TALLINN UNIVERSITY OF TECHNOLOGY
Department of Economics and Finance

Mikael Haitin

# ESG, EARNINGS, AND EARTHQUAKES: A SKEPTICAL ANALYSIS OF RESILIENCE IN MACROECONOMIC CHAOS

Bachelor's thesis

Curriculum: Applied Economics

Supervisor: Kaido Kepp

Tallinn 2025

I declare that I have prepared this thesis independently
and I have referred to all the works, important
positions and data of other authors used in its
preparation, and that I have not previously
submitted the same work for credit.

The length of the paper is 10,168 words from the introduction to the end of the summary.

Mikael Haitin ……………………………

(15/05/2025)

# TABLE OF CONTENTS

# ABSTRACT

This thesis investigates the relationship between Environmental, Social, and Governance (ESG) performance and corporate resilience during macroeconomic shocks. Building upon stakeholder theory and employing a skeptical empirical approach, the study addresses critical methodological concerns often overlooked in ESG literature – namely, endogeneity, rater bias, and flow-driven performance concerns. Using a panel dataset of 577 U.S. firms from the Russell 3000 index and quarterly observations spanning 2017 to 2024, the analysis applies a System Generalised Method of Moments (GMM) framework to assess firm profitability, proxied by EBITDA margins, in response to the COVID-19 pandemic and the Russian invasion of Ukraine. ESG scores are normalised to capture relative performance, while firm-specific controls and interaction terms account for exogenous shocks. Results reveal limited effects: while ESG variables show no significant protective effect during the pandemic, environmental performance positively correlates with profitability following the invasion, whereas social performance has an adverse impact. Governance, contrary to prevailing literature, is found to be immaterial in both periods. These findings contrast with conventional narratives around ESG benefits, underscoring the need for a more rigorous approach to modelling techniques.

Key words: ESG, macro-shocks, GMM, profitability, resilience.

# INTRODUCTION

Environmental, Social and Governance (ESG) has been subject to numerous research studies in the past two decades, with over 4,000 papers written on the topic (Clément, Robinot, & Trespeuch, 2025). There is evidence of a growing interest from investors to consider each of the ESG letters. Environmental aspects are given increasingly more attention from investors, as governments push for environmental sustainability, and climate-conscious companies with better climate-risk disclosures are valued higher (Ilhan, Krueger, Sautner, & Starks, 2023). Socially responsible companies are gaining spotlight with more examples of how corporate social responsibility (CSR) practices can increase economic value through reducing costs associated with production disruptions, consumer boycotts, lawsuits, and bankruptcy (Nofsinger, Sulaeman, & Varma, 2019). Governance issues are continuously discussed, especially with large-scale scandals, including Enron (Li Y. , 2010) and Wirecard (Heese, Wang, & Labruyere, 2021), providing extreme examples of the consequences of poor governance mechanisms.

Indeed, ESG has gained a lot of attention from investors, the public, and academics, whereas the influence of ESG on different company performance measures, like share prices, earnings, risks, and company value, has been studied especially thoroughly (Chytis, Eriotis, & Mitroulia, 2023). The pace has only increased in recent years, as the COVID pandemic and the full-scale Russian invasion of Ukraine introduced additional data points for research. The following European energy crisis further contributed to reigniting environmental concerns, as the need for new energy sources was underscored by worsened relationships between Russia and Western economies (Youngs & Goldthau, 2023).

While ESG studies vary, the majority of papers have found the relationship to be positive. Whelan and others (2021) performed a meta-analysis of over 1,000 research papers. After grouping the articles into those focused on financial metrics (including ROE, ROA, and stock performance) and investment metrics (including Jensen’s alpha, Sharpe ratio, and others), they show that 71% and 59% of studies in the respective groups found a neutral or positive relationship between ESG and company or stock performance. The meta-analysis has also revealed the asymmetric benefits of

ESG investing during economic downturns, as ESG stocks outperformed their peers during both the financial crisis of 2008 and the COVID pandemic, but not during non-crisis periods.

Still, the benefits of ESG investing as compared to conventional investing strategies remain a highly debated topic. Van der Beck (2024) shows how excessive returns from ESG investing are, in fact, driven by pressure from ESG flows rather than being performance-driven. In the absence of those flows, ESG strategies would have underperformed non-ESG strategies, and the possibility of outflows presents a risk of negative realized returns in the future. The question of the underlying rationale for these flows remains open, however, as they could stem from both investor taste and previous performance. It is certain that, to at least some extent – and likely to a significant one – these flows are performance-driven, as purely taste-based ESG flows would be expected to be negative and yield lower returns during periods of economic crisis, contrary to the empirical findings discussed earlier.

