UNIVERSITY OF GOTHENBURG
SCHOOL OF BUSINESS, ECONOMICS AND LAW

Does ESG Matter in Controversial Industries?

Evidence From Aerospace and Defense

Authors:

Ludvig Corneliusson

Filip Pankalla

Supervisor:

Jon Williamsson

Bachelor's Thesis in Finance: 15 Credits

Spring Term 2025

Abstract

This paper investigates whether Environmental, Social, and Governance (ESG) scores affect risk-adjusted stock returns and market resilience within the Aerospace and Defense (A&D) industry. The study addresses a gap in existing research, as industries like A&D are often excluded from ESG-focused studies due to their controversial nature, despite becoming an increasingly integral part of social stability following recent geopolitical instability. We expand upon previous literature by employing a panel regression approach over the period 2016-2023, and divide this period into four sub-periods. We assess both individual ESG scores E, S and G, and a composite score. Our results indicate that ESG scores do not consistently explain differences in risk-adjusted returns or return volatility. Only a weak statistical relationship could be established, indicating that higher ESG performance is associated with lower risk-adjusted returns. No significant effect on return volatility was found during the COVID-19 pandemic, challenging the notion that ESG provides resilience during market turbulence. This paper contributes to the literature by providing new evidence from the A&D sector, while also reaffirming the broader challenge of establishing clear and consistent relationships between ESG scores and financial performance.

JEL Classification: G11, G14, M14

Keywords: ESG, risk-adjusted return, market resilience, defense industry, aerospace, panel regression, COVID-19, stakeholder theory

Acknowledgements:
We would like to thank our supervisor, Jon Williamsson, for his guidance and feedback during the making of this thesis. We also appreciate the helpful comments and discussions provided by our fellow student opponents.

Table of Contents

1 Introduction 4
  1.1 Background 4
  1.2 Purpose and research questions 6
2 Theoretical Framework 7
  2.1 Efficient Market Hypothesis 7
  2.2 Capital Asset Pricing Model 7
  2.3 Sharpe Ratio 8
  2.4 Environmental, Social, and Governance 9
  2.5 Stakeholder Theory 11
3 Previous Research on ESG and Controversial Sectors 13
  3.1 ESG and firm performance 13
  3.2 ESG in Sin sectors 14
  3.3 ESG during the COVID-19 pandemic 16
4 Methodology 18
  4.1 Econometric framework and period segmentation 18
  4.2 Model specifications 18
    4.2.1 Risk-adjusted return model 19
    4.2.2 Return volatility model 19
  4.3 Calculations and assumptions 20
  4.4 Robustness considerations 21
5 Data 24
  5.1 Data collection 24
  5.2 Variable descriptions 25
  5.3 Data preparation 26
  5.4 Descriptive statistics 27
6 Empirical Results 31
  6.1 Non-event regressions 31
  6.2 COVID-19 event regressions 34
7 Discussion 37
  7.1 Results 37
  7.2 Limitations 39
8 Conclusion 41

1 Introduction

Ethical and sustainable investing has grown rapidly over the recent decades, yet empirical evidence struggles to establish its financial relevance. Industries traditionally associated with unethical business practices, such as Aerospace and Defense, have often been excluded from this line of research, while recent geopolitical events have changed this perspective. This section outlines the context and motivation for our study.

1.1 Background

In 1970, Nobel laureate and father of monetarism, Milton Friedman, popularized the idea that the primary responsibility of a corporation is to maximize returns and value for its shareholders (Friedman, 1970). This highlights the ownership-oriented sentiment of the time, which has since evolved into the idea of a corporation as an agency of all of its stakeholders. More than a decade later, R. Edward Freeman introduced a model in which firms are accountable not only to its shareholders, but also to employees, customers, suppliers, creditors, among others (Freeman, 1984). Freeman's ideas emerged during a time of great developments in organizational and management theory, largely in response to corporate management failures and exploitation that took place during the entirety of the 1980s. Practices of hostile take-overs, led mainly by private equity and hedge funds, involved taking on high levels of debt to acquire undervalued or distressed firms. These acquisitions were often followed by asset liquidations or aggressive cost-cutting with the aim of reselling for a profit, often at the expense of long-term stakeholder interests (Jensen, 1989). Eventually, corporate greed and mismanagement contributed to the largest single one-day stock market crash in history of 22,6% on Monday, October 19, 1987 (Bernhardt & Eckblad, 2013).

The demand for corporate accountability and commitment to sustainable business practices, particularly in the form of Environmental, Social and Governance (ESG) responsibility, has grown rapidly since the organizational failures of the 1980s (Kräussl et al., 2023). With each global economic crisis, each case of corporate corruption, greed and gross mismanagement, corporations that signal a strong commitment to ethical business practices, not only appeal more to ESG-investors, but also imply a lower perceived risk, potentially reducing their cost of capital by justifying a lower required risk premium (Kroll, 2023).

As indicated by Kroll (2023), ethical considerations are not the only driver of demand for ESG integration in modern investing. Strong ESG performance among firms is expected to have implications on the financial performance and stability of firms. Iannone et al. (2025), Borovkova and Wu (2020), and Wang et al. (2025) found evidence of ESG volatility resilience during periods of financial instability. Friede et al. (2015) reviewed over 2000 studies and found over 90% showed non-negative relationships between ESG scores and corporate financial performance. Furthermore, Borovkova and Wu (2020), Gavrilakis and Floros (2023), and Kroll (2023) found that ESG scores can provide information regarding stock return, though with varying degrees across sectors, regions and time periods.

Controversial sectors, or so-called “Sin” sectors (Paradis and Schiehll, 2021), that are typically associated with unethical or unsustainable business practices have long been excluded from ESG-focused research. One of these sectors is Aerospace and Defense (A&D), which historically has faced intense scrutiny for its controversial nature. However, since the Russian invasion of Ukraine in both 2014 and 2022, the topic of national security has increasingly been recognized as a component of corporate social responsibility (CSR) among investors within the context of sustainable business practices (PWC, 2022). More recently, with shifting US foreign defense policy, particularly growing signals of reduced long-term commitment to European security and collaboration, the European Union is being forced to reevaluate its own defense policies and priorities (European Commission, 2025). This changing landscape further contributes to the development of shifting perspectives on controversial sectors’ role within sustainability frameworks, and additional research into the effects of ESG in these particular industries is needed.

While the theoretical foundation linking sustainability and firm performance is evidently well established and extensively discussed in the literature, conclusive empirical evidence of the effects of ESG performance on financial outcomes remains limited. Previous research has been able to show significance within certain industries (Kroll, 2023; Borovkova & Wu, 2020; Gavrilakis & Floros, 2023), but results vary depending on time-frame, region and industry. Given the evolving view of A&D as a contributor to societal stability, particularly in light of heightened geopolitical tensions, there is a clear need to investigate whether ESG scores influence financial performance in this controversial sector.

1.2 Purpose and research questions

The purpose of this thesis is to investigate whether ESG ratings have a statistically significant effect on risk-adjusted stock market returns and market resilience in the Aerospace and Defense (A&D) industry. While ESG scores have shown varying degrees of explanatory power for firm performance depending on region, time period, and industry, limited research has specifically examined this relationship within controversial sectors. This study is particularly relevant in light of recent geopolitical developments, including shifts in US foreign policy, evolving investor perception of the industry and increased defense spending in the European Union, which have disproportionately affected the defense industry. Thus, we aim to contribute to the literature by addressing this research gap.

We aim to fulfill the purpose of this study by answering the following research questions:

1. Do ESG scores within the Aerospace & Defense sector explain differences in firms' risk-adjusted returns?
2. Did ESG scores influence financial resilience among Aerospace & Defense firms during the COVID-19 pandemic?

2 Theoretical Framework

To understand why sustainable investing might explain improved financial outcomes, this section outlines the theoretical foundations that support this perspective. It includes financial theories such as the Efficient Market Hypothesis, the Capital Asset Pricing Model, and Sharpe ratio, as well as sustainability-oriented concepts like stakeholder theory and Environmental, Social, and Governance (ESG).

2.1 Efficient Market Hypothesis

The Efficient Market Hypothesis (EMH), developed by Fama (1970), is a theory which states that if markets are efficient, asset prices reflect all the available information on financial markets. This means that stocks, bonds and other assets always have their fair value at any given time based on available information, and if there are deviations, the market quickly adjusts to the new information. In theory, in an efficient market, it should be impossible to consistently generate alpha through stock-picking, as the theory states that the assets are correctly priced at any given moment, thus no market asset is undervalued or overvalued. Fama (1970) established this theory and categorized markets into three different types based on its efficiency. The weakest efficiency market, meaning that historical price data are reflected in the assets price today. This means that historical price actions are reflected in today's price, making it impossible to consistently predict the future asset price with historical price actions. Then there is the semi-strong efficient market, meaning that all public information, such as announcements of earnings, splits and ESG-scores, and historical data are reflected in today's asset prices. This means that all available public information and historical price data is already priced in assets, indicating that ESG-related information may be priced in by market participants. Lastly, the strong efficiency market states that historical data, public and insider information is already incorporated into the asset price.

2.2 Capital Asset Pricing Model

The Capital Asset Pricing Model (CAPM) was developed by Sharpe (1964), Lintner (1965) and Mossin (1966) throughout the 1960s, and serves as the cornerstone of asset pricing theory. It builds on the idea established by the efficient market hypothesis to explain the relationship between an asset's expected return and its sensitivity to systematic risk, measured by beta (β). According to CAPM, the expected return of a security is a function of the risk-free rate, the security's sensitivity to market movements (beta), and the market risk premium, defined as:

$$E(R_i) = R_f + \beta_i(E(R_m) - R_f)$$

where $E(R_i)$ is the expected return of asset i, R_f is the risk-free rate, typically proxied by the 10-year government bond yield, $E(R_m)$ is the expected market return, and β_i is the beta of asset i. Beta is defined as (Fama & French, 2004):

$$\beta_i = \frac{Cov(R_i, R_m)}{\sigma^2(R_m)}$$

i.e., the covariance between the returns of asset i, and that of the benchmark index (such as the S&P 500), in relation to the variance of the returns of the benchmark. The interpretation is that the beta measures the sensitivity of the assets returns to that of the market, where the market acts as a proxy for systemic risk that cannot be diversified away. According to CAPM, the expected return of an asset is a function of its risk, which then serves as the theoretical foundation for examining risk-adjusted returns and the validity of subsequent conclusions derived from statistical modeling throughout this study.

However, CAPM assumes several problematic characteristics of markets and its participants worth mentioning. Firstly, it assumes investors are rational, and hold diversified portfolios only subjected to systematic risk, while in reality this is a gross generalization. Additionally, it assumes markets are efficient, while Fama determined that markets have varying degrees of informational efficiency and discovered so-called 'fat-tailed' distributions of returns, which some attribute to herd behaviour and asymmetric volatility preferences.

2.3 Sharpe Ratio

The Sharpe ratio, also introduced by Sharpe (1966), is a fundamental metric in modern portfolio theory used to evaluate returns on a risk-adjusted basis, meaning returns are adjusted for the historic volatility. More specifically, it measures the excess return per unit of risk (volatility), i.e., an asset or portfolio's expected return minus the risk-free rate of return, divided by the standard deviation. Sharpe defines it as the following:

$$Sharpe\ ratio = \frac{R_p - R_f}{\sigma_p}$$

where R_p is the return of the portfolio or asset, R_f is the risk-free rate, and σ_p is the standard deviation of the return of the portfolio or asset. The numerator, $R_p - R_f$, is referred to as the excess return, representing the return earned above the risk-free benchmark.

