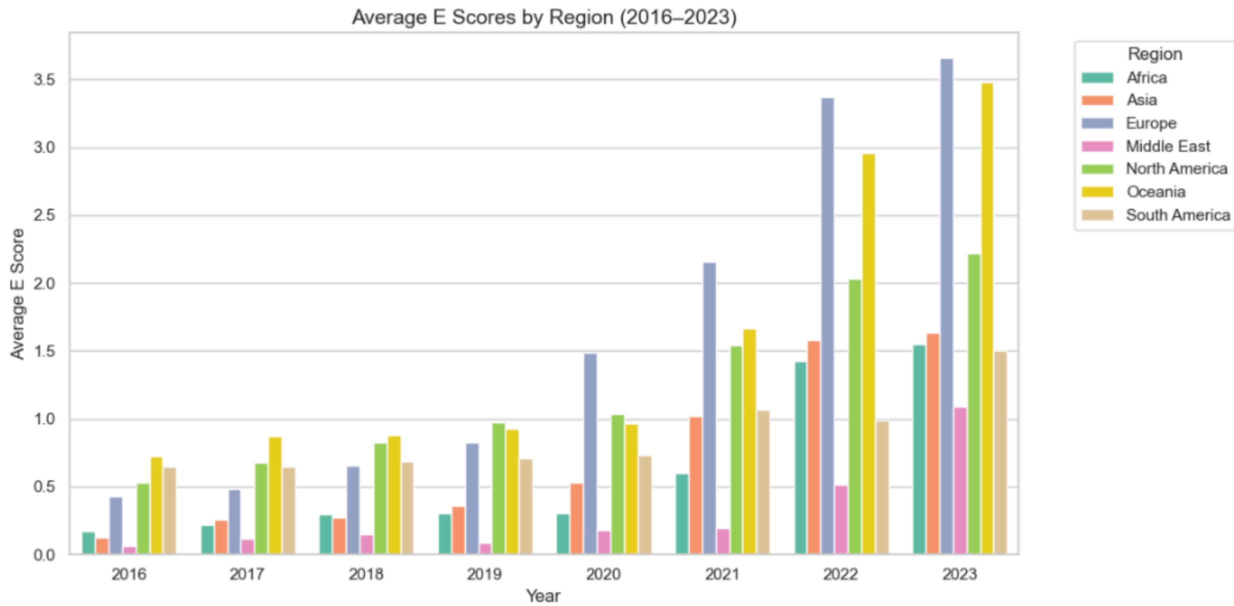
When looking at the E scores, the distribution provides critical insights into the uneven global integration of environmental considerations in banking, both in measurement and strategic prioritization. This dataset is supposed to be from 31/12/2015 to 31/12/2022, as the E score is given at the end of the calendar year, but our dataset shows dates from 2016 to 2023 as we use a one-year lag for adaptability and market response to a change in the E score. Indeed, it would not be pertinent to use the end-of-year E score from 2015 in our risk measurements 2015. Therefore, the end-of-year score in 2015 is used for 2016 risk calculations. This dataset is built on the environmental pillar of Bloomberg's ESG framework and comprises annual E score observations for the panel of 113 banks retained in the final sample. Notably, this dataset is highly asymmetrical in coverage, skewed toward developed financial systems, primarily Europe, North America, and, to a lesser extent, Oceania. This reflects a structural reality of market infrastructure and disclosure maturity, throughout the globe.

Developed regions have more listed banks due to their deeper and broader capital markets, however, more importantly, banks in these regions are significantly more likely to be covered by ESG rating agencies and to disclose standardized environmental data. Conversely, banks in Africa, the Middle East, and parts of Asia and South America are underrepresented, not because they are less relevant economically, but because ESG integration is either in its infancy or treated as secondary to other policy priorities, such as financial stability, sovereign exposure, or basic access to capital.

This divergence is evident in the data volume and the distribution of raw E scores. The overall histogram of raw scores reveals a long right tail and a dense mass near zero, reflecting many banks with scores close to zero, particularly in emerging and frontier regions, where investments are usually towards less environmental activities, but mostly fossil fuel resources, therefore exposed to climate, physical and transitional risks. These zero values do not only mean poor environmental performance but often denote non-disclosure, limited coverage, or a lack of alignment with international ESG taxonomies. The regional breakdown confirms this structural gap. Africa, for instance, has an average E score of just 0.61, with many banks holding a score of zero well into 2020, only beginning to exhibit upward trends in the last two years of the sample. Similar patterns are observed in the Middle East, where the mean E score across the period is just 0.30, the lowest of any region. In contrast, Europe leads the global landscape by a significant margin. The continent has the highest average E score (1.63) and demonstrates the steepest and most consistent year-over-year growth. From an average of only 0.43 in 2016, European banks accelerated their ESG adoption markedly post-2020, reaching a regional average of 3.66 by 2023, with institutions like HSBC, BNP Paribas, ING, and UBS emerging as frontrunners. This trend reflects regulatory pressure (e.g., EU Taxonomy, SFDR) and market expectations, where ESG performance increasingly influences investor behavior and risk modeling. North America also demonstrates strong ESG growth, albeit with greater heterogeneity. While some central U.S. banks like JPMorgan Chase, Bank of America, and Wells Fargo have substantially improved their scores since 2020, others remain stagnant or inconsistent, which helps explain the relatively high standard deviation (1.54) of North American scores across the sample. The regional average nonetheless rose from 0.53 in 2016 to 2.22 in 2023, indicating a broad but uneven ESG transformation.

Asia and Oceania occupy a middle ground. The E score uptake in Asia was sluggish before 2020 but accelerated, mainly driven by Chinese and Japanese megabanks. Despite beginning with a mean of only 0.13 in 2016, Asia reached 1.64 by 2023, with significant variability across banks and countries. Some institutions remain near zero, while others have moved closer to developed-market standards, particularly in Hong Kong and South Korea. Oceania, led by the Australian banking sector, displays one of the steepest post-2020 upward trajectories, with scores climbing from 0.73 in 2016 to 3.48 in 2023, driven by mounting investor scrutiny and the region's exposure to physical climate risks like wildfires and droughts. In South

America, the progression is more modest but still visible. The average E score climbed from 0.65 to 1.50, although high volatility and data gaps persist, especially among mid-sized institutions.



Source *: Self Plotted on Python, "Average E Scores per Region" (2016-2023, yearly E Scores from Bloomberg)*

The descriptive statistics and yearly trend plots confirm these narratives. Europe and Oceania now exhibit the highest average E scores and the highest variance, reflecting both progress and dispersion among institutions at different stages of ESG integration. In contrast, low-score regions like Africa and the Middle East show modest gains but remain relatively homogeneous, with many banks still clustering at or near zero. This heterogeneity directly affects systemic risk analysis: ESG scoring cannot be treated as a monolithic construct across global banking. The degree of environmental disclosure, regulatory pressure, and investor demand varies sharply by region, and so does an E score's interpretive weight. In Europe or Oceania, a high E score likely signals proactive environmental strategy and stakeholder alignment. A low score in Africa or the Middle East may reflect information absence, not necessarily poor environmental practice. Recognizing this asymmetry is critical to any empirical investigation into how E scores relate to financial stability and systemic risk in the banking industry.

Chapter 3 : Methodology & Risk Measures

3.1 Systemic risk Calculation Methodology

Systemic risk refers to the risk that the failure or distress of a single financial institution, or a group of interconnected institutions, could trigger instability across the entire financial system. Quantifying systemic risk is particularly challenging because it involves measuring individual exposures and the interdependencies between financial actors and their joint behavior under stress. For this reason, systemic risk requires more than traditional risk assessment tools; it calls for dynamic, market-based, and network-sensitive approaches that can capture tail events, contagion channels, and co-movements during crisis periods.

3.1.1 Traditional Risk Measures

A risk measure is a mathematical method for capturing risk. Traditional risk measures aim to quantify the potential losses that an institution may face due to adverse market movements. These metrics are typically focused on individual firm-level risk and are often based on the distribution of returns over a given time horizon. They are widely used in both regulatory frameworks and internal risk management to estimate potential capital shortfalls, liquidity buffers, or exposure to extreme market events. These traditional risk measures are a significant first step in preparing to implement other, more complex systemic risk measures.

3.1.2 Value at Risk

The VaR is one of the most widely used tools in financial risk management. It is the smallest amount that will be lost with a probability of no more than α throughout length t. The VaR risk measure answers: "How much could I lose, with a certain level of confidence, over a given period?" (Jorion, 2007; Hull, 2018). In this study, we use a 5% significance level (α = 0.05), which means we are 95% confident that losses will not exceed the calculated VaR on any given day. In more technical terms, VaR quantifies the loss distribution: it identifies the cutoff point below which only the worst 5% of losses fall. Formally, if $F_t(x)$ is the cumulative distribution function (CDF) of returns at time t, then:

$$\mathrm{VaR}_{\alpha,t} = -\, \mathbf{F^{-1}}_{\mathbf{t-1}}(\alpha)$$

This formula captures the threshold loss such that the probability of experiencing a worse outcome is just $1-\alpha$. "VaRs are just quantiles," which means the entire concept relies on estimating the correct return distribution and extracting the relevant tail percentile (Sanford and van Norden's, 2023). The pro of VaR is that it is a simple model: it summarizes risk in a single number, making it accessible for portfolio managers, traders, and regulators. It is also a core component of many regulatory frameworks, including Basel III, where banks must hold capital against potential losses measured by VaR (Basel Committee on Banking Supervision, 2011).

However, while intuitive and easy to communicate, VaR has its limitations. It does not provide any information about how bad losses can get if they exceed the VaR threshold. For example, two portfolios might have the same VaR, but one could be exposed to more severe losses in extreme scenarios. This is why measures like Conditional Value at Risk or Expected Shortfall have been developed to focus on what happens in the tail of the distribution, beyond the VaR cutoff (Rockafellar & Uryasev, 2000).

The sample analysis and display (see ANNEX 2 for all Systemic Risk measures data) show that the highest average VaR is in Europe and therefore represents the most severe risk profile with an average of -0.0342 and with largest tails as the minimum is -0.2312 and the highest being -0.0119. This shows a greater volatility for the banks in Europe but also that on average, banks on that continent tend to have larger losses with 95% confidence intervals. Europe is followed by banks in North America who show a VaR of 0.0271 on average, closely followed by Asia and Oceania. But, we clearly see a difference in standard deviation from Europe and North America compared to other world regions, which means that Europe and North America tend to have bigger movements compared to other regions. Under this metric, Europe stands out as the most systemically vulnerable region (Python self calculated based on daily log returns).

3.1.3 Expected Shortfall

ES is a systemic risk measure that can be used in combination with the VaR, because this risk measure calculates the average loss when the loss exceeds the VaR level (beyond the confidence threshold). While VaR tells us the minimum loss at a given confidence level, it stops there, it does not say anything about what happens beyond that threshold. Expected shortfall fills that gap by calculating the average loss when returns fall below the VaR level, effectively capturing what happens in the worst α% of cases (Hull, 2018). The ES explicitly focuses on fat tails and extreme values. Unlike VaR, which may be identical across banks with different exposure profiles, ES differentiates between institutions by accounting for the severity of losses beyond the threshold. Thus, two banks may report the same VaR but display different ES values if one is more vulnerable to large, infrequent losses. This makes ES particularly valuable for assessing the magnitude of extreme outcomes.

ES is sensitive to the inclusion or exclusion of low-probability, high-impact events, which can be both a strength and a limitation depending on the context (Krebs, 2021). However, this sensitivity enhances its capacity to depict downside risk more realistically than VaR alone. Given these characteristics, regulators such as the Basel Committee have incorporated ES into capital adequacy frameworks under Basel III, acknowledging its superior ability to capture systemic vulnerabilities during periods of financial turmoil (Basel Committee on Banking Supervision, 2016).

Here, Europe again emerges as the most systemically exposed region with an ES mean of -0.0514 and a minimum value of -0.3322, suggesting extremely severe losses during crisis periods, in the left tail. Europe is followed by South America and North America and their mean ES is 0.01 bigger than the one of Europe, which shows a 20% difference and reinforces the systemic exposure of European banks. For example, Middle East standard deviation is half the one of Europe, which can show that the Middle East is less exposed to the market and investors but that Europe is at the center of attention of investors and therefore more exposed to downturns. The ES results show that during market stress, European and American banks tend to suffer deeper drawdowns compared to their counterparts in possibly less integrated financial ecosystems.

3.2 Advanced Systemic Risk Measures

While traditional risk metrics such as the VaR and the ES capture the downside exposure of individual institutions, they fall short in accounting for interconnectedness and spillover effects within the financial system. To better assess the contribution of individual banks to systemic instability, more sophisticated, market-based measures have been developed. These advanced tools incorporate both tail risk and conditional dependencies, providing a more nuanced understanding of how distress in one institution may propagate through the broader network. In this section, we focus on three key measures widely used in systemic risk analysis: MES, CoVaR, Delta CoVaR, each offering a distinct perspective on financial contagion, marginal risk contributions, and capital shortfall under stress.

3.2.1 Marginal Expected Shortfall

In this thesis, we employ the MES as a primary measure to analyze the impact of environmental scores on systemic risk. MES quantifies the expected loss of a bank's equity value conditional upon severe market downturns (here represented by the S&P 500 in our sample), thus providing insight into how individual banks contribute to systemic risk during periods of significant financial stress (Acharya et al., 2017).