Some argue ESG strategies should underperform because investors are trading returns for the moral utility of sustainable investing (van der Beck, 2024). Another perspective, but with a similar conclusion, is to consider “sin” stocks as having excessive returns due to societal shunning and hence a higher cost of capital (Hong & Kacperczyk, 2009). Hart and Zingales (2022) touch on a similar issue and suggest that with governments failing to internalize externalities, such as pollution or unfair labour practices, corporations should act in line with maximizing shareholder welfare, not profit. Thus, a shareholder welfare maximization function (SWM) would become dominant and allow for a trade-off between profit and sustainable practices, suggesting that the trade-off must exist. On the other hand, the majority of research in the field of ESG is based upon either institutional or stakeholder theories (Li, Wang, Sueyoshi, & Wang, 2021) that oppose the classical shareholder view and the aim of profit maximization, and instead propose that firms should consider their ethical duties and gain legitimacy in the eyes of society.

Similarly to numerous predecessors, this thesis builds upon stakeholder theory and aims to provide empirical evidence on the effects of ESG integration on company performance during economic crises. Recent economic developments provide several interesting data points that had a material impact on all companies. As such, the following research questions are proposed:

1. Is relative ESG performance, measured through ESG scores, associated with better firm performance?

2. Does relative ESG performance, measured through ESG scores, provide protection to firms when exposed to external shocks?
3. Is there a significant relationship between relative ESG scores and firm performance, and if there is, which aspects have the most effect, and in what direction?

Although the formulated questions are quite standard in research, this thesis takes a skeptical approach to modelling the relationship. Namely, relative ESG scores are used to account for potential rater bias, discussed in the following chapter. Firm performance is measured using an accounting metric to avoid the concerns of capital inflow-driven outperformance of ESG companies. Finally, a Generalised Method of Moments (GMM) is adopted as the primary analysis tool with the aim of considering reverse causality between ESG and performance, as discussed in Chapter 3.1. To answer these questions, a system GMM was performed on a sub-sample of 443 U.S. companies from the iShares Russell 3000 ETF for two periods: from 2018 to 2021 to include the COVID-19 pandemic, and from 2021 to 2024 to include the Russian invasion of Ukraine.

The thesis starts by describing a theoretical framework for the study and defining relevant terminology. Theoretical reasoning is provided for the expected relationship, and a generalised function of company performance is formulated. An overview of previous empirical literature on the topic follows. Chapter two describes the methodology for the analysis, the data used in the analysis and the various processing techniques applied to prepare the data. Chapter three provides reviews the results of the analysis.

Bibliography was created using built-in MS Word function „Refernces“ with APA 6th edition style.

# 1. THEORETICAL FRAMEWORK AND THE FUNCTION OF FIRM PERFORMANCE

The topic of this thesis is rather controversial, and as such, it requires consideration of the underlying theoretical reasoning for why ESG could affect corporate performance and to what extent, as well as several definitions for the terminology used. The current chapter presents a theoretical framework on which this thesis is built. Starting with an overview of ESG and firm theories that provide justification for the following empirical analysis, the chapter continues to define a function of company performance, drawing from economic theory. At the end, empirical evidence for the proposed relationship is provided from previous research on the topic.

## 1.1. Goals, Sustainability and Performance of an Enterprise

First, it is important to understand what ESG is. The history of ESG starts from socially responsible investing, or SRI (also referred to as ethical investing, impact investing, sustainable investing, and others), which in itself is far from a novel idea, with roots dating back to the late 18th century. However, the concept started gaining popularity only in the 1970s and 1980s. The original idea of SRI was mainly based on “screening,” that is excluding investment in companies engaged in the manufacturing of tobacco or alcohol products, providing military equipment, running casinos, and other unethical businesses, often generalized by the term „sin stocks“ (Townsend, 2020).

Following SRI, the concept of corporate social responsibility was later born to extend the idea of screening into encouraging all corporations to actively take actions toward improving the well-being of society. Although vague, CSR could be defined as a form of self-regulated framework for companies to address the social, environmental, and economic concerns of their actions (Sheehy, 2015). To help promote and integrate CSR into corporate actions, the International Organization for Standardization (2010) published ISO 26000:2010 – a document providing voluntary guidance on the principles of sustainable development for all types of organisations internationally.