The main principle is that because market participants are risk-averse they need to be compensated for taking on risk, and therefore, risky assets are to be valued less than 'risk-free' assets with the same expected return. A higher Sharpe ratio thus indicates superior risk-adjusted performance, and conversely, lower or negative ratio suggests inefficient compensation for the risk taken. Further, another assumption of the Sharpe ratio is that returns are normally distributed. While widely used for its simplicity, it comes with several limitations affecting usability that should be acknowledged. With the standard deviation in the denominator, excess returns are penalized *uniformly* regardless of upside or downside fluctuations, while in reality, investors typically do not have a uniform preference for large fluctuations. This observation relates to the concept of loss aversion, introduced by Kahneman and Tversky (1979), which suggests that investors weigh losses more heavily than equivalent gains. Finally, estimations of excess return and volatility rely solely on historical data, effectively treating preferences for risk-adjusted returns as stationary over time, while individual risk preferences have been shown to vary across investors and time (Vissing-Jorgensen, 2003).

2.4 Environmental, Social, and Governance

ESG is a metric that evaluates companies in different non-financial metrics and it has gained traction over the last decade since many investors prefer to invest in companies that are ethical and sustainable, which has later been called ESG-investing (Pollman, 2022). Furthermore, a comprehensive study, conducted by Friede et al. (2015), found that approximately 90% out of the 2000 empirical papers studied had non-negative relationships between ESG ratings and corporate financial performance (CFP). Moreover, other research states that ESG ratings can be used as a mitigating factor for exogenous events (Wang et al., 2025).

A company's ESG rating is assessed based on three different categories: environmental, social and governance. The environmental category is evaluated based on firms' utilization of natural resources and how firms' operations affect the environment, both direct operations and across its supply chains (S&P Global, 2019). This means that the environmental pillar is evaluated based on a company's disclosure, impact and efforts to reduce its carbon emissions, i.e., issues that pose risks and opportunities for both stakeholders and shareholders.

While the environmental factor focuses on a company's effect on the planet, the social pillar focuses on a company's interaction with people and institutions (S&P Global, 2020). The social category is evaluated based on the company's relationship with its workforce, the society that it operates in and the political environment. Importantly, the social pillar is not just evaluated based on consumers' judgment of a firm, geopolitical events are also social factors. For instance, before Russia's invasion of Ukraine in 2022, the A&D sector had difficulties in getting favourable ESG-ratings due to their military connection. However, there is new evidence that institutional and retail investors have changed their perspective, for some firms in the A&D sector, by considering them socially responsible (Wille et al., 2024). Since the defense sector is crucial for national security, which is necessary for protecting peace and freedom, the view on the sectors' social responsibility has evolved by investor behaviour and by recent geopolitical events.

Lastly, the G in ESG stands for governance factors of decision-making, which includes sovereign policy, but also the allocation of liability on corporate level from board of directors, managers, shareholders and stakeholders (S&P Global, 2020). As climate risks and societal implications are often prioritized, governance factors are often overlooked. However, it is essential to be accustomed to governance risks and opportunities in decision-making. Corporate scandals, such as Volkswagen's emission tests scandal and Facebook's misuse of data occurred due to poor decision-making, which led to immense financial damages for these corporations, but also shareholders and stakeholders (S&P Global, 2020). As corporate scandals unfold with increased awareness of global diversity and unequal incomes, governance is a crucial part of ESG.

Overall, the ESG metric is favourable as it pushes firms to tackle environmental threats, societal implications and weak governance decision-making. However, ESG has its shortcomings and controversies. For instance, the rating can significantly differ between different ESG-agencies since there's no universal methodology on how to assess a rating (Dorfleitner et al., 2015). Dimson et al. (2020) found remarkably low correlations between ESG ratings across different agencies, for both the ESG composite score and for individual pillars. For example, the correlation of ESG composite score between FTSE and Sustainalytics was 59%, however, the governance score from MSCI and Sustainalytics exhibited a staggering average correlation of -2%. This inconsistency weakens the credibility for the metric and could potentially give investors, shareholders and stakeholders the wrong perception of firms ESG performance.

Furthermore, as climate change and global warming have been key concerns for decades, firms are increasingly considering climate resilience as key aspects of their strategy (Ghitti et al., 2024). Further, investors, customers and stakeholders are increasing the pressure on corporations to disclose their exposure to climate risk and how to handle those risks. However, this increased pressure on disclosing climate risks and managing risks from stakeholders, investors and shareholders can lead firms to engage in *greenwashing*, i.e., deceiving ESG-agencies, investors, shareholders and stakeholders about the firm's environmental impact (Laufer, 2003).

2.5 Stakeholder Theory

While the origins of stakeholder theory is debated, it is most commonly attributed to Freeman (1984), who formalized the idea that firms have responsibilities beyond those to its shareholders. The core principle of his theory states that a corporation that considers the interests of a broader set of stakeholders, such as employees, customers, suppliers, communities, regulators, and shareholders, can improve long-term value creation by reducing reputational, regulatory, and operational risks. In the context of ESG and sustainability, stakeholder theory helps explain how responsible corporate behaviour may translate into superior firm performance, particularly through improved risk management. According to Wang (2024), corporate ESG behaviour is a way for businesses to fulfill their responsibilities to their stakeholders. Further, the author argues that ESG performance can serve as a strategic tool to strengthen internal cohesion and external legitimacy, which in turn improves resource access, reduces financing constraints, and enhances corporate performance. These mechanisms are especially relevant in controversial sectors such as A&D, where reputational risks and stakeholder scrutiny are typically higher.

To summarize, the theoretical foundations outlined above provide a framework for understanding how risk, return and sustainable business practices interact in financial markets. CAPM and Sharpe ratio help explain the trade-off between risk and return, while the Efficient Market Hypothesis sets expectations on how ESG-related information may be priced by market participants. Environmental, Social, and Governance, and stakeholder theory provides reasons for why non-financial metrics may influence firm performance, particularly in controversial sectors, such as A&D. These theoretical perspectives are important for evaluating existing empirical findings, which vary widely across sectors and time periods, and by supporting the conclusions that can be drawn from our own analysis. The next section reviews prior studies regarding ESG and market performance, and helps position our own contribution within the existing body of research.

3 Previous Research on ESG and Controversial Sectors

Previous empirical evidence on ESG performance shows inconsistent results across regions, time periods and sectors, with limited attention given to controversial sectors such as A&D. This section reviews existing studies on ESG-related returns and market resilience, highlights gaps in sector-specific evidence, and forms the basis for our hypothesis development.

3.1 ESG and firm performance

As mentioned earlier, there is a significant research gap regarding ESG ratings impact on the financial performance in the A&D sector. However, there is abundant research on the relationship between ESG ratings and financial performance with varying results.
Friede et al.'s (2015) comprehensive study found that approximately 90% of the 2000 empirical papers studied had non-negative relationships between ESG ratings and corporate financial performance (CFP). Even though the result is generalized, it provides an overall overview of ESG's impact on financial performance. This finding is in line with Kroll's (2023) paper, which studied if higher ESG ratings provide superior stock market return. They constructed multiple portfolios based on MSCI's ESG rating system, and grouped them as leaders (AAA, AA), average (A, BBB, BB) and laggards (B, CCC), with a market capitalization-weighted portfolio approach. The study found that 10 out of the 12 regions studied exhibited that leaders had superior stock market return compared to average and laggards, with the exemption of Brazil and Germany, with no further explanation why that is. Notably, in the US market, the average ESG-portfolio in the industrial sector (e.g., A&D sector) outperformed "leaders" and "laggards". However, the study emphasizes that, when analyzing in specific regions or industries, the relationship between ESG rating and stock market return often becomes weak. This could be interpreted that ESG ratings are not included in the investment strategies in those regions and sectors, making ESG an insignificant explanatory variable for stock market return.

Contrary to Friede et al. (2015) and Kroll's (2023) findings, Gavrilakis and Floros (2023) study found a significant negative relationship between ESG ratings and stock market performance. The study was conducted on large-cap companies in six different countries (Portugal, Italy, Greece, Spain, France and Germany) from 2010 to 2020 and used combined ESG-score, since combined ESG-score is updated more frequently. The authors found that the Italian market (FTSEMIB) exhibited a significant negative relationship between ESG score and stock market return, while the other five markets had non-significant ESG coefficients. Further, the study applied the model on the Euronext100 index, which found a small but significant negative relationship between ESG scores and stock market return. Gavrilakis and Floros (2023) argued that investors in the Italian- and European market were, in that time period, willing to sacrifice stock market returns for sustainable investments. These findings are supported by Borovkova and Wu's (2020) study, which also exhibited a significant negative relationship between ESG scores and financial performance in the US market (S&P 500) and the Asian market. They suggested that these results were likely due to sustainability being a lower priority in those regions compared to other regions, such as Europe, leading to lower investor demand for highly ESG-scored stocks. Overall, the literature suggests that the relationship between ESG performance and financial performance is dependent on region and industry.

3.2 ESG in Sin sectors

Given the abundant but mixed evidence from the relationship between ESG ratings and firms' financial performance across different regions, analysing its impact on similar sectors to A&D could enhance the understanding. To better understand the impact of ESG in the A&D sector, and since there are limited available studies on A&D, it is useful to examine similar sectors that struggle to get favourable ESG scores. Specifically, vice industries or "Sin sectors", i.e., publicly traded firms whose operations may be considered unethical or immoral (Paradis and Schiehll, 2021). Therefore, a deeper understanding of what Sin sectors' relationship to ESG is, and how this, in turn, affects financial performance, is necessary.

Paradis and Schiehll (2021) examined the ESG performance for 79 sin stocks and compared it to a control group of comparable firms. Their conclusion was that Sin stocks have a significantly lower ESG composite score, but also for every individual pillar in ESG, compared to the control group. Their findings are in line with Iskenderoglu's (2025) study on the complex relationship between ESG disclosure, the value of diversification and investments in controversial sectors. The study examined how acquisitions of controversial firms (e.g., sin stocks) affected conglomerate corporations and single-sector firms ESG-ratings and its financial performance, with excess returns as proxy, i.e., the natural logarithm of the ratio between a firm's market to sales value with its estimated market to sales value. Iskenderoglu (2025) suggests that acquiring firms in controversial sectors, such as sin-stocks, significantly decreases the acquirer's ESG rating, and notably, its valuation.

Hong and Kacperczyk (2009) study on social norms and its impact on financial markets found that sin stocks tend to have 15-20% lower valuation ratios (e.g., price-to-book ratio) than comparable firms. Further, they found evidence that sin stocks outperforms comparable firms by 2,5% per year. They present evidence that sin stocks have approximately 18% lower institutional ownership. However, a further analysis was conducted on institutional ownership in sin stocks and they found that mutual funds and independent investment advisors did not hold a smaller share of sin stocks than comparable firms. They argue that lower institutional ownership cannot be associated with institutions being smarter than retail investors, since hedge funds and mutual funds are among the smartest investors. Instead, Hong and Kacperczyk (2009) suggest that lower institutional ownership in sin stocks are due to societal norms, hence some institutions avoid investing in sin stocks.