Marginal Expected Shortfall is formally defined as:

$$\mathrm{MES}_i = E(-r_i \mid r_m \leq \mathrm{VaR}_\alpha^m)$$

Where MES_i is the Marginal Expected Shortfall for bank i, r_i represents the return of bank i, r_m denotes market return, and VaR^m_α represents the VaR at the confidence level for the market return, capturing severe downturns (Brownlees & Engle, 2017). In practical terms, MES is calculated by taking a bank's average equity returns during periods when the overall market returns fall below the VaR at a certain quantile (here we use 5%). This measure then represents the bank's average losses under extreme market conditions (like a crisis) and effectively captures its systemic relevance.

Specifically, MES will serve as the dependent variable in our analysis, allowing us to empirically evaluate whether and how banks' E scores influence their systemic risk profiles. In the analysis, by correlating banks' MES values with their respective E scores, we will test our hypothesis to get a better understanding on the subject that is if higher E scores correlate with lower systemic risk contributions. This correlation analysis aims to establish whether banks that adopt robust environmental practices are less exposed to extreme financial market shocks or not.

Through integrating MES into our analytical framework, we aim to provide robust empirical evidence on the relationship between environmental sustainability practices and systemic risk. This analysis contributes valuable insights into the effectiveness of environmental initiatives post-COP21 in mitigating financial instability and enhancing the resilience of the banking sector (OECD, 2021; NGFS, 2020).

The summary statistics of MES show different results from the precedent traditional risk measures, indeed here, North America appears to have the lower (more risky) MES value, with an average of -0.0255 and same with the standard deviation, of -0.0183. Banks in North America tend to suffer more from market downturns, followed closely by Europe and South America. Other regions have relatively close to 0 MES values, which shows that banks in these regions tend to be less correlated to the market and therefore be subject to not as bad returns. Also, this could also be symptomatic of their lower integration or transparency in financial disclosures, rather than actual insulation from systemic dynamics (Battiston et al., 2017; IMF, 2020).

3.2.2 CoVaR

As seen in the precedent subpart, traditional risk measures like VaR are widely used to estimate the potential loss an individual institution might face under normal market conditions. However, VaR is a standalone risk measure that does not account for the way one institution's distress might affect the wider financial system. In response to this limitation, Adrian and Brunnermeier (2011) introduced CoVaR, as a way to measure the conditional risk of the financial system given that a specific institution is under stress.

The CoVaR measure is conceptually simple but powerful: it asks how much risk the entire system faces when a given institution is in a vulnerable state. It does this by conditioning the system's Value at Risk on the return of one specific institution. This makes CoVaR a systemic risk measure, rather than an individual one, and it is especially useful for identifying how institutions are connected through market dynamics, funding relationships, or exposure to common shocks (Brownlees & Engle, 2017).

Mathematically, CoVaR is defined as the VaR of the system, conditional on the return of a given institution being at a particular quantile:

$$\mathrm{CoVaR}_q^{\mathrm{system}|i} = \mathrm{VaR}_q\,(X^{\mathrm{system}}\,|X^{\mathrm{i}} = \mathrm{VaR}_q^{\,i})$$

Where X^{system} is the return of the system (for us the SP500), and X^{i} is the return of the individual institution i. The conditioning value X can be the institution's VaR at level α, or any other benchmark return. In practice, CoVaR is typically estimated using quantile regression, which models the relationship between the returns of the system and the returns of a specific institution at the tails of the distribution. This method allows researchers to directly capture tail dependence without assuming normality, which is an important feature given that financial return distributions often exhibit skewness and fat tails (Adrian & Brunnermeier, 2011).

The main strength of CoVaR is that it highlights interdependencies between institutions and the market as a whole. Unlike standard VaR, which focuses on an entity in isolation, CoVaR captures how financial distress in one institution might spill over to others, making it especially useful for macroprudential supervision (IMF, 2020). It has been used by central banks, regulators, and academics to track systemic vulnerabilities and assess whether some firms pose a greater threat to system-wide stability than others (Battiston et al., 2017).

When analyzing, Europe is once again at the top of the list, and displays the most critical CoVaR figures with a mean of -0.0575, and once again closely followed by North America with a mean of -0.0536. This observation confirms their pivotal role in amplifying contagion during crises. South America and the Middle East with mean of -0.0414 and -0.0359 respectively show intermediate CoVaR levels, while Asia and Oceania post the lowest conditional risk values, consistent with their low MES levels seen before. These modest CoVaR values imply that even when banks in these regions are under stress, the risk they impose on the entire system is comparatively subdued. This distinction reinforces the concept of "core-periphery" structures in financial networks, where distress in core regions triggers wider systemic consequences.

3.2.3 Delta CoVaR

Two main approaches are commonly used to assess a firm's systemic relevance, the first one is the supervisory approach which relies on firm-level accounting and regulatory data, often not publicly disclosed and in contrast, the market-based approach uses real-time, publicly available market data (such as equity prices and returns) to capture financial institutions' perceived risk and interconnectedness dynamically. One of the most prominent tools within the market-based approach is the Delta CoVaR methodology.

Proposed by Adrian and Brunnermeier (2011), the Delta CoVaR framework is an extension of the traditional VaR metric. While the VaR quantifies the loss an institution may experience at a given confidence level, it does not account for the spillover effects of that institution's distress on the rest of the financial system. Delta CoVaR addresses this limitation by focusing on the systemic impact of one institution's distress on the broader market or financial system. Specifically, CoVaR is defined as the VaR of the financial system conditional on a specific institution being under stress. Delta CoVaR is then calculated as the difference between the system's CoVaR in the distressed state and its CoVaR in a normal (median) state for that institution (Adrian & Brunnermeier, 2011).

Mathematically, this is expressed as:

$$\Delta \mathrm{CoVaR}_i^q = \mathrm{VaR}_q(\mathrm{R}_{sys} \mid \mathrm{R}_i = \mathrm{VaR}_q(\mathrm{R}_i)) - \mathrm{VaR}_q(\mathrm{R}_{sys} \mid \mathrm{R}_i = \mathrm{VaR}_{50}(\mathrm{R}_i))$$

This differential captures the marginal contribution of firm i to systemic risk. A higher Delta CoVaR implies that when an institution is in trouble, it significantly increases the downside risk for the rest of the system, demonstrating strong interconnectedness or contagion potential (Brunnermeier & Oehmke, 2013). This methodology is particularly valuable because it reflects market perceptions in real time, making it more sensitive to sudden shifts in market dynamics or stress episodes than traditional accounting-based metrics. Moreover, Delta CoVaR helps to rank institutions by their systemic importance, which is critical for regulators aiming to impose higher capital buffers on riskier firms or for investors trying to evaluate contagion risk in portfolios (IMF, 2020; Brownlees & Engle, 2017).

When looking at the summary statistics, we observe that all means are negative, which, without surprise, says that when banks in a region are in distress, they tend to amplify the systemic risk of the system. The results that we are looking at are interesting, indeed, North America and Europe show respectively higher ΔCoVaR values, which implies that banks in these regions are more systemic risk amplifiers than in the other regions when in distress. In contrast, Asia has the least negative mean ΔCoVaR at -0.0134, suggesting that Asian banks' distress contributes the least to amplifying systemic risk in their region.

The standard deviation (ΔCoVaR Std) shows how much variation there is within each region in how individual banks contribute to systemic risk. These values are quite similar across regions (ranging from about 0.022 to 0.027), indicating that while the average impact of banks in distress differs across regions, the dispersion within each region is similar. That means in each region, some banks are consistently more systemic than others.

Looking at the minimum ΔCoVaR values (around -0.10 to -0.11), these represent the most extreme cases where a bank's distress had the strongest systemic impact, highlighting that in some scenarios, individual banks could have caused significant systemic vulnerabilities. On the other hand, the maximum ΔCoVaR values are close to zero and positive for all (around 0.01), indicating that in the best scenarios, some banks' distress had almost no spillover effects on the system.

3.3 Regressions

3.3.1 Panel Regression

Once we obtained and standardized all the data on the same scale (log returns), the next part of the study is to perform a regression to understand the dynamics that exist between our variables. Firstly, a regression with only the E score will be run, followed by another regression incorporating control variables. This allows us to better explain variance and to account for influences unrelated to the E score itself. The data we have (113 banks observed over a 11-year period) has two key dimensions: a temporal dimension (years) and a transversal dimension (different companies). Therefore, we choose the panel regression framework using ordinary least squares (OLS), with standard errors clustered at the bank level to account for cross-sectional dependence within institutions. This approach is perfectly suited for our analysis because panel regressions are a widely used technique to observe the behavior of different variables over time, capturing both inter-variability across banks and intra-variability within each bank across years.

We employ a fixed effects model, including year fixed effects, which is essential when comparing across regions and for capturing critical periods such as the 2020 crisis. Based on our first hypothesis, that is banks with higher E Scores exhibit lower systemic risk, the model that we employ for the panel regression is the following :

$$\mathrm{SR}_{i,t}=\alpha+\beta_1\cdot \mathrm{E\_score}_{i,t}+\beta_2\cdot \mathrm{BankControls}_{i,t}+\beta_3\cdot \mathrm{MacroControls}_{t}+\gamma_t+\varepsilon_{i,t}$$

Where:

$\mathrm{SR}_{i,t}$ denotes the systemic risk measure (VaR, ES, MES, CoVaR, or ΔCoVaR) for bank i at year t, α is the intercept term, $\mathrm{E\_score}_{i,t}$ represents the E score of bank i in year t, $\mathrm{BankControls}_{i,t}$ is a vector of the bank-specific control variables, $\mathrm{MacroControls}_t$ is a vector of macro-level control variables that are constant across banks in a given year (e.g., three-month T-bill yields, inflation, VIX), γ_t captures year fixed effects, controlling for shocks that affect all banks simultaneously in a given year (e.g., policy changes post-COP21, global market events, macro shocks) and $\varepsilon_{i,t}$ is the error term.

To ensure robust inference in this clustered panel setting, robust covariance estimation is implemented to adjust standard errors for potential intra-bank correlation. In testing our central hypothesis, that banks with higher environmental scores tend to have lower systemic risk, the relationship we expect is positive (since systemic risk measures are signed, with values closer to zero indicating lower risk). If higher E scores are indeed associated with banks that are less systemically risky, the coefficient β_1 should be positive, indicating that better environmental performance coincides with a movement of systemic risk measures closer to zero.

3.3.2 Difference-in-Differences

One of the central methodological concerns in evaluating the impact of environmental scores on systemic risk lies in potential endogeneity, particularly from reverse causality and omitted variable bias. The reverse causality may arise if banks that are already less exposed to systemic risk voluntarily improve their environmental performance as part of a broader low-risk strategy, rather than environmental practices themselves reducing risk. For example, more conservative or well-capitalized banks might be both more resilient to systemic shocks and more likely to adopt proactive environmental policies to attract long-term investors. Additionally, some omitted variables may simultaneously influence both systemic risk and E score performance, including unobservable factors such as investor pressure (that we have discussed), where asset managers and institutional investors increasingly push for ESG compliance. Also, regulatory frameworks that vary across countries (e.g., France's Article 173 or the EU Taxonomy); and differences in national climate policies that create heterogeneous incentives for banks to strengthen environmental disclosure. Failing to account for these factors could bias the estimated relationship between E scores and risk exposure. In order to address these concerns and move toward causal inference, we implement a DiD strategy that compares banks in jurisdictions subject to stricter post-COP21 environmental regulations with those in less regulated settings, before and after policy implementation. This allows us to better isolate the effect of environmental performance from confounding influences.

The EU Taxonomy, introduced in 2020, is a regulatory framework that aims to define which economic activities can be considered environmentally sustainable. It is part of the European Union's broader strategy to promote sustainable finance and redirect capital flows toward green investments. For banks, this regulation marked a turning point by requiring greater transparency and accountability in how their activities align with environmental goals. As a result, it created a strong incentive for European financial institutions to improve their environmental performance. In the context of our thesis, we use the EU Taxonomy as a treatment in a DiD framework to examine whether an increase in environmental performance captured by E scores leads to a change in systemic risk within the banking sector.

For that DiD, we will compare banks from North America vs banks in Europe that are under the EU taxonomy from 2020. Before running any regression, we first compare the average E scores and systemic risk levels of European and U.S. banks before and after 2020 as a preliminary analysis. The descriptive statistics show that European banks experienced a sharp increase in E scores following the implementation of the EU Taxonomy, jumping from an average of 0.34 before 2020 to 2.67 after. In comparison, U.S. banks also saw an increase in E scores from 0.43 to 1.71, which is good but less pronounced than the rise in Europe. This suggests that the EU regulation may have had a real impact on how seriously banks in Europe approached environmental criteria. Then, when we look at the systemic risk mean after 2020, we observe that both European and North American banks became more systemic after 2020, probably due to the Covid 19 crisis as already said but also the return of war on European soil. When looking in detail though, we clearly see that the systemic risk increase is higher for banks in North America post EU Taxonomy, for all systemic risk measures. These trends suggest that while E scores improved in Europe, this was accompanied by a lower increase in systemic risk than other banks without this regulation (less negative systemic risk measures).