Unlike SRI – an investment strategy, and CSR – a voluntary framework for actions, ESG lacks a precise definition. ESG is an acronym for Environmental, Social, and Governance, coined in 2004 by a cohort of over 20 bank executives and the United Nations in their report Who Cares Wins, to conveniently refer to a set of problems, they argued should be considered in investment analysis. However, the report did not give a specific definition, and over time the acronym became extremely unclear, simultaneously being used to describe a set of criteria, risks, standards, investment practices, strategies, activities, and much more. The ambiguity of the term has even led to it being weaponized by both "far left" and "far right" politicians, inviting additional backlash and debate over the topic (Pollman, 2024).

Arguably the biggest challenge to ESG adoption that stems from the lack of a concrete definition, and hence standardization, is the uncertainty and variability in ESG evaluation and scoring, presenting challenges to investors and researchers. The majority of the studies that use ESG scores rely on rating providers like Bloomberg, MSCI, Refinitiv, Sustainalytics, and others (Senadheera, Gregory, Rinklebe, Farrukh, & Rhee, 2022). However, these scores are open to subjective bias due to measurement, scope, and weight divergence. It was also found that measurement is likely the main driver of rating divergence, partially driven by the *rater effect* – meaning that companies that receive a high score in one category are more likely to receive higher scores in other categories as well (Berg, Kölbel, & Rigobon, 2019). This very issue is addressed in this research in Chapter 2.1.

Despite its long history, sustainable investing remains ambiguous – especially in how ESG is defined. For the sake of consistency and clarity, this study defines ESG as *a set of criteria used to evaluate company performance in relation to environmental sustainability, social responsibility, and corporate governance*. That means companies acting in line with ESG aim to improve on the aforementioned criteria, and consequently, both SRI- and CSR-motivated practices are included in this definition. It remains unclear still why a profit-maximising company should care about ESG – a seemingly philanthropic activity – at all. However, theory doesn't necessarily contradict profit-maximisation goals with socially ethical actions, and theories on firm goals differ largely.

Shareholder theory, also known as the Friedman doctrine, states that the social responsibility of each company is to increase shareholder value, while the interests of other groups – suppliers, workers, consumers, or anyone else – should never come before the interests of the shareholders (Friedman, 1962). Shareholder theory holds that a company's only goal is to maximize profits for its owners, which is based on shareholder primacy (Smith, 1998). Critics of shareholder primacy challenge the idea that shareholders are the owners of the company (Stout, 2002), and even call the doctrine overrated altogether (Smith, 1998), for over 40 years the shareholder theory remained the dominant firm theory and the main approach to viewing companies' goals. However, the financial crisis of 2008 was a major turning point that shifted focus to stakeholder theory instead and made numerous economists and company CEOs reconsider their views (Tse, 2011).

Stakeholder theory, on the other hand, argues that firms should consider the interests of all the company's stakeholders and not just the owners (Freeman R. E., 1984). Unlike the Friedman doctrine, the stakeholder approach is much more fragmented, and the definition remains relatively vague even today. The views differ in both the rationale for stakeholder-oriented actions as well as the definition of a stakeholder itself. However, the main theme of the theory remains largely the same: *organizations should look beyond shareholder interests to acknowledge and manage the needs and demands of all stakeholders, both internally and externally, in order to ensure long-term value maximization* (Mahajan, Lim, Sareen, Kumar, & Panwar, 2023).

As previously mentioned, stakeholder theory is one of the dominant theoretical bases in the field of ESG, since both views share the same ideas to a major extent. In fact, Who Cares Wins was written upon the theme of the interdependency between a firm's success and the well-being of society and a sustainable planet, also aiming to encourage firms to consider ESG issues not merely for risk mitigation purposes, but also to secure brand reputation and create long-term value (United Nations Global Compact, 2004). And while the report did not explicitly refer to the stakeholder view, it pushed forward the same core ideas nonetheless.

Theoretically, it is argued that businesses ensuring the happiness of their employees, building a good reputation in the eyes of their customers, and creating strong relationships with suppliers – all of which also align with ESG-based practices – should have improved market performance, as they should be more efficient, exhibit higher customer stickiness, and be more likely to receive discounts (Freeman R. E., 1994). The instrumental stakeholder theory further expands on the economic benefits of ethical behaviour, suggesting that trust, cooperation, and honesty can reduce bonding and opportunism costs, as well as reduce internal free-riding risks – all contributing to competitive advantage. (Jones, 1995)

Several ways in which ESG-oriented actions can create value have been identified. More specifically, all three factors of ESG were found to have a positive causal effect on increased customer engagement and brand trust, with environment and society aspects contributing the most (Tripopsakul & Puriwat, 2022). The integration of ESG factors into corporate strategy was found to help improve efficiency by reducing production costs, lowering legal and political risks, increasing workforce productivity, and improving asset optimization (Egorov, 2023). While sound governance practices both reduce the risk of corporate failures and help improve the two other ESG criteria (O'Hare, 2022), thus simultaneously reducing risks and also having an indirect impact on the benefits from environmental and societal factors.