Although Hong and Kacperzyk (2009) suggests that sin stocks have lower institutional investors, a report from PWC (2022) indicates a changing perspective on investments in the A&D sector. Before Russia's invasion of Ukraine, institutional interest in the sector declined due to ESG concerns, despite its strong profitability. However, the perception of the industry has changed for institutional investors, since investments have increased into the A&D-sector after Russia's invasion of Ukraine. Findings from the report showed that institutional investors that had a minimum of 5 billion dollars invested in the top 16 internationally renowned A&D-companies (excluding China), 116 institutions increased their position after Russia's invasion of Ukraine. Further, only 71 institutions decreased their investment under the same timeframe. This finding contributed to the conclusion that on average, investments from institutions increased by 6% in 2022 and continued to increase in early 2023.

While sin sectors have been studied in relation to ESG and financial performance, there is lack of empirical evidence if ESG scores have a significant explanatory power on financial performance in the A&D sector. With new reports of increased institutional ownership, suggesting a change in perspective regarding its ESG scores, it is highly relevant to examine if ESG have a significant explanatory power on risk-adjusted stock market return for investors. This leads us to the formulation of our first hypothesis:

Hypothesis 1 - ESG and risk-adjusted return:

- H_0: ESG scores have no significant effect on risk-adjusted return within the Aerospace & Defense industry.
- H_A: ESG scores have a significant effect on risk-adjusted return within the Aerospace & Defense industry.

3.3 ESG during the COVID-19 pandemic

Prior research suggests that ESG scores influence firms' financial performance, though with varying results on its impact, depending on region and industry. However, we hypothesize that a better ESG score should decrease the risk for stocks, making it more resilient to exogenous events such as the COVID-19 market crash.

Wang et al. (2025) studied the influence of ESG scores on portfolio performance in the Chinese stock market, specifically during different phases of the COVID-19 pandemic. They divided the sample into pre-, during- and post-pandemic periods and constructed ESG portfolios using traditional Markowitz mean-variance optimization, maximizing risk-adjusted returns. Their findings revealed that ESG portfolios outperformed the market during the pandemic, but had no or negative impact in the surrounding periods. They speculated that the insensitivity of ESG-investors might be the reason, arguing that investors that prioritize sustainability in their investing practices have a higher tendency to hold on to their assets in times of uncertainty. They also argued that an ESG-weighted portfolio may serve as a hedge and protection against rapid economic downturn. Further, they used ESG data from Bloomberg, which provides scores for all E, S and G, and found, contrary to what we have previously seen, that S provided the largest positive prediction in stock performance, while G had a negative effect. This, they reasoned, was likely due to the shifting investor sentiment towards sustainable practices during times of social and economic stress.

In contrast to Wang et al. (2025), Demer et al. (2021) found robust evidence that for US stocks, ESG offered no explanatory power for returns in the COVID-19 crisis, once controlling for industry, market-based measures of risk, financial performance, financial position and intangible investments. Borovkova and Wu (2020) also found that US stocks may reflect a lower priority for sustainability, leading to lower investor demand for highly ESG-scored stocks.

Contrary to Demer et al. (2021), Cardillo et al. (2023) found that European stocks with higher ESG scores had a significantly mitigating effect during the COVID-19. Additionally, sustainable European stocks performed better on the stock market than less sustainable firms during 2020, with greater stock market returns, lower volatility and higher stock market liquidity. These findings are in line with Engelhardt et al. (2021) who studied if European firms with higher ESG score performed better during the COVID-19 crash. The results were that firms with better ESG scores exhibited significantly higher abnormal stock market returns and lower idiosyncratic risk, in the beginning of 2020.

The studies exhibited different results on how ESG scores impact stock market resilience during the COVID-19 pandemic. Wang et al. (2025) found that portfolios with better ESG score outperformed the market during the pandemic, specifically, the social factor. However, Demer et al. (2021) and Borovkova and Wu (2020) found no evidence in the US market of ESG scores providing market resilience, while Cardillo et al. (2023) and Engelhardt et al. (2021) found evidence in the European market. This suggests that ESG scores may reduce risk during a crisis, depending on the region and component of ESG. Ultimately, this leads us to our second hypothesis that ESG performance significantly shielded A&D firms during the COVID-19 pandemic.

Hypothesis 2 - ESG and exogenous shock resistance:

- H_0: ESG scores had no significant impact on Aerospace & Defense firms' market resilience during the COVID-19 pandemic.
- H_A: ESG scores significantly impacted Aerospace & Defense firm's market resilience during the COVID-19 pandemic.

4 Methodology

This section explains the panel regression framework used to assess the relationship between ESG scores and financial performance in the A&D sector. It covers time period segmentation and the application of different models, outlines key calculations and assumptions, and discusses robustness considerations, including the Hausman test.

4.1 Econometric framework and period segmentation

The main empirical analysis is applied using panel data through linear regression models, drawing inspiration from Borovkova and Wu (2020) and Gavrilakis and Floros (2023). In contrast to Gavrilakis and Floros who used raw stock returns, our model utilizes the risk-adjusted return measurement, Sharpe ratio, as the dependent variable and stock performance metric. This allows for capturing both excess return and the volatility of those returns into one metric (Bhatia and Kumar, 2024; Sandu, 2024; Boubaker et al., 2023).

As two hypotheses have been formulated regarding ESG scores' relationship with risk-adjusted return and market resilience, we are conducting regressions across three distinct subperiods: 2016–2019, 2020–2021 (COVID period), 2021–2023, as well as for the full period from 2016 to 2023. This approach follows a similar structure to Wang et al. (2025). However, a key distinction is that the primary objective here is not to assess the effects of ESG changes before and after COVID-19, but rather to examine whether the relationship between ESG ratings and firm performance differs across distinct market environments. Additionally, by applying our models on an isolated COVID period, we are evaluating whether ESG investors behave differently during such periods, holding on to their stocks to a higher degree, as suggested by Wang et al. (2025).

4.2 Model specifications

Here we present two models that are applied differently on our subperiods: a risk-adjusted return model that is applied to the full sample and all sub-periods (including COVID), and another tailored to assessing market resilience during the COVID-19 shock. Both rely on a panel data structure and fixed-effects estimation, while also including a variation in which individual pillar scores are changed for a composite ESG score.

4.2.1 Risk-adjusted return model

To estimate the relationship between ESG and risk-adjusted performance, the following linear panel regression models are specified:

$$(1)\ Sharpe_{it} = \alpha_i + \beta_1(logMktCap)_{it} + \beta_2(P/B)_{it} + \beta_3(Beta)_{it} + \beta_4(EBIT)_{it} + \beta_5(E)_{i,t-1} + \beta_6(S)_{i,t-1} + \beta_7(G)_{i,t-1} + \epsilon_{it}$$

$$(2)\ Sharpe_{it} = \alpha_i + \beta_1(logMktCap)_{it} + \beta_2(P/B)_{it} + \beta_3(Beta)_{it} + \beta_4(EBIT)_{it} + \beta_5(ESG)_{i,t-1} + \epsilon_{it}$$

where the dependent variable, $Sharpe$, serves as a proxy for stock performance for firm i in year t. α is the intercept, $logMktCap$ is the logarithm of market capitalization to control for firm size, and P/B (price-to-book ratio) accounts for systematic differences in valuation relative to net assets. $Beta$ captures the effect of systematic risk, while $EBIT$ (earnings before interest and taxes) is a margin and serves as a proxy for operational efficiency. E (Environmental), S (Social) and G (Governance) are all used to assess how sustainable a firm is relative to its industry peers. The second model replaces the individual pillar scores with a composite score to capture overall sustainability performance.[1]

4.2.2 Return volatility model

To assess market resilience, we apply the following panel regressions:

$$(1)\ Vol_{it} = \alpha_i + \beta_1(logMktCap)_{it} + \beta_2(P/B)_{it} + \beta_3(EBIT)_{it} + \beta_4(E)_{i,t-1} + \beta_5(S)_{i,t-1} + \beta_6(G)_{i,t-1} + \epsilon_{it}$$

$$(2)\ Vol_{it} = \alpha_i + \beta_1(logMktCap)_{it} + \beta_2(P/B)_{it} + \beta_3(EBIT)_{it} + \beta_4(ESG)_{i,t-1} + \epsilon_{it}$$

[1] For detailed variable explanations and reasoning, see Section 5.2.

where the dependent variable has been replaced by annualized return volatility, Vol, for firm i in year t. Additionally, $Beta$ has been excluded to avoid multicollinearity, given that both return volatility and $Beta$ are measurements of risk.

4.3 Calculations and assumptions

Annual return was estimated by taking the historical pricing data per stock, and calculating daily percentage returns. These daily returns were then compounded across all trading days within each year to obtain the annual total return per firm, as per the following formula:

$$Cumulative\ Return = \prod(1 + r_{daily}) - 1$$

This method captures the effect of daily compounding, reflecting the actual return experienced by investors during each calendar year. Further, annualized volatility was computed as the standard deviation of daily returns multiplied by the square root of number of trading days, approximated as 252, consistent with standard practice in financial modeling (Bodie et al., 2021):

$$Annualized\ volatility = \sigma_{daily} \times \sqrt{T}$$

To calculate Sharpe ratios, excess return was estimated by taking the annual return per stock and subtracting a risk-free rate proxy by region for each respective year. The proxy used was the 10-year government bond yield per region, taking inspiration from Borovkova and Wu (2020), and regions were subdivided based on regional frequency in our sample, in addition to practicality and data availability. Thus, we opted to divide bond yields into the following regions: US, Europe, China, UK, Canada, and Japan. The remaining countries not pertaining to these regions were grouped into an "other" category, utilizing an average yield of all other regions. This approach ensures that excess return reflects region-specific opportunity costs of capital, and is important in ensuring comparability and methodological robustness. Each respective stock was then paired with the 10-year government bond yield corresponding to the year and region for that observation, allowing for the calculation of excess returns on a regionally adjusted basis. Finally, individual Sharpe ratios were estimated by dividing the excess return by the annualized volatility.

Table 1

Average Annual Bond Yield Per Country/Region

Year	Canada	China	Europe	Japan	UK	US	Average
2016	0.0125	0.0288	0.0093	-0.0007	0.0131	0.0184	0.0136
2017	0.0178	0.0357	0.0117	0.0005	0.0124	0.0233	0.0169
2018	0.0228	0.0370	0.0126	0.0006	0.0146	0.0291	0.0195
2019	0.0159	0.0324	0.0058	-0.0011	0.0094	0.0214	0.0140
2020	0.0075	0.0296	0.0021	-0.0001	0.0037	0.0089	0.0087
2021	0.0136	0.0311	0.0020	0.0007	0.0079	0.0144	0.0116
2022	0.0277	0.0282	0.0204	0.0023	0.0245	0.0295	0.0221
2023	0.0336	0.0281	0.0327	0.0056	0.0406	0.0396	0.0300

Table 1 presents average annual government bond yields across selected countries and regions from 2016 to 2023. The data show clear differences in yield levels, with the US and China consistently exhibiting higher yields compared to other regions. Europe and Japan maintained persistently low yields throughout the period, with Japan even recording negative rates in several years.[2]

Finally, market capitalization was logarithmically transformed to capture the non-linear relationship between firm size and stock returns. Prior research suggests that the marginal effect of size on returns diminishes as firm size increases, making a logarithmic transformation appropriate. Additionally, it also reduces skewness and mitigates the influence of extreme values, ensuring that the variable behaves more consistently in the regression analysis (Hong & Kacperczyk, 2009).