The regression model employed in this analysis follows a classical DiD specification designed to estimate the impact of the EU Taxonomy regulation on the systemic risk profile of European banks, using North American banks as a control group. The regression equation takes the form:

$$y_{\mathbf{i,t}} = \boldsymbol{\alpha} + \boldsymbol{\beta_1} \cdot \mathbf{d_{i,t}} + \boldsymbol{\beta_2} \cdot \mathbf{D_i} + \boldsymbol{\beta_3} \cdot \mathbf{D_t} + \gamma' X_{\mathbf{i,t}} + \boldsymbol{\varepsilon_{i,t}}$$

In this equation, the dependent variable $y_{\mathbf{i,t}}$ corresponds to a systemic risk measure, calculated annually for each bank i over time t. The key variable of interest is the interaction term $\mathbf{d_{i,t}}$ (EU_x_Post2020), which equals 1 for European banks in the post-2020 period (i.e., after the implementation of the EU Taxonomy), and 0 otherwise. The coefficient β1 on this interaction term captures the causal effect of the EU Taxonomy on systemic risk, net of global trends and regional fixed effects.

The model also includes a treatment indicator D_i (EU), which distinguishes European banks from their North American counterparts, and a time indicator D_t (Post2020), which captures general time-related effects, such as the COVID-19 crisis or global regulatory shifts. Together, β_2 and β_3 isolate baseline differences between groups and across time. To control for confounding influences, the model incorporates a rich set of bank-specific variables $X_{i,t}$, including market volatility (VIX, T-bill rate variability, inflation) and financial characteristics such as leverage, market capitalization, profitability (ROA), and liquidity. These controls, represented by the vector γ, enhance the model's robustness by accounting for known drivers of systemic risk, ensuring that the estimated treatment effect is not biased by omitted variable concerns. Finally, standard errors are clustered at the bank level to correct for within-bank serial correlation. This structure allows for a precise estimation of how the EU Taxonomy affected the systemic stability of European financial institutions.

Also, one of the objectives of analysis is to assess the impact of the E score post COP21, basically when the E score was implemented on Bloomberg. For that, we want to compare banks that kept an E score of 0 (even after the COP21 when other banks started to increase their E score) to those who started to have an E score superior to 0 after 2015.

For that, we use a DiD model similar to the one above, a model that uses dummy variables in order to isolate the treatment effect. In our panel data, the banks showing a E score > 0 after 2015 being the treated group and the ones with constant E Score = 0 from the all period, even after 2015, the untreated (or control).

For our panel data $y_{i,t}$ we have observations on banks (i) that include the treated and control banks and over time (t) before and after the treatment. Therefore, we have dit, the dummy variable that indicates treatment, so $d_{i,t}$ = 1 for treated banks, after they are treated (starting 2016 until 2023). The objective is to find an effect on the treated group and then see if banks with Higher E score tend to be less systemic or not after 2015. Also, to ensure the robustness of our results, we will use control variables to capture other effects that may affect the systemic risk in the model, as the E score cannot solely capture the effect by itself, we will then use and try the control variables that we have in our data.

The formula for the DiD for post COP21 is the same as the one above:

$$\boldsymbol{y_{i,t} = \alpha + \beta_1 \cdot d_{i,t} + \beta_2 \cdot D_i + \beta_3 \cdot D_t} + \gamma' X_{i,t} + \boldsymbol{\varepsilon}_{i,t}$$

Where $y_{i,t}$ represents the systemic risk measure for bank i at time t, $D_i = 1$ for the treatment group, $D_t = 1$ after the treatment group is applied, $X_{i,t}$ a vector of our control variables and $\varepsilon_{i,t}$ is the error term.

Chapter 4 : Results

4.1 Panel Regression Results

When we run the panel regression for the 5 systemic risk measures, we look at the coefficient before the E score. As our systemic risk values are non absolute, a negative coefficient would mean that the systemic risk is more negative, and therefore more risky. On the other hand, a positive coefficient before the E score would mean a closer to zero systemic risk measure and therefore less risky. The first regression that we run is a simple panel regression with just the E score as a variable (see Table 1), in order to see the tendency and the relationship between the E score and the systemic risk measure. We observe that it confirms what we looked at before when analyzing the risk measures per region and the E score, as the coefficients for MES, CoVaR and Delta CoVaR are negative and statistically significant at the 5, 10 and 1% confidence, respectively. Indeed, it shows that banks with higher E scores would be more systemic as the coefficients are all negative, reflecting the preliminary analysis for E score and systemic risk measures where regions with higher E score mean were showing lower systemic risk measures in average. These simple results mean that as the E score increases, the systemic risk contribution tends to increase, we will try to dig into that throughout the analysis with different models.

TABLE 1: Panel Regression for SRs (2016–2023, E Score only)

	VaR	ES	MES	CoVaR	DeltaCoVaR
E Score	0.0003	0.0002	-0.0012**	-0.0017*	-0.0009***
	(0.000)	(0.001)	(0.001)	(0.001)	(0.000)
Years FE	Yes	Yes	Yes	Yes	Yes
Observations	896	896	896	896	896
R^2	0.409	0.436	0.514	0.578	0.909

Note:
*Significance level denoted as * for p<0.1, ** for p<0.05, and *** for p<0.01.*
Standard Errors are robust to cluster correlation (cluster)

What we suppose is that the Covid 19 crisis in 2020 will have an impact on these results, as we know that the larger banks (in Europe and North America) had the larger losses during it, and therefore it can "bias" our results because the crisis should be captured by other variables. Therefore, we run another model where we exclude the year 2020 (see Table 2), in order to compare with the first model. When looking at them, we observe a difference indeed, we see that the coefficient for E score for the VaR and ES are positive, which means that systemic risk would decrease when the E score tends to increase, but they are not statistically significant, so we cannot conclude anything on them. For the MES and CoVaR, the coefficients are still negative, but less than previously, and once again, these are not statistically significant, therefore we cannot draw any conclusions, but we observe that without the large movements of 2020, it is not sure either banks with E score are more systemic or less. The Delta CoVaR, though, shows a negative coefficient and statistically significant at 1%, showing that as the E score increases, the systemic risk contribution increases.

TABLE 2: Panel Regression for SRs (2016–2023, excluding 2020, E Score only)

	VaR	ES	MES	CoVaR	Delta CoVaR
E Score	0.0004	0.0005	-0.0007	-0.0007	-0.0009***
	(0.000)	(0.001)	(0.000)	(0.001)	(0.000)
Years FE	Yes	Yes	Yes	Yes	Yes
Observations	784	784	784	784	784
R^2	0.177	0.146	0.282	0.228	0.308

Note:
*Significance level denoted as * for p<0.1, ** for p<0.05, and *** for p<0.01.*
Standard Errors are robust to cluster correlation (cluster)

Now, after running these simple models, we look at them and we include the control variables, in order to capture some variance not captured before, and some movements not explained by the E Score itself, in order to have less bias in the model. We want to reflect the real effect of the E score on systemic risk measures with variables that can capture movements not explained before.

First, we run the model with all control variables (bank specific and macro, see Table 3) that we mentioned earlier, we clearly observe that more variance is captured for each systemic risk measure, and that some movements (e.g. crisis, etc) are captured by these variables. Again, we see that the coefficients for the VaR and ES are very close to 0 and insignificant, which means that there is no proof that the E score impacts positively or negatively these metrics, and therefore the risk and the maximum loss among the banks. But, when looking at the other systemic risk measures, we observe that the coefficients are negative, -0.0012 for the MES, -0.0020 for CoVaR and -0.0009 for the Delta CoVaR, statistically significant at the 5, 5 and 1% confidence respectively, it therefore confirms the results from the first simple regressions that say that as the E score increases, the systemic risk contributions from a bank to others increases as well, reflecting what we observed when looking at the preliminary analysis with the average systemic risk measures and the average E score per region, we observed that regions with high E scores were displaying the more negative average for all systemic risk measures, which seems counterintuitive and showing that higher environmental practices would increase the risk of spillover. We will dig into that in the future models and robustness checks in order to really understand the movements.

TABLE 3: Panel Regression for SRs (2016-2023)

	VaR	ES	MES	CoVaR	Delta CoVaR
E Score	-0.0000	-0.0002	-0.0013**	-0.0021**	-0.0009***
	(0.000)	(0.001)	(0.001)	(0.001)	(0.000)
Years FE	Yes	Yes	Yes	Yes	Yes
Observations	868	868	868	868	868
R^2	0.455	0.487	0.514	0.601	0.905

Note:
*Significance level denoted as * for p<0.1, ** for p<0.05, and *** for p<0.01.*
Standard Errors are robust to cluster correlation (cluster)

4.2 Difference-in-Differences Results

We now look at the DiD results in order to get answers for both our hypotheses, the first DiD will help answering if the introduction of the E Score as a measure itself helped reducing systemic risk within banks and the risk spillover, the second one using the EU Taxonomy as a treatment wil; help us answering the first hypothesis that is if banks with higher E Score show lower systemic risk and spillover, that is a higher E Score would have a significant and positive impact on the systemic risk measures (lowering the risk).

4.2.1 Difference-in-Differences Post COP21

For the DiD model, we want to know if the banks that adopted Environmental measures after 2015 and the COP21 agreements, tend to contribute less to the systemic risk than the ones that still have a score of 0, even in 2023.

Therefore, we split our panel data in 2 parts, as mentioned in the previous Difference-in-Differences subpart. In the VaR regression (see Table 4), the interaction term is positive and statistically significant at the 5% threshold (0.0049, p=0.013). This suggests that treated banks had higher VaR after COP21, meaning they became relatively less risky among themselves than the control banks when they started to have an E Score. Similarly, in the ES model, the interaction term is positive and significant (0.0074, p=0.038), indicating that treated banks saw a relative improvement in ES, becoming less risky after 2016, showing the same effect of the E score. The treatment dummy itself is negative and significant, showing that banks with positive E scores generally had lower ES levels.

This trend continues in the MES regression, where the interaction term is again positive and significant (0.0027, p=0.025). This result implies that banks with better environmental performance had a relative reduction in their marginal systemic risk exposure in the post-COP21 years. For CoVaR, the interaction term is positive and significant at the 10% level (0.0054, p=0.100), suggesting a marginal improvement in conditional systemic risk for treated banks, although this evidence is weaker than for MES and Delta CoVaR. And, inn the case of Delta CoVaR, the interaction term is positive and significant at the 5% level (0.0019, p=0.048), showing that banks with E scores improved their marginal contribution to systemic risk in the period after COP21.

The year fixed effects capture the broader economic context, confirming that these findings reflect the direct impact of environmental performance improvements on systemic risk. Altogether, these results suggest that higher environmental scores may have helped reduce the severity of systemic risk in the banking sector after COP21, especially for MES and Delta CoVaR. These results are different and give a different conclusion compared to the OLS panel regression, because here we do not look at the scale of the E score, indeed, as soon as banks have a E score > 0 in the 2016-2023 period, we consider them as treatment, even though some will have a score around 5 and some around 0.5. It is important to mention that because in this DiD framework, any E score above zero is treated the same regardless of whether it is just slightly above zero or significantly higher. As a result, the DiD analysis here does not capture the scale or intensity of the environmental score but rather focuses on the mere presence of a non-zero E score as a threshold effect. This explains why we see positive and significant coefficients on the interaction term, suggesting that simply having some environmental focus (even minimal) is better than having none at all.

TABLE 4: Difference-in-Differences for SRs (2012–2023)

	VaR	ES	MES	CoVaR	DeltaCoVaR
Treatment_X_Post	0.0049**	0.0074**	0.0027**	0.0054*	0.0019***
	(0.002)	(0.004)	(0.001)	(0.003)	(0.001)
Observations	1243	1243	1243	1243	1243
R^2	0.297	0.315	0.497	0.565	0.897

Note:
*Significance level denoted as * for p<0.1, ** for p<0.05, and *** for p<0.01.*
Standard Errors are robust to cluster correlation (cluster)

In contrast, the panel OLS regressions, which treat the E score as a continuous variable, actually showed a negative coefficient for E score showing that an increase in E score is associated with an increase in systemic risk. The two analyses show slightly different effects: the Difference-in-Differences approach captures a threshold effect being part of the "treatment" group (any E score) vs. not, while the OLS models highlight the continuous relationship of higher E scores potentially correlating with higher risk. It's a good reminder that the magnitude of E scores and the mere presence of an E score tell different stories, and both angles are important to interpret for your thesis.

4.2.2 Difference-in-Differences EU Taxonomy

The DiD regression results comparing European and U.S. banks before and after the introduction of the EU Taxonomy in 2020 provide strong empirical insight into the potential stabilizing effect of environmental regulation on banks. The key interaction variable, which captures the joint effect of being an EU bank and operating post-2020, shows positive and statistically significant coefficients in most systemic risk models. This indicates that systemic risk increased but less for European banks relative to their U.S. and Canada counterparts after the EU Taxonomy came into force in Europe.

Looking first at the VaR model (see Table 5), the interaction term is positive and highly significant at the 1% threshold, suggesting that while global systemic risk worsened after 2020, likely due to external pressures like the pandemic and macroeconomic shocks, EU banks were relatively more resilient than American banks. This relative improvement in resilience can plausibly be attributed to the stronger environmental regulatory framework established by the EU Taxonomy, which seems to have bolstered investor confidence or improved risk governance among EU financial institutions. We can observe the same results when looking at the ES where the coefficient for the EU interaction is also positive, though only marginally significant at the 10% threshold (nearly 5), indicating that while EU banks may not have fully avoided increased risk, they at least absorbed the shock better than banks not subject to EU taxonomy.