It can now be argued that companies incorporating ESG into their strategies can exhibit better performance, especially when exposed to exogenous macro-shocks, since their operations are subject to less uncertainty, that is, risk. To better understand this relationship, the next subsection discusses how enterprise performance is related to the macroeconomic climate, gives a definition of a macro-shock, and specifies the exact expected relationship between ESG and shock resilience.

## 1.2. Factors affecting firm performance

Macroeconomics is, in its essence, a study of the three phenomena of an economy: output, unemployment, and inflation, and the main aim of macroeconomics is to explain and characterise the movements and dynamics of these phenomena. In trying to achieve that, macroeconomics touches on a much wider range of topics, including money, interest rates, currency exchange rates, taxes, government spending, politics, and many more (Blanchard & Fischer, 1989). Various ways in which macroeconomic variables affect company performance have been identified, including relationships with monetary factors, fiscal factors and the overall economic climate.

With regards to monetary variables, Oxelheim (2003) shows that corporate performance is significantly affected by macroeconomic variables such as exchange rates, interest rates, inflation rates, and political risk, encouraging the inclusion of these factors into corporate reporting practices, as he sees that this information could be used by investors to make better decisions. Similarly, Abaidoo (2019) finds that inflation expectations, exchange rate volatility, inflation, and inflation uncertainty all have a unidirectional causal relationship to corporate performance volatility. Another study applied principal component analysis (PCA) on a sample of 116 UK-listed firms to find the relationship between return on assets (ROA) and numerous macroeconomic variables. With high explanatory power, the results showed that ROA is a function of prior period ROA and external macroeconomic variables, with key variables being the unemployment rate, real GDP, and exchange rates (Antwi & Issah, 2017).

Studies on fiscal factors, like governement expenditures and taxes have also found evidence for existing relationship between economy and firm performance. A study on Flemish municipalities found that both government spending and local taxes affect firm profitability. Unsurprisingly, the relationship was found to be positive with government spending and negative with taxes, and it was also shown that the effect of taxes is more significant than that of government spending, since, unlike expenditures, taxes affect company profitability directly (Schoenmaker, Cauwenberge, & Bauwhede, 2014). Another perspective on tax-related studies is with regards to the topic of tax risks – a theme closely tied to ESG – especially with governments increasingly getting involved in the problem of environmental deterioration through the implementation of different fiscal policies (Liu, Zhao, & Li, 2023). For example, Lin and others (2019), explored whether there is evidence of an association between CSR and tax risk. Their study showed how ethical behavior can lower tax expenses and encourage controlling for tax risks, thus improving company performance. There is also evidence that, due to uncertainty regarding regulatory policies, the risk of higher expenses for carbon-intensive companies is priced in the options market – meaning that downside protection is costlier for both environmentally unfriendly companies and investors (Ilhan, Sautner, & Vilkov, 2021).

From a perspective of an overall economic health, rather than a set of macroeconomic variables separately, the aggregate landscape plays a crucial role in enterprise performance. With consumers being at the center of every company, should their behavior change and sentiment turn pessimistic – either due to economic conditions or the political landscape (Kellstedt & Boef, 2004) – revenues will rapidly decrease. Measuring corporate performance using market-based metrics makes that relationship even more profound, as stock market returns are tied to at least industrial production, changes in risk premium, the yield curve, and inflation (Chen, Roll, & Ross, 1986).

With clear evidence of firm performance being dependent on macroeconomic variables presented, it is obvious that companies are open to the effects of macroeconomic shocks both directly and indirectly, through changes in macro-variables. It is now necessary to provide a concrete definition of a macro-shock in the context of this study. This thesis relies on a definition proposed by Ramey

(2016), namely that shocks are economically meaningful and uncorrelated primitive exogenous forces. Under Ramey's definition, a shock must satisfy three criteria:

1. Exogeneity with respect to other endogenous variables in the model,
2. Orthogonality with respect to other exogenous shocks,
3. Nature of movement – meaning unanticipated changes or news about future changes in variables.

Ramey further defines *narrative methods* as a way to identify shocks using historical records, policy documents, news reports, or expert judgments. These are shocks derived from human interpretation of real-world events, not statistically induced. A narrative shock should satisfy the same criteria of a broader shock, as outlined above. Additionally, it should be a documented event and narrative shocks are often modelled through dummy variables. This study considers macroeconomic shocks based on this framework and identifies two shocks to include in the model using narrative methods. Namely, the Covid-19 pandemic in 2020 and the unexpected invasion of Ukraine by Russia in 2022.