4.4 Robustness considerations

By following Hong and Kacperczyk (2009) and Iskenderoglu's (2025), we apply firm-clustered standard errors in our regression to account for intra-firm autocorrelation and heteroskedasticity within firms over time. Considering autocorrelation is likely to occur in our dataset since firm-specific factors such as ESG policies within firms remain the same over the years, the residuals may be correlated over time. Heteroskedasticity may also appear since the variance of residuals could differ between firms based on different characteristics, such as varying volatility based on market capitalization. This approach is also recommended by Petersen (2009).

[2] See Section 5.4 for a graphical representation of bond yields.

Because we are applying panel data with recurring firm observations, we need to take into consideration firm-specific effects that are not directly implemented in the model. These effects capture time-invariant characteristics unique to each firm. To control for such endogeneity concerns, different panel estimators can be used, primarily *fixed effects (FE)* or *random effects (RE),* depending on the assumed relationship between the firm-specific effects and the explanatory variables. Typically, one would formally test the presence of these by applying the Hausman test (Borovkova & Wu, 2020). The Hausman test, as explained by Amini et al. (2012), is conducted by testing the presence of correlation between the unobserved component, u_i, and the independent variables, X_{it}. The Hausman test statistic follows a chi-squared distribution with degrees of freedom equal to the number of explanatory variables included in the model. The test compares the coefficient estimates from the fixed effects and random effects models, and evaluates whether their differences are statistically significant:

$$H_0 : Cov(u_i, X_{it}) = 0 \text{ (Random effects appropriate)}$$

$$H_A : Cov(u_i, X_{it}) \neq 0 \text{ (Fixed effects required)}$$

Meaning, if the covariance between the unobserved component and the regressors is zero, a random effects panel regression is appropriate, while if the alternative hypothesis is true it follows that a fixed effect estimation is required (Amini et al., 2012). Following this reasoning, we applied the Hausman test during our modeling phase to determine the appropriate model specification for our data.

Table 2

Results From the Hausman Test

Test Statistic	Value
Chi-squared statistic (χ^2)	34.000
Degrees of freedom	8
p-value	0.0000
Decision	Reject Random Effects, use Fixed Effects (FE)

The results from the Hausman test shows a significant result ($p < 0,001$), indicating that the fixed effects model is preferred over random effects. This suggests that unobserved individual heterogeneity is correlated with the regressors, violating the random effects assumption.

5 Data

This section outlines the construction of the panel dataset used to examine the relationship between ESG performance and financial outcomes. ESG and accounting data were obtained from LSEG, while pricing data was retrieved via Yahoo Finance. To address methodological concerns, ESG scores were lagged, market capitalization was log-transformed, and region-specific risk-free rates were used to compute excess returns. Descriptive statistics are provided to summarize the distribution and structure of the dataset.

5.1 Data collection

As historical ESG data availability was found to be the primary limiting factor in our analysis, we collected non-pricing data for all the stocks within the Aerospace and Defense industry above a certain size, from the London Stock Exchange Group (LSEG), one of the leading financial data providers in the world (London Stock Exchange Group, n.d.). The size selection was limited to small to large capitalization firms, i.e., firms with a market capitalization larger than $250M. The reasoning for excluding firms smaller than this is to ensure the sample has reliable data and limit potential outliers that may affect the results, but also because LSEG has limited ESG data for smaller firms. In general, smaller firms tend to be more volatile than small, mid, and large capitalization firms (Stockdale, 2014).

Raw data was collected for fiscal year (FY) 2024 through 2013, with varying degrees of data availability and ESG being the primary limiting factor. Financial data is typically reported according to each firm's fiscal year-end, which may differ across firms. No adjustments were made to align fiscal years to calendar years, consistent with common practice in financial panel data studies (Fama & French, 1992; Barth et al., 2017). Ultimately, a raw sample of 318 stocks was extracted from LSEG over the period FY 2024 through 2013.

For pricing data, the open-source Python library *yfinance* was used to collect adjusted closing prices for the specified period, with data sourced from Yahoo Finance (n.d.). This approach was selected primarily for its practicality, as it allows for historical market data to be downloaded and processed easily. However, due to inconsistencies in stock coverage between the two data providers, mainly missing pricing data for small Chinese and Russian firms covered by LSEG, the sample size was reduced. From the initial set of 318 stocks from LSEG within the A&D segment, 185 could be tracked reliably with regards to pricing data.

5.2 Variable descriptions

The use of Sharpe ratio as the chosen performance metric was inspired by similar research conducted by Bhatia and Kumar (2024), Sandu (2024), and Boubaker et al. (2023). The main concern regarding the usage of a risk-adjusted performance metric is to specifically measure how investors are compensated per unit of risk, and the authors argue that this is a primary concern of ESG-oriented investors. In contrast, relying on raw stock return as a performance metric needs detailed discussions regarding what conclusions can be drawn, and can otherwise be misleading, as they may favor highly volatile stocks.

Among the independent variables, market capitalization (MktCap) is included to control for firm size (Borovkova & Wu, 2020; Gavrilakis & Floros, 2023), and logarithmically transformed to address any skewness, non-linearity or heteroskedasticity, typically associated with size control variables (Boubaker et al., 2020; Hong & Kacperczyk, 2009; Zhang, 2025). Beta is a systematic risk measurement, quantifying the firm sensitivity to the overall market. The usage and inclusion of Beta is discussed in Gavrilakis and Floros (2023), and Bhatia and Kumar (2024). Price-to-Book value (P/B) is included to capture relative firm valuation, based on Gavrilakis and Floros (2023), but also included in Borovkova and Wu (2020), and logarithmically transformed in Hong and Kacperczyk (2009). You calculate P/B by dividing the company's current stock price per share by its book value per share, and it is frequently used by investors to spot undervalued stocks.

EBIT (earnings before interest and tax) margin isolates operational efficiency, removing financing and tax effects that can vary widely across firms and countries. This helps with comparability, which is advantageous considering our sample size is multi-national. Whelan et al. (2021), and Friede et al. (2015) discuss the importance of operational efficiency in great detail, and found the most commonly used parameter to be return on assets (ROA) and return on equity (ROE). Iskenderoglu (2025), and Hong and Kacperczyk (2009) both include EBIT divided by sales as a proxy for operational efficiency.

To explore firm market resilience during the COVID-19 pandemic, we used stock return volatility (Vol) as the dependent variable, employing an approach similar to Wang et al. (2025), Broadstock et al. (2021), and Yin et al. (2023). Iannone et al. (2025) assess resilience by analyzing the volatility dynamics using EGARCH-in-mean models, and Borovkova and Wu (2020) look specifically at return losses to determine market resilience. Finally, ESG scores are all lagged by one period, following the methodological approach applied by Borovkova and Wu (2020), and Gavrilakis and Floros (2023). Gavrilakis and Floros explain this is to mitigate simultaneity bias and endogeneity concerns, as ESG scores are typically updated on an annual basis, while financial data is updated frequently.

5.3 Data preparation

Several steps were taken during the data preparation phase to ensure the dataset was suitable for regression analysis. Market capitalization (MktCap) was logarithmically transformed to reduce skewness and improve interpretability, as discussed in the previous section. ESG pillar scores and the composite ESG score were lagged by one period to mitigate simultaneity bias and endogeneity concerns. Pricing data was retrieved using the yfinance API, from which annual return, annualized volatility, and Sharpe ratio were computed. These variables, along with the financial and ESG data, were merged and structured into a panel dataset.

Firms were grouped into geographical regions, and corresponding annual regional risk-free rates were assigned to each observation. For firms that did not align with any of the main regional classifications, an "Other" category was created based on observation frequency and data availability. This group was assigned an average bond yield calculated from the primary regions (see Table 1). Non-numeric variables and observations with missing values were excluded prior to analysis. To maintain analytical integrity, no data imputation was performed. Finally, extreme values were excluded using a robust z-score approach.

5.4 Descriptive statistics

Table 3

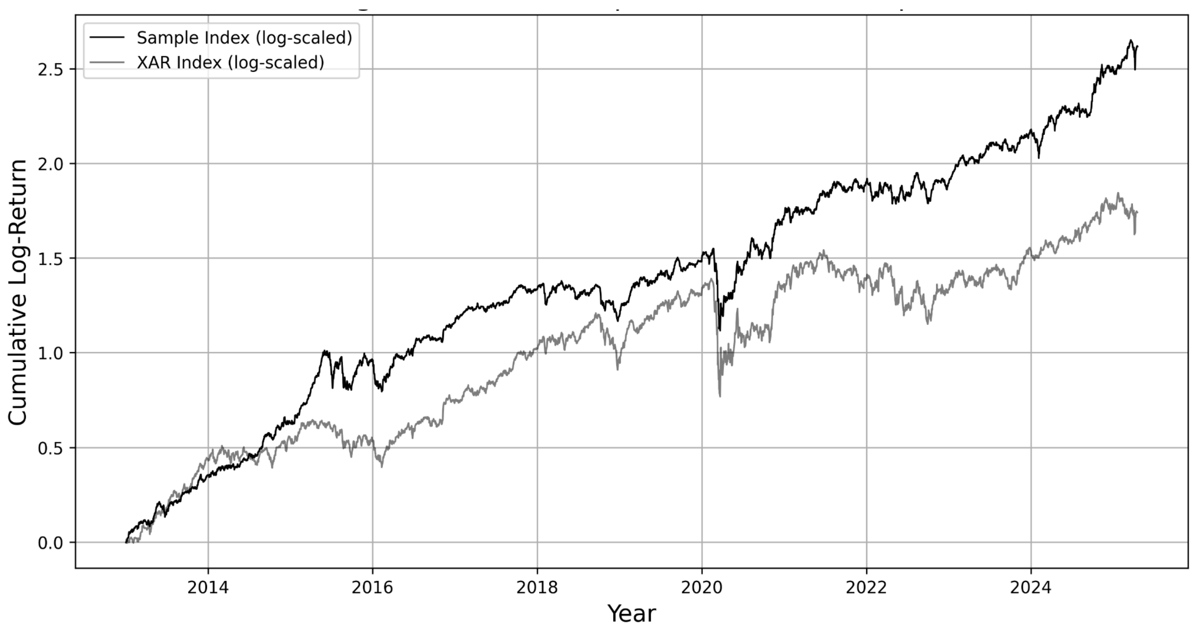
Descriptive Statistics for the Final Sample

Variable	count	mean	std	min	25%	50%	75%	max
Sharpe	463	0.4049	0.9887	-2.3228	-0.3388	0.2441	1.0629	3.4197
Vol	463	0.3385	0.1273	0.1199	0.2409	0.3188	0.4041	0.7721
logMktCap	463	1.1983	0.9640	-1.1699	0.5656	1.3594	1.9365	2.8342
P/B	463	2.5196	1.2036	0.4602	1.6163	2.3042	3.2548	5.5444
EBIT	463	0.0822	0.0427	-0.0096	0.0544	0.0769	0.1098	0.1846
Beta	463	1.0604	0.3923	0.2271	0.7625	1.0416	1.3043	1.9173
EnvScore	463	44.5219	25.8932	0.1668	22.3805	45.0295	63.3430	97.6626
SocScore	463	52.2708	23.3735	2.1077	34.5742	53.9517	71.7456	95.0213
GovScore	463	54.4719	20.6773	5.9821	40.1611	56.4621	70.5556	94.6402
ESGScore	453	50.8410	19.0432	6.9643	35.4454	50.8906	66.4188	93.0574

Table 3 summarizes the descriptive statistics for all variables included in the full period regressions. Variables such as Sharpe ratio and volatility exhibit wide spreads likely due to heterogeneity in firm performance and risk exposure. The financial variables like logMktCap and P/B ratio display somewhat right-skewed distributions. ESG scores appear broadly distributed, with greater variation in environmental and social dimensions relative to governance. Governance also exhibits a higher mean and a higher minimum value than E and S, suggesting A&D firms demonstrate stronger internal controls, transparency and management.

