Do ESG Ratings Drive Financial Performance? A Systematic Analysis of Trends and Challenges

Amelie Heinelt[1][a], Dominic Strube[1][b] and Christian Daase[2][c]

[1]Hochschule Wismar, University of Applied Sciences, Technology, Business and Design, Wismar, Germany
[2]Institute of Technical and Business Information Systems, Otto-von-Guerick University, Magdeburg, Germany

Abstract: This study examines the relationship between ESG (Environmental, Social, and Governance) ratings and financial performance through a systematic analysis of studies published between 2019 and 2024. The findings reveal that a significant correlation between ESG ratings and financial performance was only demonstrated in a portion of the studies. Regression-based models were the most frequently used methods, followed by panel data and time series analyses. However, no clear statistical relationship between the choice of methodology and the results could be established. Variations in findings are attributed to differences in ESG rating methodologies, data sources, and external factors such as macroeconomic conditions and market volatility. While ESG investments may involve short-term costs, they can contribute to long-term stability. The study highlights the need for standardized ESG ratings and consistent analytical approaches to enable more reliable conclusions.

1 INTRODUCTION

ESG ratings (Environmental, Social, and Governance) assess a company's performance in these three key areas, providing financial market participants with essential non-financial information about the sustainability of companies. Leading providers include MSCI, Sustainalytics, Thomson Reuters (formerly Asset4), Bloomberg, and Vigeo Eiris (now part of Moody's). Sustainability ratings are gaining increasing importance, serving as a strategic decision-making tool for investors and managers as well as a guide for financial investments worth trillions of dollars, such as sustainable funds (Dorfleitner et al., 2014; Hughes et al., 2021; Nazarova and Лаврова, 2022). They are used either to channel financial resources into sustainable projects out of conviction or with the expectation of achieving higher returns compared to conventional investments. According to Morningstar, global sustainable fund assets amounted to 7,659 billion USD in the second quarter of 2024. Europe dominates this market significantly with a share of 73.2% and managed assets of 5,609 billion USD. The USA follows with only 8.0%. Figure 1

shows the development of global sustainable fund assets since the beginning of 2020. These initially grew steadily but experienced a sharp decline during the COVID-19 pandemic. Since then, they have continuously recovered and continue to grow, albeit at a slower pace. Europe also leads in net fund inflows, contributing USD 10.3 billion in Q3 2024, while the United States experienced continuous net outflows throughout 2024 (Morningstar, 2024). These funds are fundamentally based on the evaluations of ESG rating agencies.

Over recent years, numerous studies have examined the question of whether ESG ratings correlate with financial performance. This article aims to investigate whether the correlation between ESG ratings and financial performance has evolved over time. In Europe, stricter regulations such as the EU Taxonomy, the Sustainable Finance Disclosure Regulation (SFDR), and the Corporate Sustainability Reporting Directive (CSRD) may have enhanced the transparency and reliability of ESG data. Additionally, ESG ratings themselves have likely improved through the integration of new data sources,

[a] https://orcid.org/0009-0008-3050-1486
[b] https://orcid.org/0000-0003-3017-5189
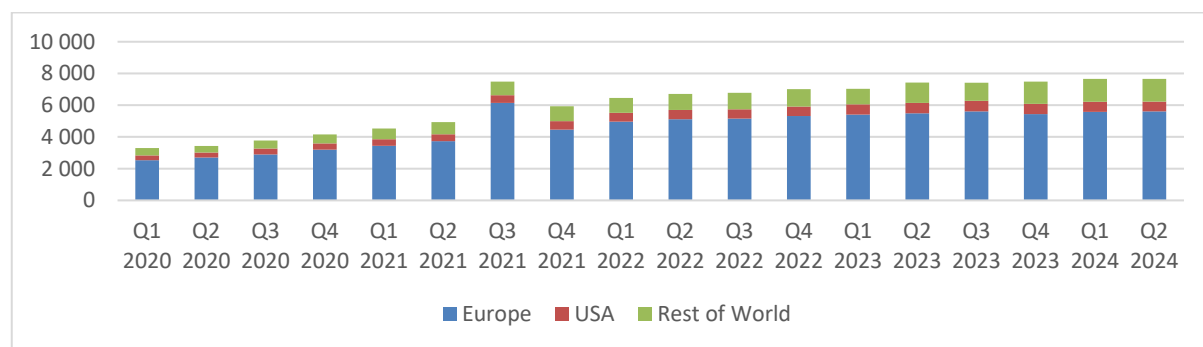[c] https://orcid.org/0000-0003-4662-7055