The MES regression is particularly revealing, the interaction between EU banks and the post-taxonomy period is positive and strongly statistically significant (0.0135) at the 1% threshold. This highlights a meaningful drop in the marginal contribution of EU banks to systemic risk following the implementation of the taxonomy. It suggests that the EU regulatory environment may have reduced exposures to high-risk sectors, therefore lowering the banks' vulnerability during financial downturns. The significance of this effect aligns closely with the hypothesis that regulatory initiatives promoting sustainability can enhance financial stability, especially in the face of systemic shocks.

Looking at the CoVaR model, the interaction term is again positive and significant (at the 10% threshold). The direction of the effect supports the broader narrative and what was observed with the other measures, also confirming the preliminary analysis: EU banks seem to have experienced less worsening of conditional systemic risk compared to U.S. banks during tough times and crisis. Finally, in the Delta CoVaR regression, the interaction effect is positive but statistically insignificant suggesting that the specific impact a bank has on the financial system during times of stress remains relatively unchanged, regardless of the regulatory context. It is possible that Delta CoVaR captures more persistent structural risks that are less responsive to policy changes in the short term.

Beyond the main Difference-in-Differences variables, the regression results also provide information on the role of financial controls. Several indicators consistently show significant associations with systemic risk. For example, higher volatility in the VIX index is associated with increased systemic risk, a finding that holds true across models and reaffirms the importance of market stress as a systemic factor, the VIX explains a large part of the variance in the model, thus justifying the high R^2 values. Also, profitability indicators such as return on assets and price-earnings ratio also tend to reduce systemic risk, suggesting that more financially stable and better valued banks contribute less to systemic instability, which makes sense here. Other variables, such as debt or debt ratios, exhibit mixed effects, reflecting the complexity of interpreting financial fundamentals in the context of systemic risk.

Overall, these results for this DiD analysis reinforce the central argument of this thesis and for the first hypothesis that regulatory efforts to integrate environmental responsibility into financial practices can have measurable effects on the stability of the banking system and within the banks themselves. Although not all measures show a high importance, an overall effect seems to be emerging with these results: the European banks, supported by the structure and expectations of the EU taxonomy, have supported systemic pressures post-2020 more effectively than their North American peers. These results suggest that the integration of environmental considerations can indeed act as a hedge against systemic risk.

TABLE 5: Difference-in-Differences for SRs (2012–2023) EU vs NA banks Post EU Taxonomy

	VaR	ES	MES	CoVaR	DeltaCoVaR
Euro_X_Post	0.0068***	0.0063*	0.0046***	0.0064*	0.0008
	(0.002)	(0.003)	(0.001)	(0.003)	(0.001)
Observations	704	704	704	704	704
R^2	0.639	0.598	0.696	0.616	0.883

Note:
*Significance level denoted as * for p<0.1, ** for p<0.05, and *** for p<0.01.*
Standard Errors are robust to cluster correlation (cluster)

4.3 Robustness Check

4.3.1 Region sample: North America & Europe

Here, we run a panel regression model, only focusing on the North America and Europe regions (see Table 6). Indeed, when analyzing the returns and E score from the other regions such as the Middle East, Africa, Asia, we observe a low correlation between these regions and Europe and North America. The analysis is restricted to these two regions because their financial sectors are more deeply integrated into global capital markets. Banks in these regions exhibit higher trading volumes, cross-border exposures, and tighter alignment with international regulatory frameworks. As such, they are more responsive to market-wide stress events and systemic fluctuations, while some banks in the Middle East or Africa may be affected by localized factors such as geopolitical instability, commodity shocks, or government intervention, which may dilute the relationship between ESG performance and traditional market-based risk indicators. When looking at the mean returns and mean E score, we could tell that regions (Europe, North America, Oceania) with higher E scores, were showing higher volatility and lower systemic risk values in average, showing a negative relationship, something that we saw previously in the OLS with all regions and banks, that high E score would lead to lower systemic risk measures' values, and therefore leading to higher systemic risk and risk spillover. These results and reasoning drove us to do an analysis of these two developed regions only, to compare banks within these regions and to understand, within these two regions, showing the higher E score mean, if a higher E score leads to higher or lower systemic risk. The model used is the same as the one in the previous panel regression. So, here, we present the results of the OLS regressions examining the relationship between E scores and the 5 systemic risk measures chosen. Focusing on Europe and North America, this analysis captures systemic patterns more directly shaped by market dynamics, making the results more interpretable and generalizable in global financial risk.

TABLE 6: Panel Regression North America & Europe only (2016-2023)

	VaR	ES	MES	CoVaR	DeltaCoVaR
E_Score	0.0011***	0.0004	0.0019***	0.0015*	0.0010***
	(0.000)	(0.001)	(0.000)	(0.001)	(0.000)
Observations	696	696	696	696	696
R^2	0.625	0.566	0.651	0.633	0.877

Note:
*Significance level denoted as * for p<0.1, ** for p<0.05, and *** for p<0.01.*
Standard Errors are robust to cluster correlation (cluster)

The VaR model results show a positive and statistically significant coefficient for the E Score (0.0011, p = 0.001), suggesting that banks with stronger environmental performance are associated with less negative, meaning less severe, extreme downside risk. This confirms the hypothesis that better E Scores may help reduce the systemic vulnerability during financial stress events. The overall model fit is relatively strong, with an R-squared of 0.632, showing that most of the variance is explained by the model and the control variables used. Several firm-level controls, such as PERatio_return, ROA_return, and Volume_return, also show significant effects in expected directions, reinforcing that profitability and market conditions influence banks' tail risk exposure. Notably, macroeconomic controls such as inflation, Treasury Bill volatility (Tbill3m_std), and VIX_std are all highly significant, highlighting the sensitivity of VaR to shifts in broader market uncertainty.

For the ES model, the E Score coefficient is positive but statistically insignificant (0.0004, $p = 0.611$). This suggests that while E Scores may be associated with reducing average extreme losses, the effect is not robust across all tail risk measures. The model's explanatory power remains substantial, with an R-squared of 0.573, and again, the macro variables retain strong statistical significance. Interestingly, PERatio_return and ROA_return exhibit significant adverse effects, indicating that higher profitability and valuation ratios are associated with reduced ES, which is consistent with financial theory.

In the case of MES, the E Score has a substantial and statistically significant positive coefficient (0.0019, $p < 0.001$) at the 1% level. This again suggests that banks with better environmental practices within thses two regions tend to contribute less to systemic losses during distress periods. The R-squared of 0.657 is one of the highest among all models and this reinforces the systemic resilience benefit of ESG integration in market-integrated economies.

The regression analysis for CoVaR yields a positive and marginally significant coefficient for the E Score (0.0015, $p = 0.063$), which falls just below the conventional 10% significance threshold. Although this result does not reach the stricter 5% level, a p-value under 0.10 is typically interpreted as strong evidence against the null hypothesis, indicating a statistically plausible association between higher E scores and lower systemic risk contribution.
In the context of CoVaR, a higher value reflects greater conditional system-wide losses when an individual institution is under distress. Consequently, a positive coefficient on the E Score suggests that as a bank improves its environmental performance, its CoVaR becomes less negative, moving closer to zero, indicating a reduction in the severity of spillover risk the bank imposes on the financial system during downturns. In other words, banks with better environmental credentials are less likely to amplify system-wide distress when facing financial stress themselves.

Although the statistical strength of this relationship is modest, the directionality is consistent with the hypothesis that environmentally proactive banks are more resilient and less systematically dangerous. This result supports the notion that sustainable banking practices may offer external benefits, not only enhancing firm-level stability but also contributing to broader financial system robustness, showing greater trust of investors when the market is in distress.

Finally, the Delta CoVaR regression, which measures the marginal impact of a bank's distress on the system, shows a highly significant and positive E Score coefficient (0.0010, $p < 0.001$) at the 1% threshold. This reinforces the idea that environmentally stronger banks are less likely to amplify systemic shocks under stress. The model shows excellent fit, with an R-squared of 0.877, and VIX_std is again a dominant predictor, confirming that volatility expectations play a central role in systemic propagation, capturing the major events and crisis as the Covid 19 crisis or the war in Europe for example.

Overall, the regression results confirm that for banks in Europe and North America, environmental performance measured through E Scores has a statistically significant and positive association with lower systemic risk across multiple metrics, especially those focused on extreme or marginal impacts like VaR, MES, and Delta CoVaR. The findings highlight that environmental engagement can stabilize financial markets, particularly for institutions operating in globally integrated and liquid environments. These effects are less pronounced or insignificant in measures like ES and CoVaR, suggesting that while ESG can improve resilience, its protective effects are more concentrated in high-risk scenarios than in average systemic dependence.

4.3.2 Covid 19 sample: North America & Europe, excluding 2020

In this section, we refine the analysis by excluding the year 2020 from the regression sample and keeping Europe and North America regions, to compare with the previous regression (see Table 7). This decision is grounded in both methodological and economic reasoning. The year 2020 was marked by the onset of the COVID-19 pandemic, which triggered an unprecedented and globally synchronized market shock. Systemic risk measures during this period were not only unusually volatile but also heavily driven by exceptional policy responses, temporary market closures, and liquidity interventions that may not reflect the structural risk dynamics of the banking system. Including such a year could introduce noise, obscure long-term relationships, and potentially bias the estimation of the E Score’s effect on systemic risk, even though some control variables as the VIX capture the movements well, as mentioned just before. By removing 2020, the analysis focuses on periods that are more representative of regular financial market functioning, thereby improving the interpretability and robustness of the results.

TABLE 7: Panel Regression North America & Europe only (2016-2023, excluding 2020)

	VaR	ES	MES	CoVaR	DeltaCoVaR
E_Score	0.0009***	0.0002	0.0020***	0.0018***	0.0009***
	(0.000)	(0.001)	(0.000)	(0.001)	(0.000)
Observations	636	636	636	636	636
R^2	0.529	0.413	0.333	0.176	0.311

Note:
*Significance level denoted as * for p<0.1, ** for p<0.05, and *** for p<0.01.*
Standard Errors are robust to cluster correlation (cluster)

The VaR regression excluding 2020 still yields a positive and statistically significant coefficient for the E Score (0.0009, p = 0.003) at the 1% level. This indicates that higher E Scores are associated with reduced downside risk for banks with stronger environmental performance. Compared to the full-sample model with 2020 included, the magnitude of the coefficient is slightly lower, and the R-squared decreases from 0.632 to 0.538. This drop is expected, as the extreme observations in 2020 contributed significantly to overall variation. Importantly, the significance of the E Score remains, supporting the robustness of the finding that environmental strength improves risk resilience under normal market conditions.

For the ES model, the E Score is once again statistically insignificant (0.0002, p = 0.819). This suggests that even with 2020 being removed, there is insufficient evidence to conclude that E Scores influence the average loss in extreme scenarios, which may point to the fact that E Scores mainly help in reducing the tail end of risk, not necessarily the mean shortfall during extreme events, that may be driven by other factors, such as fear and that investors may not be looking at the ESG components (who become secondary) when facing depression or extreme events.

For the MES regression, the E Score remains positive and highly significant at the 1% level (0.0020, $p < 0.001$), confirming that environmentally stronger banks continue to exhibit lower marginal risk exposure, even when 2020 is excluded. The MES is particularly useful in capturing how a single institution contributes to systemic risk under stress. The significance of the E Score in this model aligns with the notion that better environmental management is associated with a lower likelihood of distress that coincides with or amplifies broader market turmoil.

Similar result for the CoVaR regression which also yields a statistically significant E Score coefficient (0.0018, $p = 0.007$), maven more significant than from the previous model as it is significant at the 1% threshold this time, reinforcing the hypothesis that banks with higher environmental performance contribute less to the conditional risk of the system. This result holds even after excluding 2020, which suggests that the systemic dependence captured by CoVaR is sensitive to ESG quality in a structurally consistent way.

Finally, the Delta CoVaR model reports a positive and highly significant at the 1% threshold, as the E Score coefficient is 0.0009, and $p = 0.002$. Delta CoVaR captures the marginal systemic impact of a bank when it moves from a normal state to distress. The consistent significance of E Scores in this specification indicates that environmentally stronger banks are less likely to exacerbate systemic dislocations when stressed. This result supports the growing argument that ESG integration, particularly on the environmental dimension, acts not only as a corporate responsibility strategy but also as a systemic risk mitigation mechanism.

In summary, after excluding the atypical year of 2020, the E Score maintains a positive and statistically significant relationship with the majority of systemic risk measures, particularly VaR, MES, CoVaR, and Delta CoVaR, with reinforced significance. This confirms the robustness of earlier findings and underscores the value of environmental performance in strengthening systemic resilience.