Figure 1

Cumulative Log-Transformed Returns for S&P 500 A&D ETF "XAR" and Our Sample



As Figure 1 shows, our sample and the S&P 500 A&D ETF “XAR” have experienced a long-term upwarding trend from 2013 to 2025. Especially after Russia's invasion of Ukraine in 2022, our sample-index has surged in value and reached new all time highs, which is likely due to an increased investor interest in the industry. However, as seen in the figure, the exogenous COVID-19 pandemic event contributed to a stock market crash for the A&D sector in 2020 but recovered almost immediately.

Figure 2

Sample Log-Transformed Returns Between 2016-2024

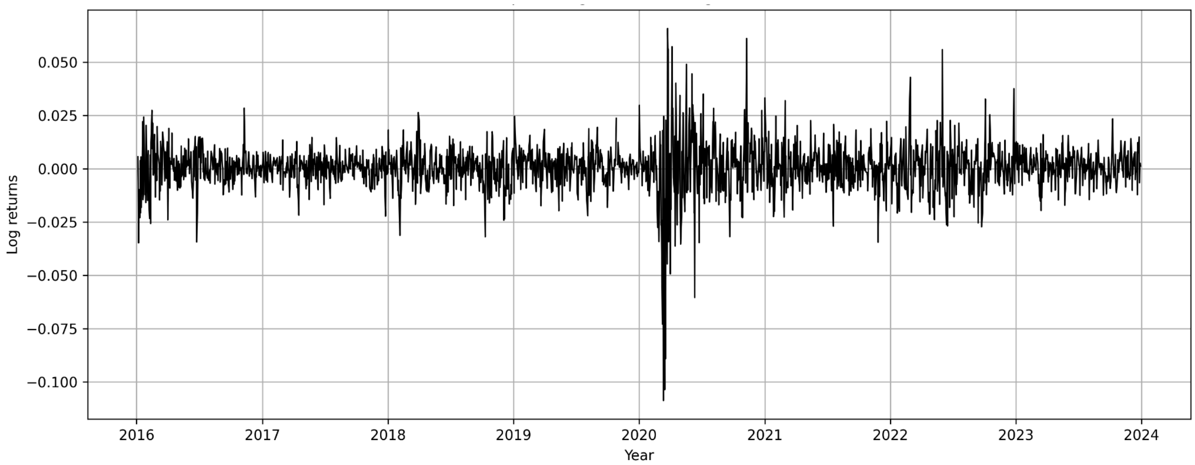


Figure 2 shows the log-transformed returns of our equal weighted sample-index with data spanning from 2016 to 2024. The figure displays trading days with significant returns, i.e., days with increased volatility, and is reflected by large spikes both positive and negative. The most noticeable fluctuation in the figure is in the beginning of 2020, which aligns with the exogenous COVID-19 event and stands out in the figure with significant large spikes of returns, both positive and negative. Further, increased volatility can be observed after COVID-19, which aligns with increased geopolitical instability.

Table 4

Correlation Matrix

	Sharpe	Vol	logMktCap	P/B	EBIT	Beta	EnvScore	SocScore	GovScore	ESGScore
Sharpe	1.0000	-0.1562	0.1866	0.0556	0.2424	-0.0335	0.1069	0.1216	-0.0331	0.0764
Vol	-0.1562	1.0000	-0.2367	-0.0352	-0.1166	0.0458	-0.1363	-0.0775	-0.0780	-0.1124
logMktCap	0.1866	-0.2367	1.0000	0.1924	0.1970	0.0493	0.5098	0.4307	0.1586	0.4656
P/B	0.0556	-0.0352	0.1924	1.0000	0.0277	0.0102	0.0997	0.0816	0.0738	0.1035
EBIT	0.2424	-0.1166	0.1970	0.0277	1.0000	-0.1325	0.0160	0.0900	-0.0487	0.0313
Beta	-0.0335	0.0458	0.0493	0.0102	-0.1325	1.0000	-0.0352	-0.0601	0.0499	-0.0283
EnvScore	0.1069	-0.1363	0.5098	0.0997	0.0160	-0.0352	1.0000	0.7638	0.3561	0.8569
SocScore	0.1216	-0.0775	0.4307	0.0816	0.0900	-0.0601	0.7638	1.0000	0.3414	0.8923
GovScore	-0.0331	-0.0780	0.1586	0.0738	-0.0487	0.0499	0.3561	0.3414	1.0000	0.6515
ESGScore	0.0764	-0.1124	0.4656	0.1035	0.0313	-0.0283	0.8569	0.8923	0.6515	1.0000

Table 4 provides some insight into our dataset. Sharpe ratio appears to be moderately correlated with EBIT and logMktCap, suggesting that more profitable and larger firms tend to have higher risk-adjusted returns. ESG scores show weak correlation with Sharpe ratio, but are negatively related to volatility, suggesting that ESG performance may contribute to lower risk. The ESG pillars are highly correlated with each other, especially between environmental and social scores. Beta shows weak correlations with all variables, consistent with its role as a market-based risk measure rather than one influenced by firm characteristics or ESG.

Figure 3

10-Year Government Bond Yields Per Sample Region Between 2016-2023 (Averages)

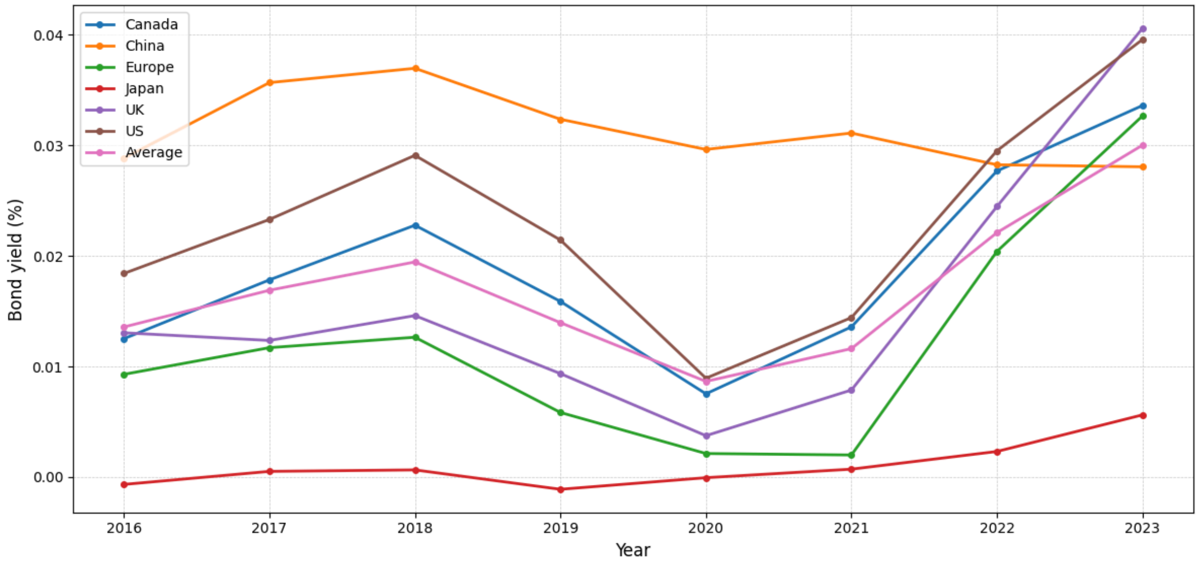
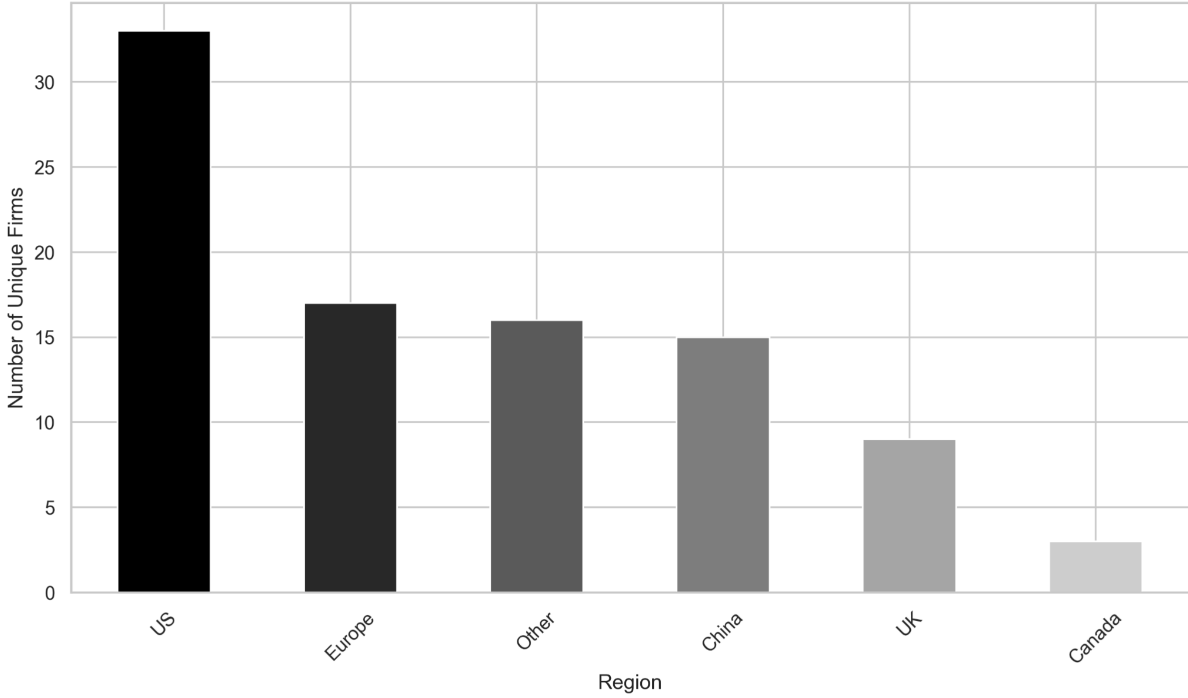


Figure 3 illustrates average 10-year government bond yields across regions between 2016-2023. The figure shows trends that reflect average borrowing costs over the year, smoothing short-term fluctuations while capturing broader monetary policy. What can be noted is a collapse in bond yields following the 2020 pandemic shock in western regions, while Japan and China seem to have survived almost perfectly unscathed. The drop was followed by a relatively quick recovery beginning in 2021, as inflationary pressures emerged and central banks responded by tightening policy. This reversal is particularly visible in the US and the UK, where yields exceeded 3,5-4% on average in 2023.

Figure 4

Sample Distribution by Region



Note: Category "Other" includes Australia, South Korea, Singapore, Israel, Brazil, Turkey, and India.