Figure 1: Development of Global Sustainable Fund Assets in USD Billion by Region (Q1 2020 – Q2 2024).

advances in methodologies, and technologies like artificial intelligence.

The study also examines whether specific factors, such as the choice of ESG rating provider, geographic focus, or analytical methods, affect study outcomes. Europe, recognized as a leader in sustainable finance, may exhibit distinct results due to its strong regulatory framework and investor demand, which can drive returns. Analytical and methodological choices (e.g., linear regression, event studies, or portfolio construction) are also critical factors influencing outcomes and are therefore included in this analysis.

The following research questions arise:

RQ 1: Have the correlation results between ESG ratings and financial performance changed since 2019?

RQ 2: How do different ESG rating providers, methodological choices, and geographic differences influence the financial performance of ESG investments?

A detailed data table of the evaluated studies is available upon request, as the full analysis exceeds the article's space limitations.

2 METHODOLOGY

The purpose of this paper is to analyse the pre-existing literature on the ESG rating and its impact on the financial performance of companies.
The focus was on proven correlation by studies that have already been conducted as well as how the different ESG rating providers, methodological choices and geographic differences impact the financial performance. These were then further analysed in this article. This analysis was aimed at identifying any themes and recurring trends in these

studies to understand if these have changed throughout the years. Furthermore, this can be used as a basis for further research and identify gaps in the existing literature.

2.1 Source Selection

For the literature search a systematic approach was employed. This approach can be used to synthesise scientific evidence and answer one or more research questions on a prior established topic. It is supposed to further academic research by building on already existing literature and their results. This approach makes it possible to use empirical methods combined with a traditional literature review (Lame, 2019).

The search was conducted through the academic database Scopus. This database was chosen as it is one of the largest academic research abstract-databases and should therefore provide the greatest selection of literature on the topic. The search strings that were used was "(esg OR "environmental, social, governance" OR "sustainability" AND "Rating") AND ("financial performance" OR "stock performance" OR "stock returns") AND (correlation OR relationship)". The main filter that was used was the timeframe. The search was condensed down to articles that were published from the year 2019 to 2024. This specific time frame was chosen as there was a visible increase in studies published on this topic since 2019.

2.2 Search and Screening Process

From all the articles that could be found through the search described above, only the 98 most cited articled were used for analysis. This decision was made to ensure a wide selection of literature that represent the foundational research in the topic, as well as identifies the core concepts and trends. For this analysis mainly the methodology section and

discussion as well as results were taken into consideration.

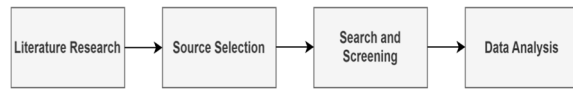The graphic below is visualizing the research process for this paper.

Figure 2: Research process.

2.3 Data Analysis

An analysis of the methodology, discussion and results section was performed. There were predetermined themes that were used in the analysis of the data. In this case the themes included:

- The timeframe of the study
- The data sources used for ESG ratings and financial data
- Geographical and sectoral factors
- The results of the study
- If there was a correlation found or not

An analysis of the data collected from the 98 articles was used to identify the themes and trends, as well as see any changes in them throughout the years. The data was coded in an excel document to visualize the trends that could be seen in the different themes used to analyse.