Connecting the previous discussion together, it can be formulated that the function of profitability of a company is dependent on company-specific factors, including ESG, as well as macroeconomic variables, or, more formally:

$$Prof = \beta_0 + \sum \beta_i (firm - specfic\ factors)_i + \sum \beta_j (macro - factors)_j + shock + \varepsilon \quad (1)$$

where
$Prof$ – is company profitability,
$firm - specfic\ factors$ – are factors specific to each company such as size, leverage, management, human capital, etc,
$macro - factors$ – are macroeconomic variables such as monetary factors, fiscal factors, overall economic health and others,
$shock$ – is a binary dummy,
$\varepsilon$ – is the error term.

It must be noted that, although macro-factors are, among everything else, influenced by the shock variable, in the proposed equation shocks still satisfy Ramey's criteria, as they remain primitive external events with unidirectional causality in relation to macro-factors. That means shocks are still external events, not caused by macroeconomic conditions, yet affecting them later. This, however, brings up *ex post* multicollinearity concerns between the two variables, potentially creating challenges in modelling the proposed relationship. Hansen and Wernerfelt (1989), have shown that economic factors affecting company profitability are roughly independent of firm-specific factors and that firm-specific variables have twice the explanatory power of economic variables. Following their findings, macro-factors could be excluded from equation (1), which would simultaneously achieve two things:

1. Address potential *ex post* collinearity concerns between the shock variable and macro-factors,
2. Amplify the effects of shocks by aggregating direct and indirect effects on company profitability under one variable.

With that a generalised function of company profitability can be formulated as:

$$Prof = \beta_0 + \sum \beta_i(firm - specific\ factors)_i + shock + \varepsilon \quad (2)$$

Under this definition, macro-factors are assumed to be exogenous to the model and become part of the error term. Assuming near-independence of macro-factors in relation to firm-specific factors, only limited, if any, omitted variable bias is introduced into the equation. While this definition might seem contradictory to the previous discussion on the relationship between macroeconomic variables and company performance, in the context of this research, it allows this study to specifically focus on the aggregate effects of shocks on profitability and thus better understand if ESG can provide protection against macroeconomic risks in general, that is, shocks become a proxy for systematic risk. This, in turn, necessitates systematic risk definition and invites an overall discussion on risk.

In a generally accepted definition, risk always entails two main concepts: exposure and uncertainty. Should any of these two concepts be absent, the event in question shall not be considered a risk (Holton, 2004). Companies are constantly exposed to a variety of risks that differ in their severity and likelihood. Unsurprisingly, it is important for managers and shareholders to know their exposure.

Numerous statistical risk measures exist, most common of which are standard deviation, Value-at-Risk (VaR) and Expected Shortfall (ES), which can be applied to both market-based and accounting performance indicators. From these measures, standard deviation or volatility captures the general level of uncertainty, while VaR and ES are related to the concept of tail-risk (Kou & Peng, 2016). All of those measures can also include either systematic or idiosyncratic risk components, depending on what indicator they are applied to. Systematic or market risk refers to a risk that simultaneously affects many market participants with high severity of losses and spreads throughout the system. Idiosyncratic or firm-specific risk is a type of risk that is different for each company, that is, uncorrelated between companies (Pérignon, Hurlin, Colliard, & Benoit, 2017), and stems from management decisions – for example, the amount of debt financing that a company uses.

In the context of this study, profitability is used as an indicator of company performance. Profitability is affected by both general market conditions as well as management decisions, and hence includes both systematic and idiosyncratic risk components. Thus, statistical measures of risk calculated from profitability reflect total risk associated with a company. When considering the effects of shocks on company profitability, exposure to systematic risk is implied – ESG, being a firm-specific factor, is therefore expected to limit that very exposure, for example through efficiency or brand loyalty, as previously discussed. Some empirical evidence from previous studies on the topic supporting and debating these conclusions is presented in the next and final subchapter of this section.

## 1.3. Previous empirical literature

The empirical research on the topic of ESG and firm performance has been very thorough – over 1,000 studies were published during the period from 2015 to 2020 alone. The majority of papers were able to find a positive relationship, measuring performance using both accounting and market-based metrics. In general, research indicates that ESG works best as part of a long-term corporate strategy, and plain disclosure of ESG performance is not associated with better returns (Whelan, Atz, Van Holt, & Clark, 2021).