Figure 4 shows the number of unique firms by region in the final sample. The US dominates the sample, followed by Europe, Other regions, China, the UK, and Canada. The distribution highlights a regional skew, with US firms and Europe dominating the sample, while Canada has much fewer firms included.

6 Empirical Results

Below we present the main empirical findings from our panel regressions. We begin with the non-event regressions and the composite ESG model variation, followed by the application of both models on the COVID-19 period. Below each regression table, we comment on the statistical significance of individual variables and evaluate the explanatory power of the models across subperiods.

6.1 Non-event regressions

Table 5

Non-Event Regressions: Risk-Adjusted Return (ESG Pillar Scores)

	2016-2019	2021-2023	2016-2023
const	-1.690 (1.267)	-1.668 (1.212)	-0.321 (0.488)
logMktCap	1.257** (0.497)	0.695 (0.518)	0.606*** (0.217)
P/B	-0.007*** (0.001)	0.197 (0.133)	-0.005*** (0.001)
EBIT	0.486 (1.805)	0.006 (2.644)	4.150*** (1.199)
Beta	0.454 (0.287)	0.067 (0.820)	0.099 (0.225)
EnvScore	0.010 (0.012)	0.019 (0.013)	0.006 (0.007)
SocScore	-0.029 (0.018)	-0.008 (0.017)	-0.012 (0.009)
GovScore	0.009 (0.009)	-0.001 (0.010)	-0.005 (0.005)
No. Observations	176	226	463
R-squared	0.158	0.157	0.076
Number of firms	53	85	87
F Statistic (Robust)	83.069 (p=0.000)	2.795 (p=0.010)	5.389 (p=0.000)

Note: *p<0.1; **p<0.05; ***p<0.01

Our panel regression estimates for the period 2016–2019 shows that both firm size and relative valuation were statistically significant predictors of risk-adjusted returns. Specifically, larger firms and those with higher P/B ratios were associated with higher Sharpe ratios, implying that market valuation and scale contributed positively to performance in this period. Operational efficiency, systematic risk sensitivity and the ESG components did not exhibit statistically significant effects. The model is statistically significant overall, with a robust F-statistic of 83,1 ($p < 0,01$), and accounted for approximately 15,8% of the within-firm variation in Sharpe ratios.

In the more recent years of 2021-2023, none of the variables showed significance. Overall, the model remains significant, with a robust F-statistic of 2,8 (p < 0,01), and explained 15,7% of the within-firm variation in Sharpe ratios over the period.

For the full period 2016-2023, firm size, relative valuation and operational efficiency were significant determinants of risk-adjusted performance. As for ESG pillars, none of the lagged components could be shown to have explanatory power in terms of variability in performance over the period. Overall, the model is highly statistically significant, as indicated by the robust F-statistic of 5,4 with a p-value below 1%, and explained approximately 7,6% of the variation in Sharpe ratios.

Overall, none of the individual ESG pillar scores showed significance for any of the periods, and traditional financial metrics, particularly firm size, valuation and operational efficiency, offered stronger explanatory power for risk-adjusted returns. All models were statistically significant at the 1% level, as indicated by the robust F-statistics. This suggests that despite ESG scores lacking explanatory power, the model as a whole captured meaningful variation in Sharpe ratios.

Table 6

Non-Event Regressions: Risk-Adjusted Return (ESG Composite Score)

	2016-2019	2021-2023	2016-2023
const	-1.755 (1.450)	-0.061 (1.252)	-0.280 (0.475)
logMktCap	1.244** (0.509)	0.635 (0.500)	0.637*** (0.223)
P/B	-0.007*** (0.001)	0.211* (0.125)	-0.005*** (0.001)
EBIT	1.515 (2.186)	0.997 (2.694)	4.160*** (1.127)
Beta	0.538 (0.355)	0.098 (0.770)	0.141 (0.222)
ESGScore	-0.012 (0.022)	-0.027 (0.018)	-0.015** (0.008)
No. Observations	173	219	453
R-squared	0.129	0.154	0.078
Number of firms	51	86	88
F Statistic (Robust)	118.076 (p=0.000)	2.989 (p=0.014)	7.724 (p=0.000)

Note: *p<0.1; **p<0.05; ***p<0.01

Table 6 presents regression results after substituting individual ESG pillars for a composite ESG score. For the period 2016-2019, size and relative valuation again shows significance in determining risk-adjusted return. Specifically, larger firms were associated with higher

Sharpe ratios, while an increase in relative valuation was associated with lower Sharpe ratios. Beta and operational efficiency were not providing information regarding the variation in Sharpe ratios over the period, and the model explained about 12,9% of the variation with a robust F-statistic of 118,1 ($p < 0,00$).

In 2021-2023, relative valuation was weakly significant, and positively associated with Sharpe ratios, while the ESG score remains insignificant. The model was significant overall, explaining about 15,4% of the within-firm variation and a robust F-statistic of approximately 3,0 ($p < 0,014$). As we can see, this period continues to have few significant variables.

As for the full sample period (2016-2023), traditional financial metrics such as firm size, relative valuation and operation efficiency were again significant determinants of risk-adjusted return. However, this time the composite ESG score was associated with a negative effect on Sharpe ratios and showed significance at the 5% level ($p < 0,05$). This implies that a better ESG score was associated with lower risk-adjusted stock market return for A&D firms over the period.

In summary, the regressions this time using a composite ESG score were again statistically significant across all periods. Traditional financial metrics were better at explaining variance in Sharpe ratios, while the ESG score failed to consistently provide explanatory power. Only a negative effect of -0,015 at the 5% level could be established over the full sample period, suggesting that A&D firms with higher ESG scores suffered in terms of risk-adjusted return. Beta failed to show significance in any of the periods.

6.2 COVID-19 event regressions

Table 7

COVID-19 Event Regressions: Risk-Adjusted Return and Return Volatility (ESG Pillar Scores)

	2020-2021
const	0.158 (1.756)
logMktCap	1.492** (0.591)
P/B	0.087 (0.078)
EBIT	6.477** (2.659)
Beta	-1.125 (0.682)
EnvScore	-0.023 (0.020)
SocScore	-0.010 (0.019)
GovScore	-0.006 (0.014)
No. Observations	126
R-squared	0.331
Number of firms	68
F Statistic (Robust)	5.385 (p=0.000)

Note: *p<0.1; **p<0.05; ***p<0.01

	2020-2021
const	1.417*** (0.242)
logMktCap	-0.304*** (0.080)
P/B	0.014 (0.018)
EBIT	-1.924*** (0.552)
EnvScore	-0.006 (0.005)
SocScore	-0.005 (0.004)
GovScore	0.002 (0.002)
No. Observations	126
R-squared	0.315
Number of firms	68
F Statistic (Robust)	7.281 (p=0.000)

Note: *p<0.1; **p<0.05; ***p<0.01

Note: Left regression uses Sharpe ratio as the output variable, and the table on the right uses return volatility.

Through application of our model during COVID-19, the results indicate that firm size and operational efficiency remained significant drivers of risk-adjusted performance. In contrast, relative valuation, Beta, and all lagged ESG components were not statistically significant. The model is overall highly significant, with a robust F-statistic of 5,4 (p < 0,01), and explained a relatively high 33% of the within-firm variation in Sharpe ratios.

Setting the output variable to return volatility during the COVID-19 period shows that firm size and operational efficiency are significant determinants of return volatility. Specifically, larger firms and those with higher EBIT margins experienced lower volatility. The remaining variables were not exhibiting statistical significance. The model was strongly significant overall, with a robust F-statistic of 7,3 (p < 0,01), and explained approximately 31,5% of the within-firm variation in return volatility.

Overall, both our models explained approximately 30% of the within-firm variation in both Sharpe ratios and return volatility in the A&D sector over the COVID period. ESG scores failed to provide any significant explanatory power, while firm size and operational efficiency both had significant beneficial effects on risk-adjusted return and volatility.

Specifically, larger firms with higher EBIT margins were associated with lower return volatility and higher Sharpe ratios.

Table 8

COVID-19 Event Regressions: Risk-Adjusted Return and Return Volatility (ESG Composite Score)

	2020-2021
const	-0.778 (1.421)
logMktCap	1.512** (0.605)
P/B	0.065 (0.077)
EBIT	6.203** (2.820)
Beta	-0.985* (0.514)
ESGScore	-0.020 (0.024)
No. Observations	126
R-squared	0.310
Number of firms	68
F Statistic (Robust)	4.869 (p=0.001)

Note: *p<0.1; **p<0.05; ***p<0.01

	2020-2021
const	1.231*** (0.235)
logMktCap	-0.308*** (0.074)
P/B	0.014 (0.023)
EBIT	-2.162*** (0.533)
ESGScore	-0.003 (0.004)
No. Observations	126
R-squared	0.261
Number of firms	68
F Statistic (Robust)	8.292 (p=0.000)

Note: *p<0.1; **p<0.05; ***p<0.01

Note: Left regression uses Sharpe ratio as the output variable, and the table on the right uses return volatility.

Table 8 substitutes individual ESG scores for the composite ESG score and applies the models again to the COVID-19 period (2020-2021). Our risk-adjusted return model explained 31% of the variation in Sharpe ratios, with size and operational efficiency again significant determinants. Beta showed weak significance at the 10% level, and was associated with lower risk-adjusted return, i.e., firms more sensitive to systematic risk performed worse. The model significance was high as indicated by the robust F-statistic of 4,9 ($p < 0,00$).

After applying our second model on the COVID-19 period, with return volatility as the dependent variable, size and operational efficiency exhibited significant resilient effects on return volatility. The ESG score and relative valuation provided no significant impact on volatility. The overall model is robust with an F-statistic of 8,3 and explained 26,1% of the variation in annualized volatility within firms.

Overall, the COVID-19 regressions using the composite ESG score confirm that firm size and operational efficiency were the main drivers of both risk-adjusted performance and volatility, while ESG and valuation metrics remained insignificant. Beta showed weak explanatory power for Sharpe ratios. Both models were statistically robust and explained a meaningful portion of within-firm variation, reinforcing the greater utility of fundamental financial characteristics over ESG scores in periods of market stress.