There was also an analysis of the different statistical methods used in the original studies taken into consideration for our results.

Finally, a chi-square test was used to investigate the connection between the methodological approaches and the correlation results. This test is also known as the Pearsons chi-square and is most used to test associations between to categorical variables such as the existence of a correlation and the statistical method.

The formula to calculate the chi-square test is also known as (Onchiri, 2013):

$$X^2 = N \cdot \frac{(AD - BC)^2}{(A + C)(A + D)(A + B)(A + D)}$$

2.4 Limitations

Just as any method there are a few limitations that must be considered. As the studies that were used were limited to articles written in English, there might be relevant findings that were excluded due to studies in other languages were not investigated for this analysis.

Another limitation is the fact that only peer revied journal articles were used, with more in-depth research and analysis of industry or company reports, there might have been more, or other insights that could have been seen.

Lastly, the fact that there is a large variation in data sources within the articles that have been used for analysis, it is possible to assume that there are potential inconsistencies which can result in difficulties when directly comparing them to one another.

2.5 Justification

There was a visible increase in studies on this topic, starting in 2019 and very few before then. This can indicate that there is either a relevance in this topic or there have been significant findings in studies which led to an increase in further studies and research on the topic. The analysis of a wide variety of articles with different focus points made sure to ensure diverse perspectives and the generalizability of the findings in this analysis. Furthermore, the thematic analysis provides a structured approach to combine findings from quantitative and qualitative studies.

Figure 3: Frequency of Different Methodological Research Approaches.

Methodological Research approach	Frequency
Regression-Based Technique	41
Panel Data and Time Series Methods	14
Machine Learning and Predictive Analytics	5
Portfolio and Risk Analysis	3
Factor Analysis and Causal Inference	2
Multivariate and Descriptive Methods	11
Event and Impact Studies	2
Other	15

3 RESULTS

In total, 98 studies were analysed, with approximately 39% demonstrating a significant correlation between ESG ratings and financial performance. The most commonly used ESG ratings include Bloomberg, MSCI, Thomson/Refinitiv, Compustat, and Sustainalytics. However, the selection of ratings

varied widely across studies, with multiple data providers often used in combination. This variability makes it difficult to establish a clear pattern or attribute results to any specific rating source. The use of different providers reflects the diversity in data availability and methodological preferences among researchers.

The analysis of 98 studies reveals that regression-based techniques, with 41 mentions, are the most frequently employed method to determine the relationship between ESG ratings and financial performance. This popularity may stem from their suitability for quantifying relationships between independent variables (e.g., ESG ratings) and dependent variables (e.g., financial performance), as well as their ability to control for confounding factors. Far behind are panel data and time series methods (14), which are often applied to account for both temporal trends and firm-specific differences, making them particularly useful for longitudinal studies.

Multivariate and descriptive methods were employed 11 times, serving as tools for exploring data structures and identifying patterns. Less commonly used are specialized approaches such as machine learning and predictive analytics (5), portfolio and risk analysis (3), as well as factor analysis and causal inference and event and impact studies (each with 2 mentions). The "Other" category (15) reflects alternative methods.

To investigate whether there is a connection between the methodological approach and the correlation results, a chi-square Test was performed. The test yielded a Chi-Square value of 0.59 and a p-value of 0.999, suggesting no statistically significant relationship between the methods used and the likelihood of observing a correlation. This indicates that the choice of methodology does not seem to influence whether a study identifies a significant relationship between ESG ratings and financial performance.

An analysis of the correlation between the study period and correlation results also showed no significant findings. Interestingly, a slight negative correlation of -0.216 was observed, suggesting that more recent studies tend to report fewer correlations. However, the p-value of 0.219 is well above the significance threshold of 0.05, meaning no statistically significant relationship can be established. This result might reflect evolving methodologies or changing perceptions of the relationship over time but requires further exploration.