Sustainable investing strategies have been found to be associated with better returns during economic or social crises as well, providing a certain degree of protection or resilience to adverse economic conditions. Fernández, et al. (2019) provided evidence that German green mutual funds were able to outperform their peers during the 2008 financial crisis. Another study arrived at similar conclusions in the case of London-listed FTSE stocks. SRI investment portfolios showed higher resilience in 2007–2009 and recovered their losses more quickly than conventional portfolios (Wu, Lodorfos, Dean, & Gioulmpaxiotis, 2015).

More recent research on the Covid-19 pandemic has further underlined the positive effects associated with ESG on company performance. Wang, Jiao, and Ma (2024) have analysed 4,527 Chinese A-share listed companies and found both a statistically significant and positive relationship between ESG-responsible performance and corporate resilience during the pandemic, using economic value added (EVA) as a measure of firm performance and including bankruptcy risk as a tail-risk metric. ESG has also been shown to provide resilience by reducing stock price volatility introduced by the pandemic, thereby having a stabilising effect on stocks in a crisis (Zhou & Zhou, 2021).

And while most of the empirical research shares these views, the debate remains. Bae, et al . (2021) analysed U.S. companies during the Covid pandemic and did not find a statistically significant relationship between CSR and stock market performance. Evidence was also presented by Demers, et al. (2021) that, after accounting for all the control variables in the regression, the effects of ESG ceased to be significant, bringing investments in intangible assets to the spotlight instead.

Van der Beck's (2024) findings about flow-driven ESG returns can also be mentioned as a counter argument, while Giakoumelou, et al. (2022) indirectly support those findings by showing how ESG scores serve as an indicator of managerial competence and risk aversion – attracting capital inflows and positioning ESG as a "safe haven" for investors during uncertain times, such as the 2008 financial crisis and the 2020 Covid pandemic.

In general, the reasearch indicates the existance of a positive and significant relationship between ESG and financial performance, but this view is not supported by all studies. This research analyses the same relationship and attempts to address the three problems often presented as reasons for positive relationships – flow-driven performance, rater-bias and endogeneity concerns. The first issue is addressed by using profitability as a measure of firm performance that is accounting-based, not stock price-derived. Secondly, Generalised Method of Moments is chosen as the analytical tool, which helps address endogeneity. Thirdly, relative ESG scores are included in the model to reduce the effects of rater-bias. A detailed discussion on the data used, methodology, analysis, and results is provided in the following chapters.

# 2. METHODOLOGY & DATA

## 2.1. Relationship specification and modelling approach

The relationship to be analysed is that of firm performance, ESG scores, and their interaction during macro-shocks. In Chapter 1.2, a generalised function of firm performance was presented in Equation (2). The firm-specific factors in that function can be broken down into ESG-related and other control variables to arrive at:

$$Prof = \beta_0 + \sum \beta_i (ESG)_i + \sum \beta_k (shocks)_k + \sum \beta_j (controls)_j + \varepsilon \tag{3}$$

where
$ESG$ – are E, S and G performance measures,
$controls$ – are firm-specific control variables,
$shocks$ – are macro-shocks defined using narrative methods.

Under this specification, firm profitability is used as a measure of company performance, and to consider potential bias of market-derived metrics – namely, flow-driven outperformance concerns – an accounting performance measure, EBITDA margins, is used. Profitability is defined as a function of environmental, social, and governance performance, other firm-specific variables, and external shocks that capture the direct effects on company operations, as well as indirect effects through changing the overall economic landscape. The error term includes unaccounted-for firm-specific controls, economic factors, and noise, whereas ESG measures are expected to be roughly orthogonal to the economic factors and likely correlated with the omitted controls, which is also the case for control variables included in the equation.

The suspected correlation between regressors and omitted variables is an endogeneity concern that can result in simultaneity bias, violating standard OLS assumptions and leading to underestimated standard errors and inadequate coefficients (Whited & Roberts, 2013). One of the main reasons for endogeneity is reverse causality – it might be that firms with better performance have more resources to invest in ESG-related activities, improving their ESG scores—not the other way around. Corporate and ESG studies are plagued by endogeneity concerns (Netter, Linck, & Wintoki, 2012), and pooled OLS, fixed effects, or random effects regressions are likely to produce biased and inconsistent coefficients. Furthermore, with standard regression techniques, unit roots are another important problem that needs to be considered, or otherwise risk spurious regression. Appendix 4 further elaborates on the problem of non-stationarity in the data used in this study.