Figure 5

Regression Residual Plot: Risk-Adjusted Return (2016-2023)

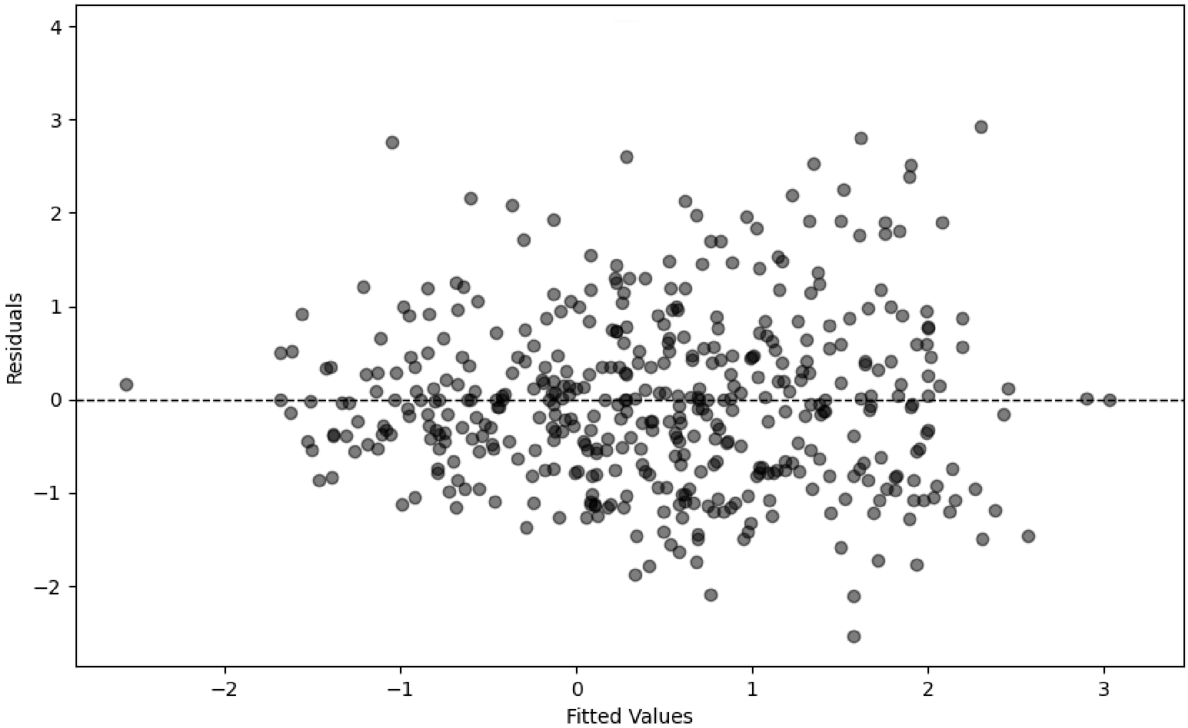


Figure 5 presents residuals plotted for the full regression period (2016-2023). The residuals appear to be generally centered around zero and exhibit no clear pattern, supporting the linearity assumption. The spread is relatively constant, but a slight skewness towards the right side may suggest mild heteroskedasticity. As firm-clustered standard errors are employed, these issues should be properly accounted for.

7 Discussion

This section interprets the empirical findings in relation to prior research on ESG and firm performance. We evaluate the statistical significance of key variables, compare our results to existing studies, and highlight the stronger explanatory power of traditional financial metrics relative to ESG scores. Finally, we address limitations concerning data coverage, model specification, and the generalizability of our results.

7.1 Results

Across all regressions, firm size (logMktCap) and operational efficiency (EBIT margin) provided the greatest explanatory power with regards to both risk-adjusted returns and return volatility within the A&D industry. The results indicate firm size and efficiency are positive factors in performance, which confirms previous results from earlier research (Boubaker et al., 2020; Zhang, 2025). These findings contradict Gavrilakis and Floros (2023) and Borovkova and Wu's (2020) results, that larger firms are associated with lower stock market return. However, Borovkova and Wu's (2020) used raw stock return as a proxy for financial performance, and smaller firms tend to exhibit greater volatility (Stockdale, 2014), a factor that Sharpe ratio accounts for. This held across most time periods, including the COVID-19 period, and was also associated with lower return volatility, indicating higher market resilience among larger and more efficient A&D firms, which aligns with results from Wang et al. (2025) and Demer et al. (2021). The period between 2021-2023 proved to be an inefficient period, as only relative valuation showed weak significance with regards to returns. This could possibly be explained by the post-COVID recovery phase, where markets may have priced in expectations and policy signals more aggressively, reducing the explanatory power of firm-level fundamentals. As noted by Wang et al. (2025), ESG-driven effects tend to peak during periods of acute crisis but fade in the recovery phase, where speculative sentiment, macroeconomic policy shifts, and sector-specific shocks dominate price movements.

Price-to-book ratio (P/B) showed mixed results. Over the pre-COVID period (2016-2019), and the full sample period (2016-2023), we found evidence that relative valuation had a significant negative effect on risk-adjusted returns. This suggests that firms with higher valuation suffered in terms of having weaker returns relative to volatility. Similar results were found in Broadstock et al. (2021), while Gavrilakis and Floros (2023) found a significant positive relationship. The mixed results are thus relatively consistent with evidence from previous research. The effect of relative valuation could not be determined to have a significant effect during and post-COVID. Market sensitivity (Beta) was not statistically significant in explaining returns or volatility in any model, other than weak significance over the COVID-19 period in determining Sharpe ratios. This suggests that systematic market risk does not meaningfully drive performance in the A&D sector, when controlling for ESG and firm-specific characteristics. This result is corroborated by Hong and Kacperczyk (2009), who reference the persistent insignificance of Beta in previous research.

Regarding ESG, our findings are relatively one-sided. Individual ESG pillars were not found to be statistically significant determinants of risk-adjusted performance or return volatility in any subperiod regressions. However, the composite ESG score showed a significant negative relationship with Sharpe ratio over the full sample period at the 5% significance level. In terms of our defined hypotheses, this result would mean a rejection of the first hypothesis, which assumes the effect of ESG on risk-adjusted returns within the A&D sector is zero. We have sufficient evidence to say that, over this period, ESG had an effect. These results are in accordance with other research (Borovkova & Wu, 2020; Gavrilakis & Floros, 2023), which found empirical results that better ESG performance decreases stock market return. Comparatively, Kroll (2023) demonstrated that the highest ESG rated portfolios outperformed portfolios with lower ESG score, measured in raw stock market return, in 10 out of 12 regions, but provided no further explanation. We cannot, however, reject the second null hypothesis, which assumes that ESG does not have an effect on return volatility. We do not have enough evidence to meaningfully conclude that the effect was not, in fact, zero.

The application of our findings in a theoretical context, involves challenging the view that higher ESG performance corresponds to higher risk-adjusted returns. Stakeholder theory posits that responsible firms create long-term value by reducing reputational, regulatory and operational risks (Freeman, 1984), which helps strengthen internal cohesion and reduce financing costs (Wang, 2024). As evidenced in this thesis, the negative relationship between ESG scores and market performance, and the difficulties in establishing a significant effect, thus challenges this standpoint. An explanation for this contradictory relationship may be that ESG factors are mispriced or less important in sin sectors. Gavrilakis and Floros (2023) found similar results and offered the possible explanation that the negative effect between ESG and financial performance is likely due to sustainability being a lower priority in different regions and sectors. This reasoning could also apply to the findings in this study, specifically, the negative relationship between ESG performance and risk-adjusted stock market return in the A&D sector.

Additionally, for investors seeking greater market resilience through ESG investing, our findings yet again show limited utility. Several studies have found varying results regarding market resilience and ESG scores, depending on region, industry and time period. However, our results from the A&D sector do not support this view. During the COVID-19 pandemic period, ESG scores were not found to have a significant effect on stock return volatility. We offer the possible explanation that, within the A&D sector, firm-level ESG performance may have played less of an important role relative to broader systemic and geopolitical concerns that dominated investor decision-making during this time. Furthermore, if ESG effects are primarily realized through reputational or stakeholder-driven channels, these mechanisms may have been less relevant in the A&D sector.

A possible practical implication for investors seeking superior risk-adjusted returns or market resilience, is that ESG considerations may offer limited value as a screening tool. Investors may find greater utility in focusing on traditional financial indicators, such as size or efficiency. A similar conclusion is also made by Revelli and Viviani (2015), and La Torre et al. (2020). This perspective is easily challenged, however, by proponents of sustainable investing offering the explanation that ESG rating agencies are unreliable and their methodologies may not be sufficient (Kroll, 2023).

7.2 Limitations

While this study contributes to filling the research gap in today's literature by examining ESG ratings impact on risk-adjusted stock market return within the A&D industry, there are certain limitations that need to be acknowledged. The data availability within the A&D sector was limited and there were discrepancies between data providers, since we extracted historical price data from *yfinance* and ESG ratings from LSEG, hence constraining some firms to be included in the sample. The sampled period between 2016-2023 is sufficient but extending the time period would provide more robustness to our results. However, the absence of data for both ESG ratings and historical price data pre 2016 were limited. Some firms in the A&D sector that were included in the sample, are not purely defense, i.e., some firms could be purely aerospace. This could influence the results since both financial and ESG performance may differ between pure defense and aerospace firms.

To calculate Sharpe ratio, excess return was estimated by taking the annual return for every stock and subtracting it with a 10-year government bond yield dependent on its region. However, we only included bond yields from six different regions: US, Europe, China, Canada and Japan. Firms that were included in our sample but did not pertain to any of those regions used the “other” bond yield, which was the average yield for the six countries. The rationality of only utilizing government bonds from those six regions was due to limited data availability for firms operating in other regions, which may limit this thesis.

Another limitation is the absence of a universal methodology for ESG ratings, which may result in different ratings for the same firm between different agencies (Dimson et al., 2020). This study utilized ESG ratings from LSEG but the non-universal methodology weakens the credibility for the metric. This approach could potentially lead to different perceptions of firms ESG performance and empirical results depending on which agency that is being used.

8 Conclusion

This section summarizes the key findings from our analysis of ESG scores and financial performance in the Aerospace and Defense sector. ESG scores exhibited limited explanatory power compared to traditional metrics like firm size and efficiency. We reflect on the implications for ESG integration in controversial industries, reiterate key methodological limitations, and suggest future research on ESG rating divergence and underexplored sectors.

This thesis investigates the relationship between ESG performance and stock market performance. Specifically, we employ a panel regression approach over the period 2016-2023, and assess the explanatory power of both individual ESG pillar scores and composite scores on risk-adjusted returns and return volatility in the Aerospace and Defense (A&D) sector. Our findings suggest that ESG scores do not consistently explain variation in risk-adjusted returns or return volatility in the A&D sector. The only exception is a negative and statistically significant coefficient for the ESG composite score over the full sample period, providing evidence that higher ESG performance corresponds to lower risk-adjusted returns in certain industries. No significant relationship was found relating to ESG performance having a stabilizing force on return during large market fluctuations, as both individual pillar scores and the composite score showed insignificance.

The implications of our results largely challenge the concept of ESG performance translating to superior market performance. Our findings contribute to the growing body of research highlighting the difficulty in establishing a consistent and significant relationship between ESG and firm-level financial outcomes. Moreover, by isolating the A&D sector, a controversial industry often excluded from ESG analysis, it also adds to the limited empirical evidence on how ESG behaves in controversial sectors. Investors, ESG data providers, and policymakers may all benefit from our findings by re-evaluating how ESG metrics are interpreted and used in controversial industries similar to A&D.

We also recognize the inconsistencies in ESG ratings across different data providers to be a primary limiting factor in the comparability and reliability of ESG scores. Low correlation between providers undermines the validity of the interpretation, and we encourage future research to examine the implications of ESG rating divergence and how it affects investment decisions and financial outcomes. Comparison of the predictive power of different data providers may also be an interesting topic for further exploration. Finally, expanding this line of research into other controversial sectors would strengthen the empirical foundation and contribute to a broader understanding of how ESG functions within these industries.

References

Amini, S., Delgado, M. S., Henderson, D. J., Parmeter, C. F., Carter Hill, R., Newey, W. K., Baltagi, B. H., & White, H. L. (2012). Fixed vs Random: The Hausman Test Four Decades Later. In *Essays in Honor of Jerry Hausman (Vol. 29, pp. 479–513).* Emerald Group Publishing Limited. https://doi.org/10.1108/S0731-9053(2012)0000029021

Barth, M. E., Cahan, S. F., Chen, L., & Venter, E. R. (2017). The economic consequences associated with integrated report quality: Capital market and real effects. *Accounting, Organizations and Society, 62,* 43–64. https://doi.org/10.1016/j.aos.2017.08.005

Bauer, R., Pavlov, B., & Schotman, P. (2004). Panel data models for stock returns: the importance of industries. *Working paper.*

Bernhardt, D. & Eckblad, M. "Stock Market Crash of 1987," *Federal Reserve History,* October 22, 2013, https://www.federalreservehistory.org/essays/stock-market-crash-of-1987.