Many of the studies with a demonstrated correlation focus on global markets or multi-sectoral

analyses, such as oil and gas or real estate. These broad approaches aim to capture general trends and cross-industry insights. Few studies specifically examine individual countries or regions, such as China or BRICS nations, making it challenging to establish a clear geographic preference or draw region-specific conclusions. Similarly, no clear relationship was observed between the choice of ESG ratings and the results. The data show a wide distribution of sources, with no single ESG data provider dominating the studies with demonstrated correlations. This diversity underscores the complexity of the topic and the importance of considering multiple perspectives in ESG research.

4 DISCUSSION

Our results show that there is no statistically significant relationship between the study period, the choice of methodology, and the correlation results of the studies. Due to the challenges, no statistical relationship can be mathematically proven in relation to region or rating provider in our case. However, it appears that neither of these aspects serves as a main driver for positive results.

The reasons for this are manifold and are most likely rooted in the design of the rating methodologies themselves. Numerous studies demonstrate that the breadth and diversity of ESG factors, the subjectivity of their evaluation, and differing assessment methods result in vastly divergent ratings for the same company. This means that ESG ratings exhibit high inconsistency due to low correlations among them, which can be attributed to the lack of standardized methods for measuring ESG performance (Berg et al., 2019; Chatterji et al., 2016; Dimson et al., 2020; Zumente and Lāce, 2021). The clear relationship between a high ESG rating and a higher level of sustainability remains ambiguous. Some studies suggest that the mere volume of available data positively correlates with ESG ratings, indicating that insufficient sustainability data could lead to a downgraded rating (Drempetic et al., 2020). Although approximately 39% of the analysed articles indicate a positive correlation, the lack of standardized criteria makes it difficult to compare results across studies or draw reliable conclusions. Furthermore, a correlation does not necessarily imply causation, which is challenging to establish given the mentioned limitation. Nonetheless, a correlation can be used to detect existing relationships between variables and can be of use to guide further studies in the are of study. It is also valuable for predictions, when a clear

underlying cause cannot be identified. Lastly, these correlations can be used to generate further hypothesis which can be proven by causal research. Only through the establishment of standardized definitions of sustainability and uniform measurement methods will it be possible in the future to conduct more reliable and robust investigations.

Another significant factor is that many studies use capital market indicators as their dependent variable. In particular, the stock market is influenced by a multitude of complex factors, such as macroeconomic developments, geopolitical events, and speculative behaviour. Crises like the COVID-19 pandemic have significantly increased the volatility of capital markets in industrialized nations (Baek et al., 2020; Ozkan, 2021). Similarly, the Russia-Ukraine conflict had a substantial impact on stock returns and market volatility, leading to high inflation and rising interest rates (Ahmed et al., 2023; Wu et al., 2023). Against this backdrop, the effect of ESG ratings may be overshadowed by these broader influences, making both temporal and geographic comparisons difficult, as macroeconomic and geopolitical factors vary significantly between countries.

Furthermore, it remains questionable whether sustainability leads to a short-term improvement in financial performance. Investments in social and governance aspects may result in companies incurring higher short-term costs, for instance, through stricter compliance regulations, improved working conditions, or more comprehensive reporting. While these measures contribute to long-term stability and the company's reputation, they can reduce returns in the initial phase. Thus, sustainability may not always provide immediate financial benefits but rather represents a strategic decision aimed at long-term stability and corporate responsibility.