The Generalised Method of Moments (GMM) addresses this issue by exploiting moment conditions based on valid instruments – variables that are correlated with endogenous regressors but uncorrelated with the error term. This is done through first-differencing the equation to remove unobserved individual fixed effects from the model, and then using lagged levels of the endogenous variables as instruments for the differenced equation. However, this approach eliminates any time-invariant variables, like shock dummies, from the model. System GMM is an extension that estimates two equations: the original equation in first differences and the equation in levels, instrumented with lagged differences of the endogenous variables (Arellano & Bover, 1995). This approach allows time-invariant variables to be kept in the model, which is crucial for this analysis, as shocks are included through binary dummies.

System GMM necessitates accurate specification of instrumented variables. The main idea is to instrument variables endogenous to the model; an additional supporting starting point is to test the variables for unit roots. As system GMM instruments variables with lagged first differences, the model not only allows for weak non-stationary processes but also implies the need to instrument them to avoid spurious regression. One final system GMM specific is that it works best on panels with small to moderate time dimensions $T$, or the model could risk overfitting (Ibid). The original dataset in this study has $T = 32$, usually considered large, and so two separate models were estimated with smaller periods, as described later.

After describing the methodological approach to the analysis, the following sub-chapter defines the specific variables to be included in modelling equation (3).

## 2.2. Variable definitions

As discussed in sub-chapter 1.1, using raw ESG scores is subject to potential *rater bias*, which is one of the concerns that this thesis aims to take into account. The best solution, also proposed by Berg et al. (2019), is to calculate average scores from several sources. However, due to data availability limits, that solution remains out of reach. This research follows the approach of Cornett et al. (2016) which indirectly helps address potential bias of score providers – min-max normalisation across entities using:

$$(X_{it} - \min(X_t)) \ / \ (\max(X_t) - \min\ (X_t)) \tag{4}$$

where

$X_i$ – is either E, S or G score.

This calculation results in relative ESG performance across firms, which is less subject to biased ratings, since if bias exists, it affects all the firms in a similar way and relative measures cancel it out – at least to some extent.

The next variable that needs definition is macro-shocks, which are commonly included in the model as dummy variables. In literature, dummy shock variables can have different definitions depending on the context, methodology, and other specifics. As this study aims to capture both instant (direct) effects and lagged (indirect) effects on operations, the chosen time windows for dummies were relatively long. For the COVID-19 pandemic, the dummy was defined for the first three quarters of 2020, as it was found that companies received most damage over quarters two and three, while the first quarter included instant effects (Bloom, Fletcher, & Yeh, 2021). For the Russian invasion, the dummy was defined for the whole year 2022, as economic consequences – including supply chain disruptions, trade restrictions, loss of confidence, higher energy prices, and many more – remained material throughout 2022 (Liadze, Macchiarelli, Mortimer-Lee, & Sanchez, 2023).

Finally, drawing from other studies, including Bae et al. (2021), Ding et al. (2021), Chiaramonte et al. (2022) and numerous others, the debt-to-assets ratio, cash-to-total-assets ratio, and a logarithm of total assets were chosen as firm-specific control variables. A summary of all the regressors used in the model is presented in the table below.

Table 1. A summary of regressors used in the model

| Variable | Type | Definition | Literature |
|---|---|---|---|
| ESG scores | Independent | Min-max normalised scores | Cornett et al. (2016) |
| leverage | Control | Total Debt / Total Assets | Ding et al. (2021) |
| cash ratio | Control | Cash and equivalents / Total Assets | Ding et al. (2021) |
| size | Control | Natural logarithm of total assets | Ding, et al. (2021) |
| covid | Dummy | 1 for Q1-Q3 2020, 0 for everything else | Bloom, et al. (2021) |
| war | Dummy | 1 for 2022, 0 for everything else | Liadze, et al. (2023). |

Source: compiled by author

Additionally to these regressors, interaction terms between E, S, and G scores and the shock variables were added to the model, which is a common practice to separate the effects of regressors with and without the shock present. The next sub-chapter follows with an overview of the initial data, sources, and data preparation techniques.