4.3.3 Falsification Test

To evaluate the hypothesis that the introduction of environmental practices following the 2015 COP21 agreement has a positive effect on reducing systemic risk, we employed a DiD regression framework that allowed us to estimate the causal impact of an intervention by comparing the change in outcomes over time between a treated group and a control group. In our case, the treated group consists of banks that received non-zero environmental scores (E scores) starting in 2016, while the control group includes banks for which no E score is available throughout the period. By focusing on changes in the systemic risk indicators we use, we can assess whether the emergence of environmental scoring is associated with a structural decline in risk exposure within the banking sector. The main DiD estimates reveal that the interaction term Treatment X Post is statistically significant across all these measures, suggesting a consistent link between the implementation of environmental practices and lower levels of systemic risk. However, as with any DiD approach, the credibility of the causal interpretation relies heavily on the parallel trends assumption, which requires that, in the absence of treatment, the treated and control groups would have experienced similar trends in the outcome variable. Since this condition cannot be verified directly, it is standard practice to conduct falsification tests to assess its possibility. These tests work by simulating the treatment in a period when it should not have had any effect; in our case, we randomly set a fake treatment year to 2018. Then, we run the regression and examine whether a significant treatment effect still appears and with a similar coefficient or not.

In this study, this test offers valuable insights into the robustness of our findings, indeed, the interaction term for this falsification specification is statistically insignificant for ES, CoVaR, and Delta CoVaR, which reinforces the credibility of the original results and supports the notion that spurious correlations or pre-treatment dynamics do not drive the observed effects. This strengthens the internal validity of the causal claim that environmental scoring reduces systemic risk. While the effect for VaR is statistically significant, and MES shows marginal significance ($p = 0.072$), these exceptions may be due to metric-specific sensitivities, data noise, and other factors. It is also possible that VaR, as a tail-risk measure focused on threshold exceedance, reacts more sharply to evolving risk perceptions that precede formal scoring.

The use of falsification tests in this context is aligned with best methodological practices in applied econometrics and corporate finance research. Such tests are crucial in rejecting alternative explanations and ensuring that the timing of the estimated treatment effect aligns with the actual policy or institutional change under investigation. They can serve as a methodological safeguard, particularly when randomized experiments are infeasible and treatment assignment is not entirely exogenous. By demonstrating that systemic risk measures remained unchanged during the "placebo" period for most metrics, we provide strong empirical support for the assumption that the reduction in systemic risk observed after 2016 is attributable to the integration of environmental practices, rather than to unrelated pre-existing trends.

TABLE 8: Falsification Test : 2018 as the falsified year

	VaR	ES	MES	CoVaR	DeltaCoVaR
E_Score	0.0041**	0.0052	0.0021*	0.0042	0.0016
	(0.002)	(0.003)	(0.001)	(0.004)	(0.002)
Observations	1199	1199	1199	1199	1199
R^2	0.473	0.487	0.486	0.586	0.893

Note:
*Significance level denoted as * for p<0.1, ** for p<0.05, and *** for p<0.01.*
Standard Errors are robust to cluster correlation (cluster)

The consistency of the Difference-in-Differences results, combined with the reassuring evidence from the falsification test, provides robust support for the hypothesis that the post-COP21 institutional shift toward environmental accountability has had a meaningful effect on reducing systemic risk in the global banking sector. This finding can contribute to the growing and existing literature on sustainable finance by offering quantitative evidence that environmental performance is now a reputational or ethical concern but is also materially linked to financial system stability.

4.3.4 Quantile Regression & Results

Following the unexpected results of the panel regressions we ran and the contrary results between these and the DiD, in order to understand and compare how environmental performance influences systemic risk across the risk distribution, we apply the quantile regression at three specific points: the 10th percentile (q = 0.1), the 25th percentile (q = 0.25), and the 75th percentile (q = 0.75). Unlike the traditional OLS regression, which estimates the average effect of predictors on the dependent variable, the quantile regression captures the heterogeneity of these effects across different distribution segments. This approach is relevant in systemic risk research, where understanding the behavior of institutions in the tails of the risk distribution, particularly in extreme loss scenarios, is critical. By examining these distinct quantiles, we can identify whether the relationship between E Scores and systemic risk is consistent or more pronounced under specific risk conditions.

The selection of these quantiles is grounded in financial risk theory. The 10th percentile corresponds to the lower tail of the distribution, where systemic losses are most severe. Estimating the impact of E Scores at this level allows us to observe whether environmental performance plays a protective role during periods of extreme stress. The 25th percentile captures a moderately risky environment, reflecting less extreme but relevant vulnerability, while the 75th percentile shifts the focus to higher returns, not being risky ones, testing whether the same explanatory power persists when not looking at low returns only. This layered perspective offers a more refined understanding of the drivers of systemic risk.

At the 10th percentile (see Table 9), the results show that environmental performance statistically impacts several systemic risk measures. The coefficient for the E Score in the VaR model is 0.0010, with a t-value of 2.850 and a p-value of 0.004, indicating a strong and statistically significant positive relationship at the 1% level. This suggests that higher E Scores are associated with lower VaR in the tail of the distribution, implying that stronger environmental performance helps mitigate extreme losses, at the company's level. A similar effect is observed for the MES, where the E Score has a coefficient of 0.0012 ($t = 2.775$, $p = 0.006$) statistically significant at the 1% level, again supporting the hypothesis that banks with higher environmental standards face lower downside exposure in severe stress scenarios. For CoVaR, the coefficient is 0.0023 with a t-value of 2.197 and a p-value of 0.028, statistically significant at the 5% level, it shows that when the bank is doing poorly, it has less impact on the global financial system as the E score is higher. These findings confirm that E Scores significantly reduce systemic spillovers in tail events. In contrast, the results for ES and ΔCoVaR are not statistically significant at this quantile, suggesting that while E Scores help cushion against individual and marginal systemic risk, their explanatory power may not extend equally to all risk measures in extreme conditions.

TABLE 9: Quantile Regression for q=0.1 (2016-2023)

	VaR	ES	MES	CoVaR	DeltaCoVaR
E_Score	0.0010***	0.0004	0.0012***	0.0023**	0.0000
	(0.000)	(0.001)	(0.000)	(0.001)	(0.000)
Observations	1191	1191	1191	1191	1191
R^2	0.3547	0.3791	0.4314	0.4797	0.7738

Note:
*Significance level denoted as * for p<0.1, ** for p<0.05, and *** for p<0.01.*
Standard Errors are robust to cluster correlation (cluster)

When moving to the 25th percentile (see Table 10), the effect of E Scores remains positive but becomes less pronounced. The coefficient for the VaR and MES is equal to 0.0006 for both, with significance at the 1 percent and 10 percent levels, respectively. This indicates that even in moderately risky environments, environmental performance continues to have a beneficial, though weaker, influence on systemic risk. However, the statistical significance observed in the CoVaR models at q = 0.1 is no longer present at q = 0.25, indeed the coefficients for the ES, CoVaR and Delta CoVaR are non statistically significant in the regression, despite their positive value. These results at this quantile suggest that the influence of E Scores on systemic risk is strongest in the most adverse market scenarios and gradually diminishes as one moves toward the median of the risk distribution. Nonetheless, the directionality of the coefficients remains consistent, reinforcing the idea that environmental performance contributes positively to the systemic risk in the banking industry.

TABLE 10: Quantile Regression for q=0.25 (2016-2023)

	VaR	ES	MES	CoVaR	DeltaCoVaR
E_Score	0.0006***	-0.0001	0.0006*	0.0002	0.0002
	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)
Observations	1191	1191	1191	1191	1191
R^2	0.2895	0.3255	0.3031	0.3736	0.6088

Note:
*Significance level denoted as * for p<0.1, ** for p<0.05, and *** for p<0.01.*
Standard Errors are robust to cluster correlation (cluster)

Now, looking at the 75th percentile (see Table 11), the model's explanatory power declines more substantially. The pseudo R-squared falls to 0.1375, indicating that the variance in the model is less well captured at this level. Among all systemic risk measures, only VaR retains a statistically significant relationship with the E Score, though the strength of the coefficient is weaker compared to the lower quantiles. No significant effects are found for MES, CoVaR, ES, or ΔCoVaR. This significant drop implies that the impact of environmental performance is concentrated in the lower part of the risk distribution, where stress is most acute. In normal conditions, E Scores appear to have limited influence and does not have similar impact in the model. These results highlight an important asymmetry: while environmental performance is a buffer against severe tail risks, it does not exert the same explanatory power when risk levels are low.

TABLE 11: Quantile Regression for q=0.75 (2016-2023)

	VaR	ES	MES	CoVaR	DeltaCoVaR
E_Score	-0.0004*	-0.0004	-0.0000	-0.0000	-0.0009
	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)
Observations	1191	1191	1191	1191	1191
R^2	0.1375	0.1569	0.1073	0.1362	0.2672

Note:
*Significance level denoted as * for p<0.1, ** for p<0.05, and *** for p<0.01.*
Standard Errors are robust to cluster correlation (cluster)

Taken together, the quantile regression results reveal a consistent and meaningful insight. As measured by E Scores, the beneficial effects of environmental performance are most evident in the lower tail of the systemic risk distribution. Higher E Scores significantly reduce systemic risk exposure at the 10th percentile for most of the systemic risk measures, where the financial system is most vulnerable. This protective effect weakens as we move toward more moderate or higher-risk scenarios. Such findings have important implications for both banking institutions and financial regulators. They suggest that ESG integration, particularly in the environmental dimension, can play a crucial role in promoting sustainability and enhancing systemic resilience. By reinforcing environmental performance, banks may strengthen their capacity to withstand severe financial shocks, thus contributing to broader financial stability.

Conclusion

This thesis explores whether banks that perform better environmentally, as measured by their E Scores, contribute less to systemic risk in the global banking sector in the years following COP21. We applied various financial econometric tools, including panel regressions, Difference-in-Differences analysis, and quantile regressions, supported by several robustness checks on our dataset. The goal was to understand if better environmental practices from banks tend to lower or increase systemic risk spillover.

The initial panel regression results painted a surprising picture, but not in contradiction with some existing research (e.g. Dong, 2025). In the models using E Scores as a continuous variable, higher scores were associated with more negative values of some systemic risk measures like MES, CoVaR, and Delta CoVaR. Since more negative values indicate higher risk, this implied that banks with better E scores might be more systemic, which seemed counterintuitive based on the hypothesis we wanted to answer. However, once we accounted for 2020, a year marked by the COVID-19 crisis and extreme market volatility, and introduced control variables like bank size, profitability, and macroeconomic factors, this relationship became much weaker and, in some cases, disappeared. This showed that the E Score's effect can be easily misinterpreted without adequately accounting for broader shocks and fundamental differences between banks.

To get a better picture, we turned to the DiD model, where instead of looking at the E Score as a number, we checked whether banks had adopted environmental practices after COP21. The results were much more precise, showing that banks that adopted E policies, saw statistically significant improvements in systemic risk measures and became less systemic than their peers. These results suggest that environmental engagement, regardless of intensity, helped banks become more resilient than those that did nothing and were not investing in environmental practices.

We ran another DiD model including an “incentive” being the EU Taxonomy that was adopted in Europe in 2020, a regulation incenting banks to adopt environmental practices, and we saw that indeed E Score increased much more after that period in Europe than in North

America. We observed as a result that the treated group (Europe bank & EU Taxonomy) saw its systemic risk increasing less than banks in America, with statistically significant results, showing the negative relationship between the E Score and the systemic risk and the spillover risk.

We then added a quantile regression analysis to understand better how this relationship plays out across different levels of risk, trying to get another look at the hypothesis after our contrary results of the panel regression and the DiD. The significant findings here showed that higher E Scores were associated with lower risk at the end of the distribution, where banks face the most severe losses. This was especially true for metrics like VaR, MES, and CoVaR at the 10th percentile. As we moved to more moderate scenarios at the 25th percentile or less risky periods at the 75th percentile, the effect of E Scores became weaker or disappeared altogether, telling us that the benefit of environmental performance is most substantial during market stress. In more normal or calm periods, E Scores have a limited impact and significance, this asymmetry is a key insight that shows that environmental efforts help the most when things are rough.

To test the strength of these results, we narrowed our focus to just Europe and North America once again for another panel regression, where banks are more globally connected. The relationship between E Scores and systemic risk remained significant and consistent for most metrics. We also re-ran the models while excluding 2020 to remove the extreme market noise of the pandemic. Again, the results increased, especially for VaR, MES, CoVaR, and Delta CoVaR; these consistent results across different regions and time frames give more confidence that our observations are not just a statistical accident but a real and meaningful trend.

Putting it all together, the picture that emerges is nuanced. On the one hand, the continuous panel regressions suggest that increasing E Scores does not always translate to lower risk, especially when other factors or extreme events are not fully considered. On the other hand, the DiD and quantile regression models show that even minimal environmental engagement helps banks become less systemic, particularly in periods of high financial stress, something that we also observed when excluding banks outside of Europe or North America. This difference in interpretation matters, it reminds us that how we measure and model sustainability influences our conclusions. These results carry practical implications for regulators, investors, and bank managers, encouraging even low levels of environmental commitment in banks can make a difference during crises. At the same time, it is important to keep improving how we assess ESG integration, especially by looking beyond scores and into the quality and consistency of environmental policies.