5 CONLUSIONS

Ultimately this study demonstrates that there is no statistical relationship between the study period, methodological choices, or the geographical focus, as initially suspected. The variety in different ESG rating providers and methodological approaches added to the complexity of this analysis. This issue highlights the need for more standardized ESG rating to be able to definitively draw conclusions from these ratings and their impact on different factors such as the financial performance of a company. These findings further suggest that the variability in ESG ratings can be a contributing factor to the inconsistent results across studies. This is further driven by the

differing methodologies, as well as subject evaluations. Approximately 39% of the studies that were observed showed a positive correlation between the ESG ratings and financial performances. It is crucial to note, that correlation does not imply causation. Variability in ESG ratings, mainly driven by the different methodologies used as well as differences in the factors considered, have a significant impact on the inconsistent results across the different studies. There were approximately 39% of studies that showed a correlation between the ESG rating and financial performance, it must be mentioned, that a correlation does not imply an automatic causation between these factors. There is a variety of external factors which can manipulate the impact of ESG ratings on financial performance. Some examples for these factors can be macroeconomic conditions as well as the market volatility. There are definite higher short-term costs due to investments into the ESG aspects which have to be considered. These can lead to more ling-term stability however there is no certainty, that it causes financial improvements. There should be a clear understanding of the relationship between the ESG ratings, and the financial performance achieved. This can be supported by standardized ratings and methodologies as well as analysis which account for internal as well as external factors.

REFERENCES

Ahmed, S., Hasan, M. M., & Kamal, M. R. (2023). Russia–Ukraine crisis: The effects on the European stock market. *European Financial Management*, *29*(4), 1078–1118.

Baek, S., Mohanty, S. K., & Glambosky, M. (2020). Covid-19 and stock market volatility: An industry level analysis. *Finance Research Letters*, *37*, 101748.

Berg, F., K&ouml, lbel, J., & Rigobon, R. (2019). Aggregate Confusion: The Divergence of ESG Ratings. *SSRN Electronic Journal.* Advance online publication.

Chatterji, A. K., Durand, R., Levine, D. I., & Touboul, S. (2016). Do ratings of firms converge? Implications for managers, investors and strategy researchers. *Strategic Management Journal*, *37*(8), 1597–1614.

Dimson, E., Marsh, P., & Staunton, M. (2020). Divergent ESG Ratings. *The Journal of Portfolio Management*, *47*(1), 75–87.

Dorfleitner, G., Halbritter, G., & Nguyen, M. (2014). Measuring the Level and Risk of Corporate Responsibility - An Empirical Comparison of Different ESG Rating Approaches. *SSRN Electronic Journal.* Advance online publication.

Drempetic, S., Klein, C., & Zwergel, B. (2020). The Influence of Firm Size on the ESG Score: Corporate

Sustainability Ratings Under Review. *Journal of Business Ethics*, *167*(2), 333–360.

Hughes, A., Urban, M. A., & Wójcik, D. (2021). Alternative ESG Ratings: How Technological Innovation Is Reshaping Sustainable Investment. *Sustainability*, *13*(6), 3551.

Lame, G. (2019). Systematic Literature Reviews: An Introduction. *Proceedings of the Design Society: International Conference on Engineering Design*, *1*(1), 1633–1642.

Morningstar. (2024). *Global Sustainable Fund Flows: Q3 2024 in Review*. https://www.morningstar.com/lp/global-esg-flows

Nazarova, V., & Лаврова, В. (2022). Do ESG Factors Influence Investment Attractiveness of the Public Companies? *Journal of Corporate Finance Research / Корпоративные Финансы | ISSN: 2073-0438*, *16*(1), 38–64.

Onchiri, S. (2013). Conceptual model on application of chi-square test in education and social sciences. *Educational Research and Reviews*, *2013*, 1231–1240.

Ozkan, O. (2021). Impact of COVID-19 on stock market efficiency: Evidence from developed countries. *Research in International Business and Finance*, *58*, 101445.

Wu, F., Zhan, X., Zhou, J., & Wang, M. (2023). Stock market volatility and Russia–Ukraine conflict. *Finance Research Letters*, *55*, 103919.

Zumente, I., & Lāce, N. (2021). ESG Rating—Necessity for the Investor or the Company? *Sustainability*, *13*(16), 8940.