Bhatia, V., Kumar, S. (2024). The Impact of ESG Investing on Portfolio Performance: An Empirical Study of Emerging Markets. *Journal of Informatics Education and Research,* 4(3), 2024. 3314-3325. https://doi.org/10.52783/jier.v4i3.1916

Bodie, Z., Kane, A., & Marcus, A. J. (2021). *Investments* (Twelfth edition, International student edition.). McGraw-Hill Education.

Borovkova, S., & Wu, Y. (2020). ESG versus financial performance of large cap firms: The case of EU, U.S., Australia and Southeast Asia. *Probability & Partners.* https://probability.nl/wp-content/uploads/2020/08/Refinitive_ESG_Analysis_in_4_Regions.pdf

Boubaker, S., Le, T. D. Q., Manita, R., & Ngo, T. (2023). The trade-off frontier for ESG and Sharpe ratio: a bootstrapped double-frontier data envelopment analysis. *Annals of Operations Research*, *347*(1), 717–741. https://doi.org/10.1007/s10479-023-05506-z

Broadstock, D. C., Chan, K., Cheng, L. T. W., & Wang, X. (2021). The role of ESG performance during times of financial crisis: Evidence from COVID-19 in China. *Finance Research Letters, 38*, 101716. https://doi.org/10.1016/j.frl.2020.101716

Cardillo, G., Bendinelli, E., & Torluccio, G. (2023). COVID-19, ESG investing, and the resilience of more sustainable stocks: Evidence from European firms. *Business Strategy and the Environment, 32(1),* 602–623. https://doi.org/10.1002/bse.3163

Demers, E., Hendrikse, J., Joos, P., & Lev, B. (2021). ESG did not immunize stocks during the COVID-19 crisis, but investments in intangible assets did. *Journal of Business Finance & Accounting, 48(3–4),* 433–462. https://doi.org/10.1111/jbfa.12523

Dimson, E., Marsh, P., & Staunton, M. (2020). Divergent ESG ratings. https://doi.org/10.17863/CAM.55949

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

Engelhardt, N., Ekkenga, J., & Posch, P. (2021). ESG Ratings and Stock Performance during the COVID-19 Crisis. *Sustainability, 13(13),* 7133. https://doi.org/10.3390/su13137133

European Commission. (2025, April 30). *Commission mobilises €910 million to boost European defence and close capability gaps* [Press release]. https://ec.europa.eu/commission/presscorner/detail/en/ip_25_1116

European Commission. (2025, April 10). *Commission presents first European Defence Industrial Strategy and proposes a regulation to strengthen the European defence industry through common procurement* [Press release]. https://ec.europa.eu/commission/presscorner/detail/en/ip_25_793

Fama, E. F., & French, K. R. (2004). The Capital Asset Pricing Model: Theory and Evidence. *The Journal of Economic Perspectives*, *18*(3), 25–46. https://doi.org/10.1257/0895330042162430

Fama, E. F., & French, K. R. (1992). The Cross-Section of Expected Stock Returns. *The Journal of Finance (New York), 47(2),* 427–465. https://doi.org/10.1111/j.1540-6261.1992.tb04398.x

Fama, E. F. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *The Journal of Finance, 25(2),* 383–417. https://doi.org/10.2307/2325486

Freeman, R.E., (1984). *Strategic Management: A Stakeholder Approach.* Boston: Pitman.

Friede, G., Busch, T., & Bassen, A. (2015). ESG and financial performance: aggregated evidence from more than 2000 empirical studies. *Journal of Sustainable Finance & Investment, 5(4),* 210–233. https://doi.org/10.1080/20430795.2015.1118917

Friedman, M. (1970, September 13). The social responsibility of business is to increase its profits. *The New York Times Magazine.* https://www.nytimes.com/1970/09/13/archives/a-friedman-doctrine-the-social-responsibility-of-business-is-to.html

Gavrilakis, N., & Floros, C. (2023). ESG performance, herding behavior and stock market returns: evidence from Europe. *Operational Research, 23(1),* 3. https://doi.org/10.1007/s12351-023-00745-1

Ghitti, M., Gianfrate, G. & Palma, L. The agency of greenwashing. *J Manag Gov 28,* 905–941 (2024). https://doi.org/10.1007/s10997-023-09683-8

Hong, H., & Kacperczyk, M. (2009). The price of sin: The effects of social norms on markets. *Journal of Financial Economics, 93(1),* 15–36. https://doi.org/10.1016/j.jfineco.2008.09.001

Iannone, B., Duttilo, P., & Gattone, S. A. (2025). Evaluating the Resilience of ESG Investments in European Markets During Turmoil Periods. *Corporate Social-Responsibility and Environmental Management.* https://doi.org/10.1002/csr.3225

Iskenderoglu, C. (2025). The value of diversification: ESG and investment in controversial industries. *Finance Research Letters, 76,* 106956. https://doi.org/10.1016/j.frl.2025.106956

Jensen, M. C. (1989). Eclipse of the Public Corporation. In *Harvard business review* (Vol. 67, Number 5, p. 61). Harvard Business Review. https://doi.org/10.2139/ssrn.146149

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47(2),* 263–291. https://doi.org/10.2307/1914185

Kräussl, R., Oladiran, T., & Stefanova, D. (2024). A review on ESG investing: Investors' expectations, beliefs and perceptions. *Journal of EconomicSurveys, 38,* 476–502. https://doi.org/10.1111/joes.12599

La Torre, M., Mango, F., Cafaro, A., & Leo, S. (2020). Does the ESG Index Affect Stock Return? Evidence from the Eurostoxx50. *Sustainability*, *12*(16), 6387. https://doi.org/10.3390/su12166387

Laufer, W. S. (2003). Social Accountability and Corporate Greenwashing. *Journal of Business Ethics, 43(3)*, 253–261. https://doi.org/10.1023/a:1022962719299

Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *The review of economics and statistics, 47(1),* 13–37. https://doi.org/10.2307/1924119

London Stock Exchange Group. (n.d.). About us. LSEG. https://www.lseg.com/en/about-us

Mossin, J. (1966). Equilibrium in a capital asset market. *Econometrica, 34(4),* 768–783. https://doi.org/10.2307/1910098

Nunes, C., & Recine, J. (2023). *ESG and global investor returns study.* Kroll. https://www.kroll.com/en/insights/publications/cost-of-capital/esg-global-investor-returns-study

Paradis, G., & Schiehll, E. (2021). ESG Outcasts: Study of the ESG Performance of Sin Stocks. *Sustainability, 13(17),* 9556. https://doi.org/10.3390/su13179556

Petersen, M. A. (2009). Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches. *The Review of Financial Studies*, *22*(1), 435–480. https://doi.org/10.1093/rfs/hhn053

Pollman, E. (2022). The making and meaning of ESG. *U of Penn, Inst for Law & Econ Research Paper No.* 22-23. SSRN. https://ssrn.com/abstract=4219857

Revelli, C., & Viviani, J.-L. (2015). Financial performance of socially responsible investing (SRI): what have we learned? A meta-analysis. *Business Ethics (Oxford, England)*, *24*(2), 158–185. https://doi.org/10.1111/beer.12076

Sandu, D.M. (2024). The Effect of ESG Scores on Portfolio Performance. Evidence from Europe. *Proceedings of the International Conference on Business Excellence, 18(1),* 2024. 442-450. https://doi.org/10.2478/picbe-2024-0038

Sharpe, W. F. (1964). Capital asset prices: a theory of market equilibrium under condition of risk. *The Journal of finance (New York), 19(3),* 425–442. https://doi.org/10.2307/2977928

Sharpe, W. F. (1966). Mutual Fund Performance. T*he Journal of Business (Chicago, Ill.), 39(1),* 119–138. https://doi.org/10.1086/294846

S&P Global. (2019). Understanding the “E” in ESG. S&P Global. https://www.spglobal.com/en/research-insights/market-insights/understanding-the-e-in-esg

S&P Global. (2020). What is the “G” in ESG? S&P Global. https://www.spglobal.com/en/research-insights/market-insights/what-is-the-g-in-esg

S&P Global. (2020). What is the “S” in ESG? S&P Global. https://www.spglobal.com/en/research-insights/market-insights/what-is-the-s-in-esg

Stockdale, John. (2014). On the Volatility of Very Small Exchange-Traded Operating Companies. http://dx.doi.org/10.2139/ssrn.2409130

Vissing-Jorgensen, A. (2003). Perspectives on Behavioral Finance: Does "Irrationality" Disappear with Wealth? Evidence from Expectations and Actions. *NBER Macroeconomics Annual*, *18*, 139–194. https://doi.org/10.1086/ma.18.3585252

Wang, C. (2024). The Relationship between ESG Performance and Corporate Performance - Based on Stakeholder Theory. *SHS Web of Conferences, 190,* 3022. https://doi.org/10.1051/shsconf/202419003022

Wang, S., Cheng, H. C., Wang, J., & Yick, H. Y. (2025). The performance of ESG portfolios: Evidence from the Chinese market under COVID-19. *Economic Modelling, 143,* 106958. https://doi.org/10.1016/j.econmod.2024.106958

Whelan, T., Atz, U., Van Holt, T., & Clark, C. (2021). ESG and Financial Performance: Uncovering the Relationship by Aggregating Evidence from 1,000 Plus Studies Published between 2015-2020. *NYU Stern Center for Sustainable Business and Rockefeller Asset Management.* *https://www.stern.nyu.edu/sites/default/files/assets/documents/NYU-RAM_ESG-Paper_2021%20Rev_0.pdf*

Wille, J. H., Förster, N., Keller, A., Kutschera, H.-J., & Reiff, N. (2024). Defending the future: The evolving role of ESG in the defense industry amid the war in Ukraine. Strategy& (PwC). https://www.strategyand.pwc.com/de/en/industries/aerospace-defense/evolving-role-esg-in-defense-industry.html

Yahoo Finance. (n.d.). *Yahoo Finance.* https://finance.yahoo.com

Yin, X.-N., Li, J.-P., & Su, C.-W. (2023). How does ESG performance affect stock returns? Empirical evidence from listed companies in China. *Heliyon, 9(5)*, e16320–e16320. https://doi.org/10.1016/j.heliyon.2023.e16320

Zhang., L-S., (2025). The impact of ESG performance on the financial performance of companies: evidence from China's Shanghai and Shenzhen A-share listed companies. *Frontiers in Environmental Science*, *13*. https://doi.org/10.3389/fenvs.2025.1507151

Generative AI Usage

AI tools were used in a limited and responsible manner in the creation of this thesis. ChatGPT was used for identifying weaknesses in our texts, providing suggestions for grammar improvements, clarity and structure. These suggestions were only implemented by reworking our original texts and never used verbatim. For managing references, we utilized NotebookLM, which simplifies the process, but never replaces original content. For programming purposes, Microsoft Copilot was used to identify and resolve coding issues, such as syntax errors and poorly functioning code, but never plagiarised in any way.