Ultimately, this thesis shows that environmental performance shapes systemic risk in the global banking sector. Not all effects are significant or straightforward, and context matters a lot. However, overall, more environmentally engaged banks tend to be better prepared when facing a crisis or market downturn, with banks with higher E scores being less systemic and a reduced spillover risk. That makes a strong case for pushing sustainability as a moral choice and a sound strategy to manage financial risk. This opens the door to a broader discussion surrounding recent geopolitical developments and shifts in market sentiment that have challenged the momentum of ESG integration in banking. One pivotal moment came right after Donald Trump's inauguration when the United States announced its withdrawal from the Paris Agreement. This decision marked a symbolic retreat from coordinated international climate efforts and signaled a potential rollback of environmental priorities at the federal level. Just before his inauguration, six major American banks announced their decision to withdraw from the United Nations Net-Zero Banking Alliance (NZBA) to align with the future administration and avoid right-wing criticisms (The Guardian, 2025). It shows that the ESG commitments seem to depend on pressure from investors and politics and that sometimes, the empirical research and conclusions drawn might not be enough to convince the banks. Also, across the Atlantic, sustainable finance has been facing headwinds recently. According to the Centre for Sustainability and Excellence (CSE), European sustainable investment funds recorded historic outflows in 2024, with investors pulling billions of euros from ESG-labeled products. This reversal highlights an erosion of confidence in ESG markets, partly driven by growing criticism over greenwashing, unclear metrics, and regulatory fragmentation. The appeal of ESG investing now appears more fragile and driven by some events, such as geopolitical instability. As banks might turn their back to ESG frameworks, the spillover risk might increase as banks are more exposed to rapidly growing risks.

References

Acharya, V. V., Cooley, T., Richardson, M., & Walter, I. (2010). *Regulating Wall Street: The Dodd-Frank Act and the new architecture of global finance*. Wiley.

Acharya, V. V., Pedersen, L. H., Philippon, T., & Richardson, M. (2017). Measuring systemic risk. *The Review of Financial Studies, 30*(1), 2–47. https://academic.oup.com/rfs/article-abstract/30/1/2/2682977?redirectedFrom=fulltext

Adrian, T., & Brunnermeier, M. K. (2011). *CoVaR* (NBER Working Paper No. 17454). National Bureau of Economic Research. https://www.nber.org/system/files/working_papers/w17454/w17454.pdf

Aevoae, G., Dragotă, I. M., & Ratiu, R. V. (2023). Does ESG performance reduce systemic risk for banks? Evidence from a global sample. *Sustainability, 15*(3), 2422. https://www.mdpi.com/2071-1050/15/3/2422

Allen, F., & Gale, D. (2000). Financial contagion. *Journal of Political Economy, 108*(1), 1–33. https://www.journals.uchicago.edu/doi/10.1086/262109

Bank for International Settlements (BIS). (2021). *Quarterly Review – March 2021*. https://www.bis.org/publ/qtrpdf/r_qt2103.htm

Basel Committee on Banking Supervision. (2011). *Basel III: A global regulatory framework for more resilient banks and banking systems*. Bank for International Settlements. https://www.bis.org/publ/bcbs189.htm

Basel Committee on Banking Supervision. (2013). *Global systemically important banks: Updated assessment methodology and the higher loss absorbency requirement*. https://www.bis.org/publ/bcbs255.htm

Basel Committee on Banking Supervision. (2016). *Minimum capital requirements for market risk*. Bank for International Settlements. https://www.bis.org/bcbs/publ/d352.htm

Battiston, S., Mandel, A., Monasterolo, I., Schütze, F., & Visentin, G. (2017). A climate stress-test of the financial system. *Nature Climate Change, 7*(4), 283–288. https://www.nature.com/articles/nclimate3255

Bloomberg. (2023). *Bloomberg ESG Scores: Overview & FAQ*. Bloomberg L.P.

Brunnermeier, M. K. (2009). Deciphering the liquidity and credit crunch 2007–2008. *Journal of Economic Perspectives, 23*(1), 77–100. https://www.aeaweb.org/articles?id=10.1257/jep.23.1.77

Brunnermeier, M. K., & Oehmke, M. (2013). Bubbles, financial crises, and systemic risk. In G. M. Constantinides, M. Harris, & R. M. Stulz (Eds.), *Handbook of the Economics of Finance* (Vol. 2, pp. 1221–1288). Elsevier. https://doi.org/10.1016/B978-0-44-459406-8.00018-4

Capotă, A., Cristea, M., & Nistor, I. (2022). Investor behavior and ESG: Do sustainable funds exhibit more stable flows? *Journal of Financial Stability, 58*, 100966. https://www.sciencedirect.com/science/article/abs/pii/S1572308921001248?via%3Dihub

Carney, M. (2015). *Breaking the tragedy of the horizon – Climate change and financial stability*. Bank of England. https://www.bankofengland.co.uk/speech/2015/breaking-the-tragedy-of-the-horizon-climate-change-and-financial-stability

Cerqueti, R., Ferraro, G., & Magazzino, C. (2021). ESG compliance and systemic risk: Evidence from investment funds. *Finance Research Letters, 39*, 101661. https://www.sciencedirect.com/science/article/abs/pii/S1544612320302579?via%3Dihub

ChatGPT (OpenAI, 2024) was used to assist with Python coding

Curcio, D., Pancotto, L., & Quagliariello, M. (2024). ESG investment and systemic risk in the European financial system. *European Financial Management*. https://onlinelibrary.wiley.com/doi/10.1111/eufm.12407

Delis, M. D., Iosifidi, M., & Ongena, S. (2019). Bank lending to carbon-intensive firms and climate policy. *Journal of Financial Intermediation, 44*, 100828.

Dong, Y. (2023). ESG portfolios and systemic risk: The unintended consequence of sustainable investing. *Journal of Banking & Finance, 147*, 106269.https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5110773&utm_source

Dong, Y. (2025). Green herding and financial fragility: Evidence from institutional portfolios. *Journal of Financial Stability* (Forthcoming).

European Central Bank. (2022). *Climate stress test report*. https://www.bankingsupervision.europa.eu/ecb/pub/pdf/ssm.ClimateST2022report.en.pdf

European Commission. (2021). *Sustainable finance and the EU taxonomy: Commission takes further steps to channel money towards sustainable activities*. https://ec.europa.eu/commission/presscorner/detail/en/IP_21_1804

Financial Stability Board (FSB). (2010). *Reducing the moral hazard posed by systemically important financial institutions (SIFIs)*. https://www.fsb.org/wp-content/uploads/r_101111a.pdf

Financial Stability Board (FSB). (2011). *Policy measures to address systemically important financial institutions*. https://www.fsb.org/2011/11/r_111104bb/

Galletta, S., & Mazzù, S. (2023). ESG controversies and bank risk: Evidence from European financial institutions. *Journal of International Financial Markets, Institutions and Money, 82*, 101732. https://www.sciencedirect.com/science/article/abs/pii/S1042443122002049?via%3Dihub

Gangi, F., Meles, A., Monferrà, S., & Mustilli, M. (2019). Does corporate social responsibility help the survivorship of banks during financial crises? *Corporate Social Responsibility and Environmental Management, 26*(4), 1394–1405. https://onlinelibrary.wiley.com/doi/10.1002/csr.1751

Gidage, L., Kumar, A., & Bansal, R. (2024). ESG and systemic risk: Evidence from emerging markets. *International Review of Economics & Finance*. https://doi.org/10.1016/j.iref.2024.01.007

Gorton, G. (2010). *Slapped by the invisible hand: The panic of 2007*. Oxford University Press.

Grammarly. (n.d.). *Grammarly: Free AI writing assistance*. https://www.grammarly.com/ (used for grammar and wording)

Hull, J. C. (2018). *Risk management and financial institutions* (5th ed.). Wiley.

Ilhan, E., Sautner, Z., & Vilkov, G. (2021). Carbon tail risk. *The Review of Financial Studies, 34*(3), 1540–1571. https://academic.oup.com/rfs/article-abstract/34/3/1540/5864038?redirectedFrom=fulltext

International Monetary Fund (IMF). (2020). *Global Financial Stability Report*. https://www.imf.org/en/Publications/GFSR/Issues/2020/10/13/global-financial-stability-report-october-2020

Jorion, P. (2007). *Value at Risk: The new benchmark for managing financial risk* (3rd ed.). McGraw-Hill.

Koenker, R., & Bassett, G. (1978). Regression quantiles. *Econometrica, 46*(1), 33–50. https://www.jstor.org/stable/1913643?origin=crossref

Krueger, P., Sautner, Z., & Starks, L. T. (2020). The importance of climate risks for institutional investors. *The Review of Financial Studies, 33*(3), 1067–1111. https://academic.oup.com/rfs/article-abstract/33/3/1067/5735302?redirectedFrom=fulltext

Maddala, G. S., & Kim, I.-M. (1998). *Unit roots, cointegration, and structural change*. Cambridge University Press.

Mishkin, F. S. (2011). Over the cliff: From the subprime to the global financial crisis. *Journal of Economic Perspectives, 25*(1), 49–70. https://www.aeaweb.org/articles?id=10.1257/jep.25.1.49

Monasterolo, I., & Raberto, M. (2019). The EIRIN flow-of-funds behavioural model of green fiscal policies and green sovereign bonds. *Ecological Economics, 144*, 228–243. https://www.sciencedirect.com/science/article/abs/pii/S0921800916306000?via%3Dihub

Network for Greening the Financial System (NGFS). (2020). *Climate scenarios for central banks and supervisors*. https://www.ngfs.net/en/climate-scenarios-central-banks-and-supervisors

Ng, S., & Perron, P. (2001). Lag length selection and the construction of unit root tests with good size and power. *Econometrica, 69*(6), 1519–1554. https://onlinelibrary.wiley.com/doi/abs/10.1111/1468-0262.00256

OECD. (2021). *Climate finance provided and mobilised by developed countries: Aggregate trends updated with 2019 data*. https://www.oecd.org/en/publications/climate-finance-provided-and-mobilised-by-developed-countries-aggregate-trends-updated-with-2019-data_03590fb7-en.html

Perron, P. (1989). The great crash, the oil price shock, and the unit root hypothesis. *Econometrica, 57*(6), 1361–1401. https://www.jstor.org/stable/1913712?origin=crossref

Refinitiv. (2023). *ESG scoring methodology guide*. https://www.refinitiv.com/en/sustainable-finance/esg-scores

Rockafellar, R. T., & Uryasev, S. (2002). Conditional value-at-risk for general loss distributions. *Journal of Banking & Finance, 26*(7), 1443–1471. https://www.sciencedirect.com/science/article/abs/pii/S0378426602002716?via%3Dihub

Sanford, D., & van Norden, S. (2023). *Risk modeling and financial stability* [Lecture 2]. HEC Montréal.

Shakil, M. H. (2021). ESG controversies and firm risk: Evidence from the energy sector. *Energy Economics, 94*, 105066. https://www.sciencedirect.com/science/article/abs/pii/S0140988320304060?via%3Dihub

Stock, J. H., & Watson, M. W. (2019). *Introduction to econometrics* (4th ed.). Pearson Education.

Task Force on Climate-related Financial Disclosures (TCFD). (2017). *Final report: Recommendations of the Task Force on Climate-related Financial Disclosures*. https://www.fsb-tcfd.org/publications/final-recommendations-report/

United Nations Environment Programme Finance Initiative (UNEP FI). (2021). *Net-Zero Banking Alliance*. https://www.unepfi.org/net-zero-banking/

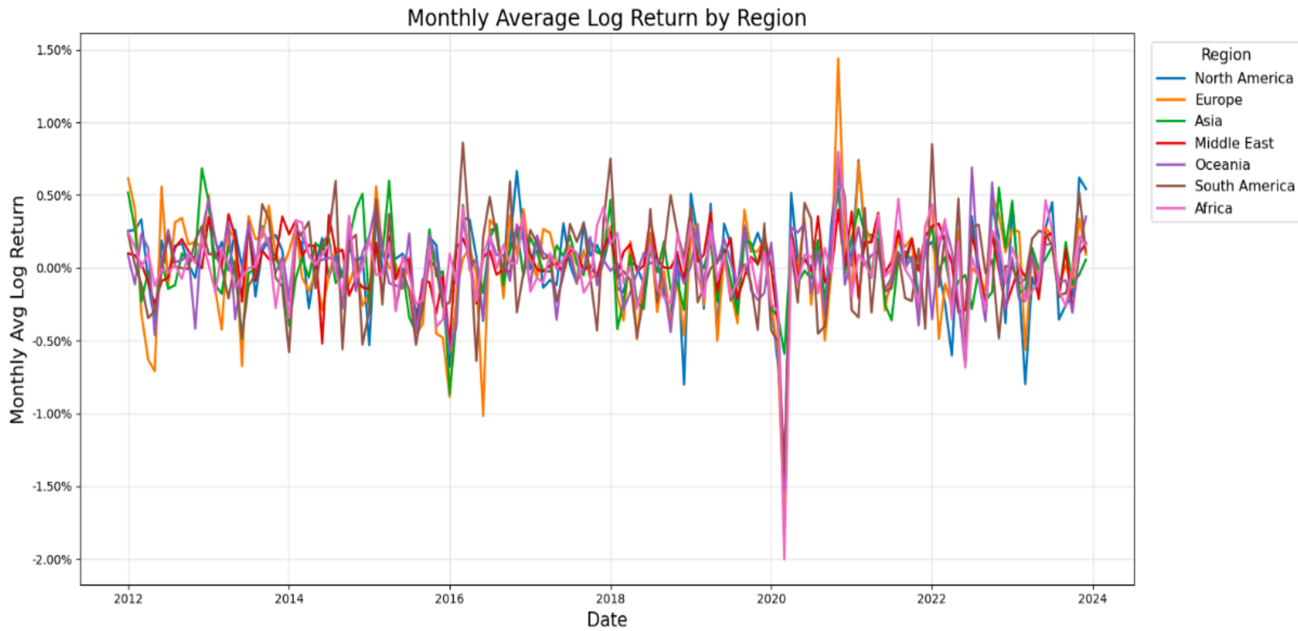
United Nations Framework Convention on Climate Change (UNFCCC). (1998). *Kyoto Protocol to the United Nations Framework Convention on Climate Change*. https://unfccc.int/kyoto_protocol

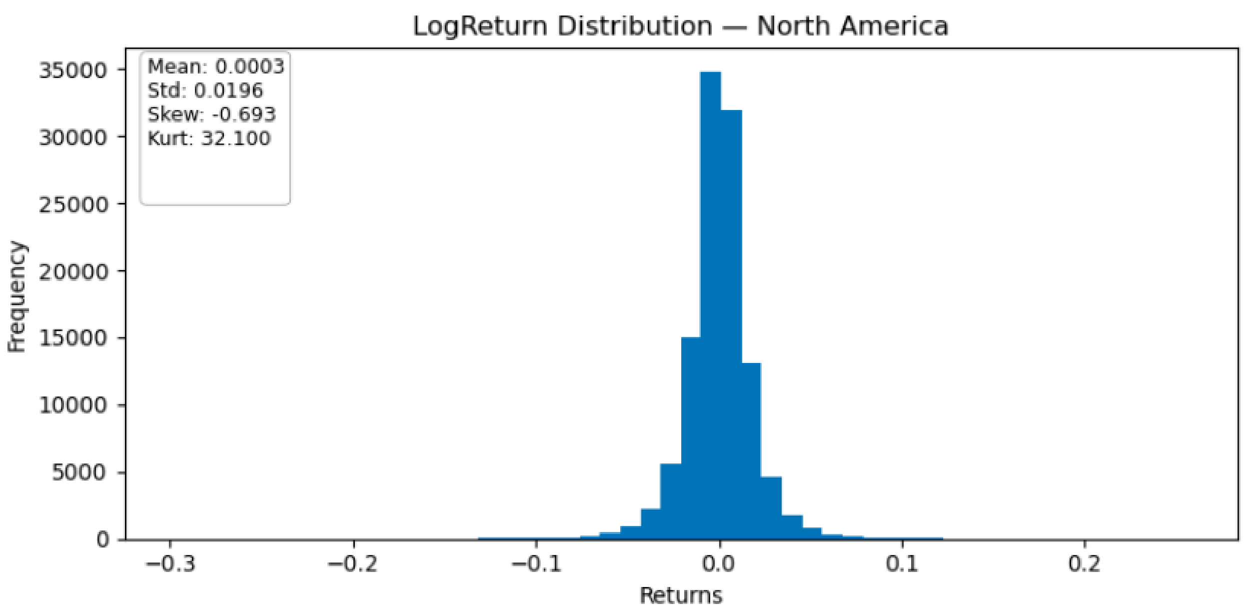
United Nations Framework Convention on Climate Change (UNFCCC). (2015). *The Paris Agreement*. https://unfccc.int/process-and-meetings/the-paris-agreement/the-paris-agreement

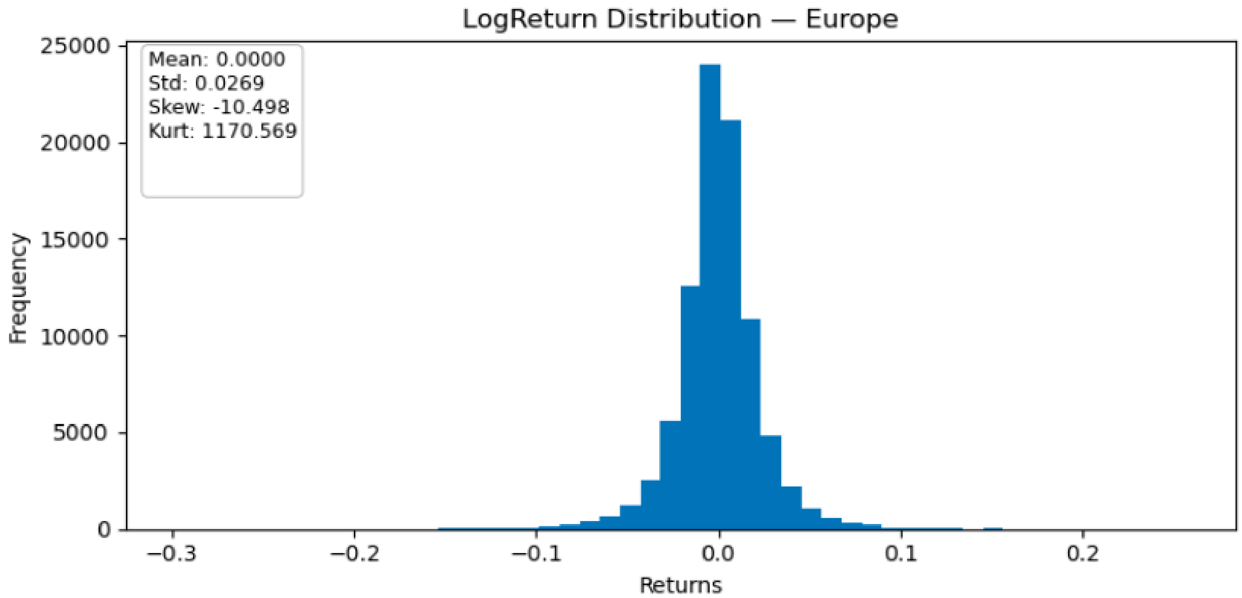
White, H., Kim, T. H., & Mammen, E. (2015). Testing conditional independence via quantile regression. *Journal of Econometrics, 185*(2), 259–273. https://www.sciencedirect.com/science/article/abs/pii/S0304407614001742?via%3Dihub

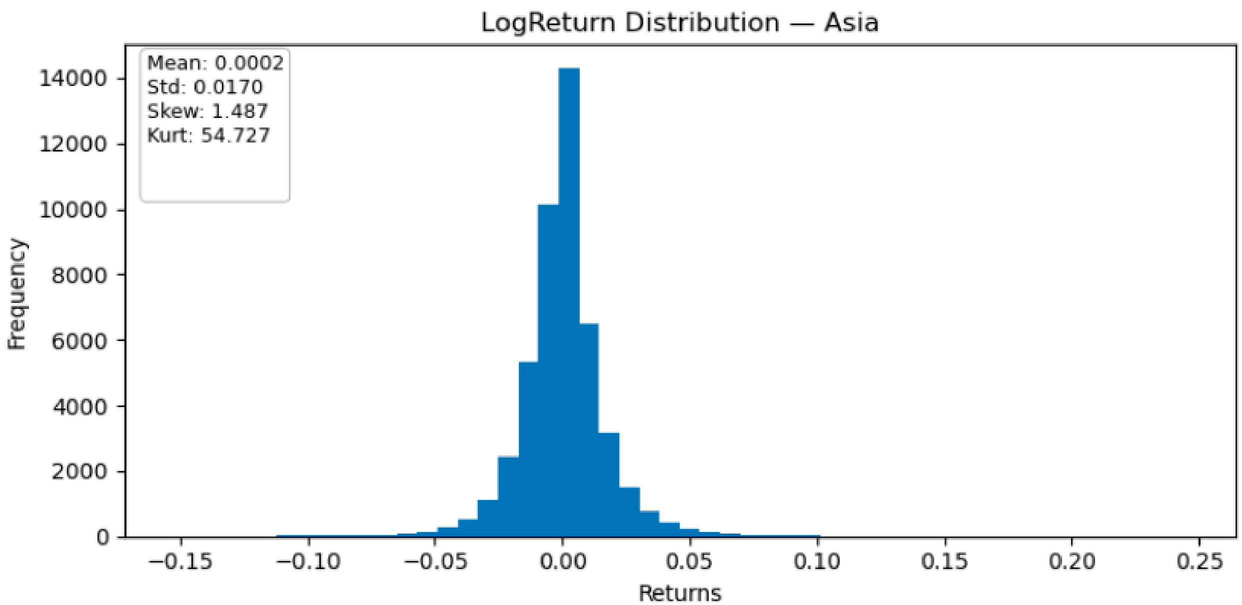
Wong, W., & Zhang, X. (2022). Media coverage of ESG incidents and firm value. *Journal of Corporate Finance, 72*, 102146. https://www.sciencedirect.com/science/article/abs/pii/S0929119921002686?via%3Dihub

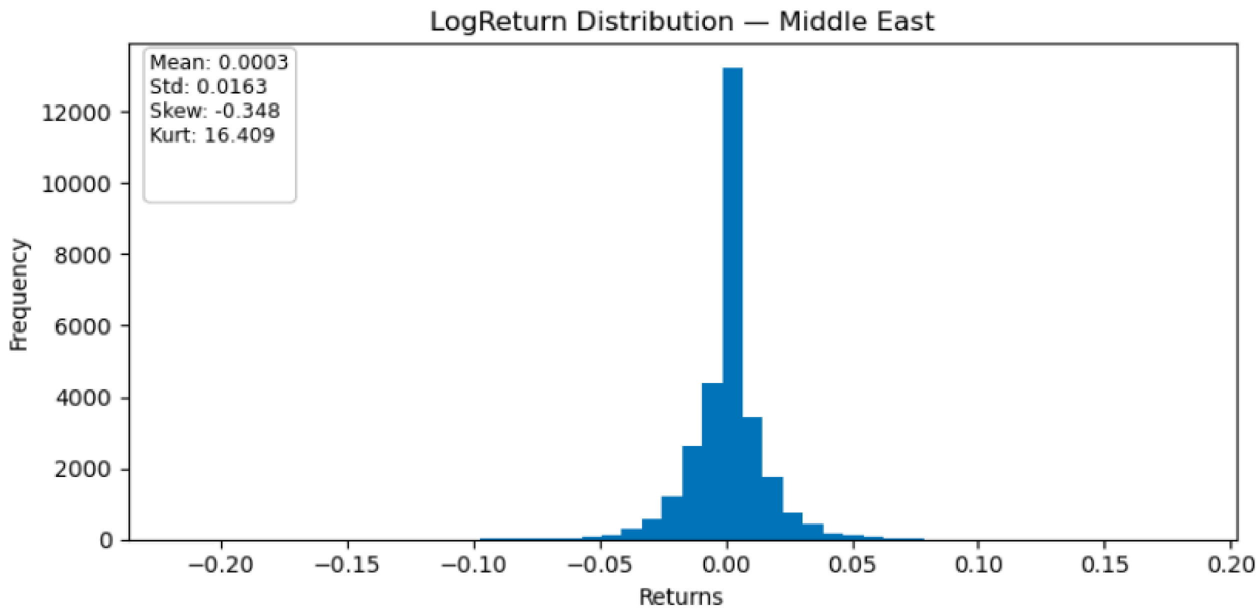
Annex 1

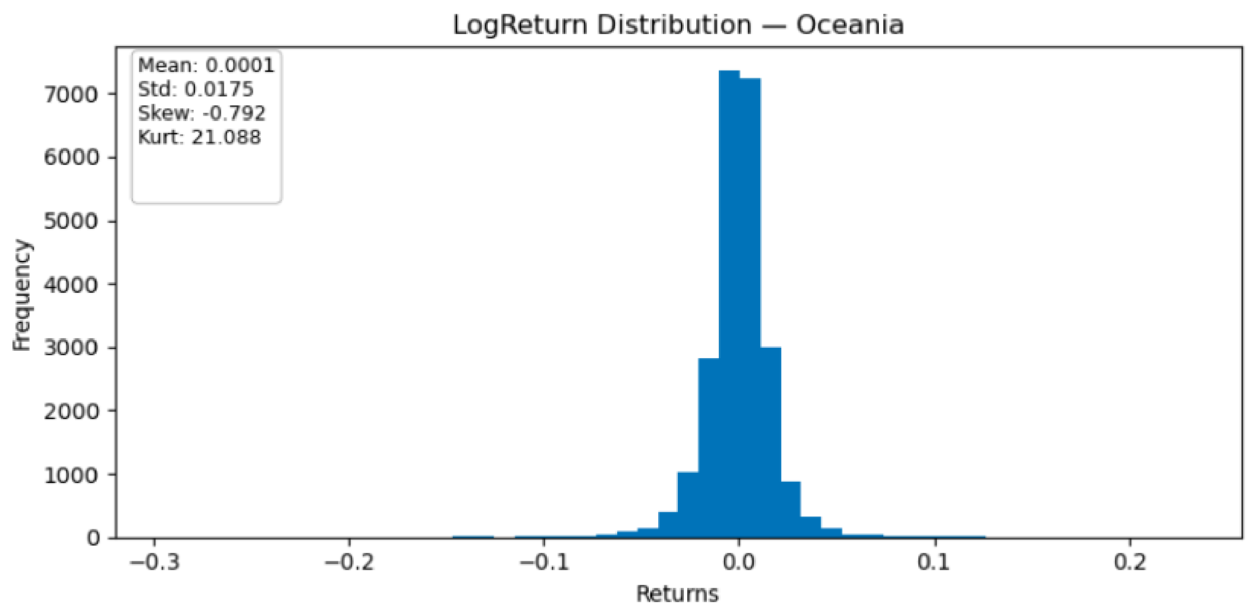

Monthly Average Log Return by Region
Monthly Avg Log Return
1.50%
1.00%
0.50%
0.00%
-0.50%
-1.00%
-1.50%
-2.00%
2012
2014
2016
2018
2020
2022
2024
Date
Region
North America
Europe
Asia
Middle East
Oceania
South America
Africa

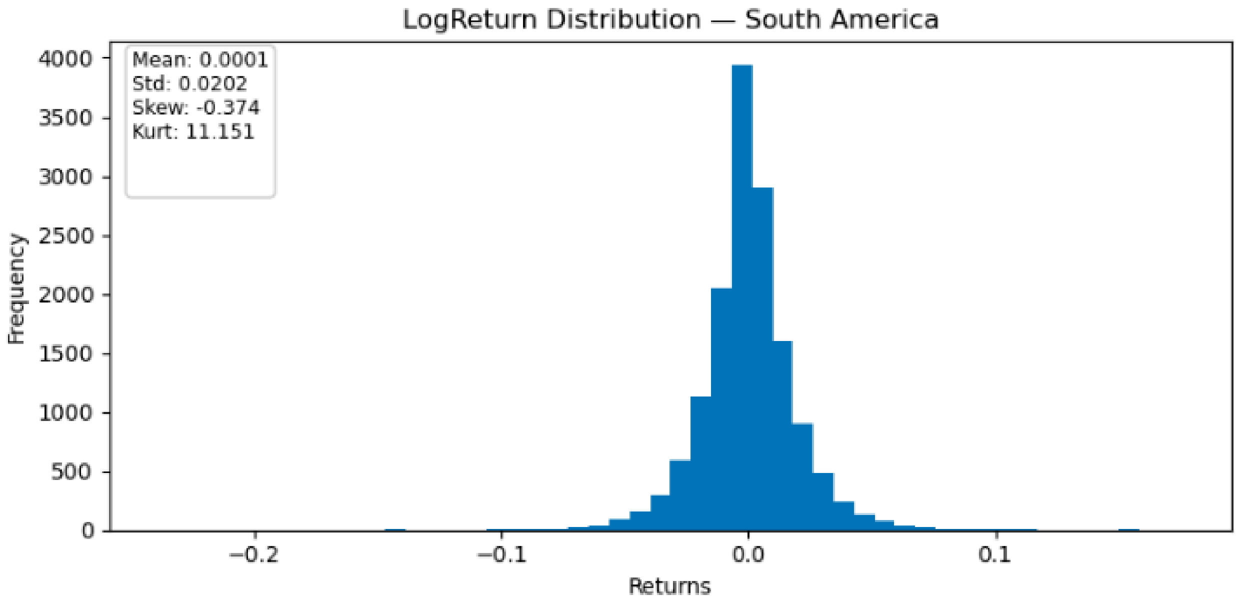

LogReturn Distribution — North America
Mean: 0.0003
Std: 0.0196
Skew: -0.693
Kurt: 32.100
Frequency
35000
30000
25000
20000
15000
10000
5000
0
−0.3
−0.2
−0.1
0.0
0.1
0.2
Returns

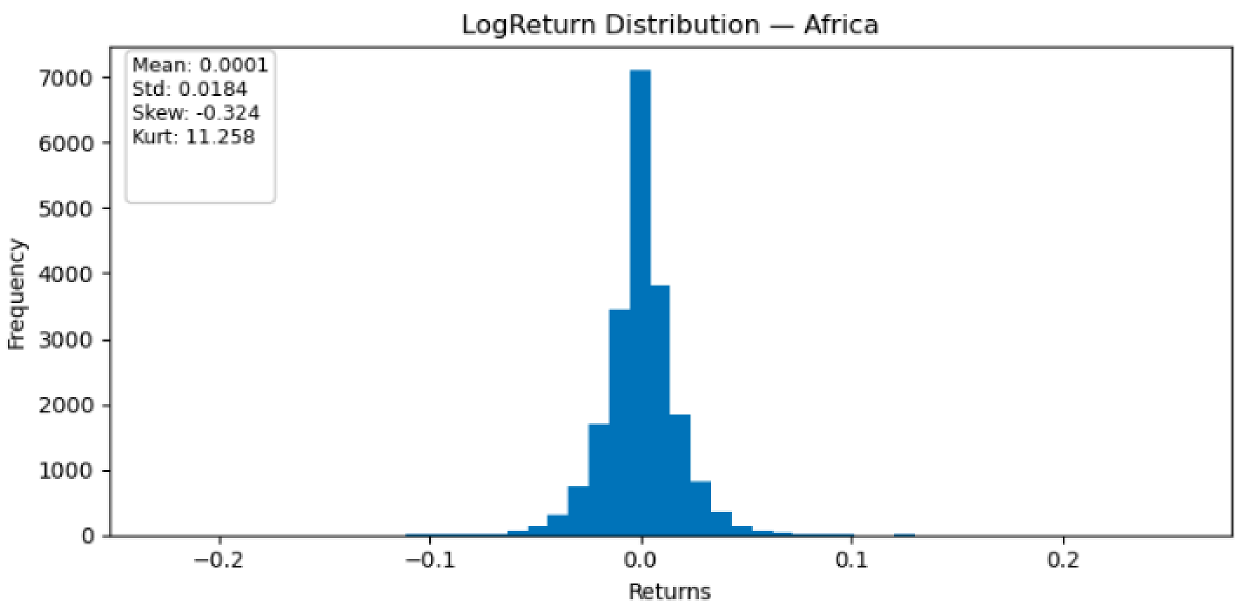

LogReturn Distribution — Europe
Mean: 0.0000
Std: 0.0269
Skew: -10.498
Kurt: 1170.569
Frequency
Returns

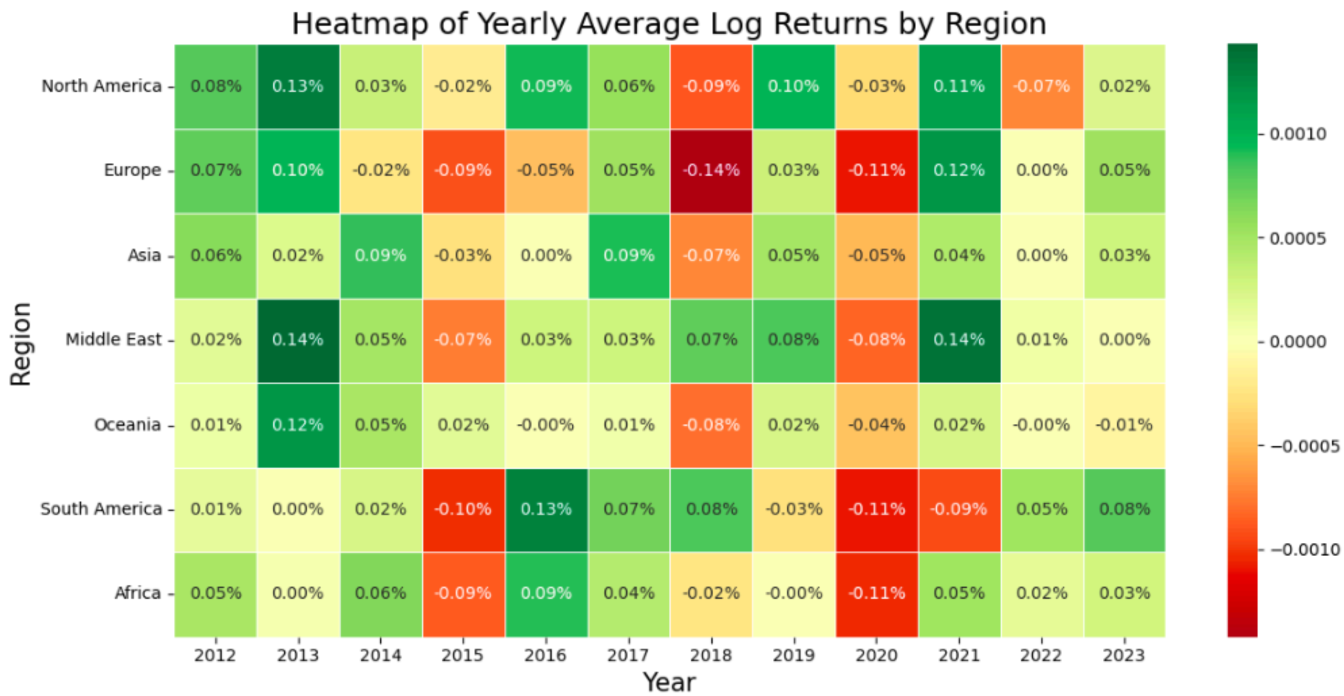

LogReturn Distribution — Asia
Mean: 0.0002
Std: 0.0170
Skew: 1.487
Kurt: 54.727
Frequency
Returns


LogReturn Distribution — Middle East
Mean: 0.0003
Std: 0.0163
Skew: -0.348
Kurt: 16.409
Frequency
0
2000
4000
6000
8000
10000
12000
−0.20
−0.15
−0.10
−0.05
0.00
0.05
0.10
0.15
0.20
Returns


LogReturn Distribution — Oceania
Mean: 0.0001
Std: 0.0175
Skew: -0.792
Kurt: 21.088
Frequency
0
1000
2000
3000
4000
5000
6000
7000
−0.3
−0.2
−0.1
0.0
0.1
0.2
Returns


LogReturn Distribution — South America
Mean: 0.0001
Std: 0.0202
Skew: -0.374
Kurt: 11.151
Frequency
Returns
0
500
1000
1500
2000
2500
3000
3500
4000
−0.2
−0.1
0.0
0.1


LogReturn Distribution — Africa
Mean: 0.0001
Std: 0.0184
Skew: -0.324
Kurt: 11.258
Frequency
Returns
0
1000
2000
3000
4000
5000
6000
7000
−0.2
−0.1
0.0
0.1
0.2



***Source** : Self Plotted on Python based on daily log returns for banks on Bloomberg Terminals*

Annex 2

Summary Statistics of Annual 5% VaR by Region (2012–2023)

	Region	VaR Mean	VaR Std	VaR Min	VaR Max
0	North America	-0.0271	0.0120	-0.0803	-0.0095
1	Europe	-0.0342	0.0173	-0.2312	-0.0119
2	Asia	-0.0236	0.0084	-0.0568	-0.0066
3	Middle East	-0.0226	0.0085	-0.0492	-0.0065
4	Oceania	-0.0246	0.0093	-0.0539	-0.0112
5	South America	-0.0281	0.0103	-0.0513	-0.0121
6	Africa	-0.0266	0.0103	-0.0622	-0.0091

Summary Statistics of Annual 5% Expected Shortfall (ES) by Region (2012–2023)

	Region	ES Mean	ES Std	ES Min	ES Max
0	North America	-0.0401	0.0209	-0.1803	-0.0139
1	Europe	-0.0514	0.0306	-0.3322	-0.0184
2	Asia	-0.0344	0.0120	-0.0767	-0.0128
3	Middle East	-0.0360	0.0149	-0.0822	-0.0124
4	Oceania	-0.0376	0.0175	-0.0904	-0.0157
5	South America	-0.0414	0.0176	-0.1033	-0.0169
6	Africa	-0.0382	0.0174	-0.1200	-0.0135

Summary Statistics of Annual MES by Region (2012–2023)

	Region	MES Mean	MES Std	MES Min	MES Max
0	North America	-0.0255	0.0183	-0.1170	0.0122
1	Europe	-0.0190	0.0158	-0.0902	0.0124
2	Asia	-0.0053	0.0072	-0.0331	0.0062
3	Middle East	-0.0059	0.0103	-0.0502	0.0138
4	Oceania	-0.0063	0.0096	-0.0430	0.0107
5	South America	-0.0142	0.0144	-0.0739	0.0024
6	Africa	-0.0103	0.0150	-0.0930	0.0058

Summary Statistics of Annual CoVaR by Region (2012–2023)

	Region	CoVaR Mean	CoVaR Std	CoVaR Min	CoVaR Max
0	North America	-0.0536	0.0391	-0.3119	-0.0104
1	Europe	-0.0575	0.0436	-0.3456	-0.0076
2	Asia	-0.0301	0.0211	-0.1051	-0.0029
3	Middle East	-0.0359	0.0340	-0.1663	-0.0033
4	Oceania	-0.0345	0.0262	-0.1318	-0.0083
5	South America	-0.0414	0.0358	-0.1973	-0.0091
6	Africa	-0.0365	0.0331	-0.2181	-0.0050

Summary Statistics of Annual ΔCoVaR by Region (2012–2023)

	Region	ΔCoVaR Mean	ΔCoVaR Std	ΔCoVaR Min	ΔCoVaR Max
0	North America	-0.0245	0.0243	-0.1069	0.0063
1	Europe	-0.0196	0.0247	-0.1036	0.0139
2	Asia	-0.0134	0.0226	-0.0969	0.0159
3	Middle East	-0.0137	0.0254	-0.1031	0.0136
4	Oceania	-0.0171	0.0267	-0.1015	0.0066
5	South America	-0.0161	0.0265	-0.1015	0.0086
6	Africa	-0.0140	0.0271	-0.1015	0.0123