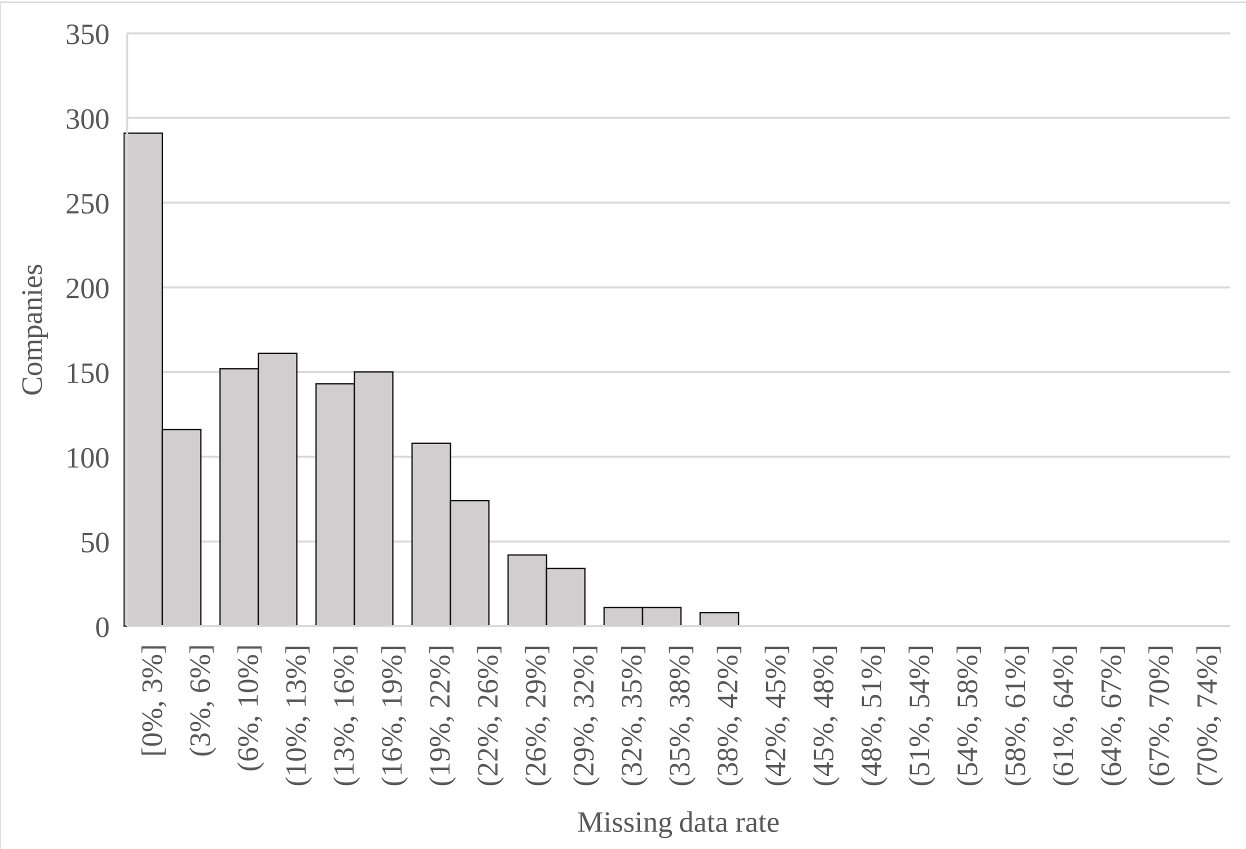
## 3.3. Data and sources

This study analyses the proposed relationship on a sample of 577 companies included in the IWV exchange-traded fund (ETF) (Haitin, 2025). IWV is a ticker under which the iShares Russell 3000 ETF is traded, which aims to track the performance of the Russell 3000 index. The index is composed of a broad selection of U.S. equities. This choice allows for an unbiased selection of companies, limited to a single region – the United States. ETF holdings data, company financials, and ESG scores have been retrieved using APIs provided by Financial Modelling Prep (FMP). FMP aggregates financial data from the Securities and Exchange Commission (SEC), Board of Governors, and Federal Reserve System. FMP's ESG scores are estimated using natural language processing (NLP) models to extract a selection of keywords from company filings, which is a practice included in the methodologies of rating agencies like Standard & Poor's and Sustainalytics.

Latest available information on the holdings of IWV served as an initial selection of companies. This original selection included a total of 2,644 companies. For each company, the following data with quarterly frequency from 2017 to 2024 has been retrieved: cash and cash equivalents, total debt, total assets, total revenues, EBITDA (earnings before interest, taxes, depreciation, and amortisation), environmental score, social score, and governance score. All financial variables were measured in U.S. dollars. The three ESG scores were presented scaled on a continuous range from 0 to 100.

While the FMP database is extensive and attractively priced, working with their API quickly revealed a less flattering reality: frequent inconsistencies, unexpected errors, and a generous helping of anomalous data entries. The informal terminology introduced later reflects a very real challenge in empirical work – the reliability of input data. This study openly acknowledges the limitations posed by such issues and has applied a range of corrective measures to extract meaningful signals from an otherwise noisy and, at times, disorderly dataset. A detailed description of data processing is presented in Appendix 1. The steps included cleaning the data from duplicates, removing companies with too many consecutively missing data points, imputing missing values, *declowning* the dataset, adjusting and removing outliers, and excluding companies with inadequate values.

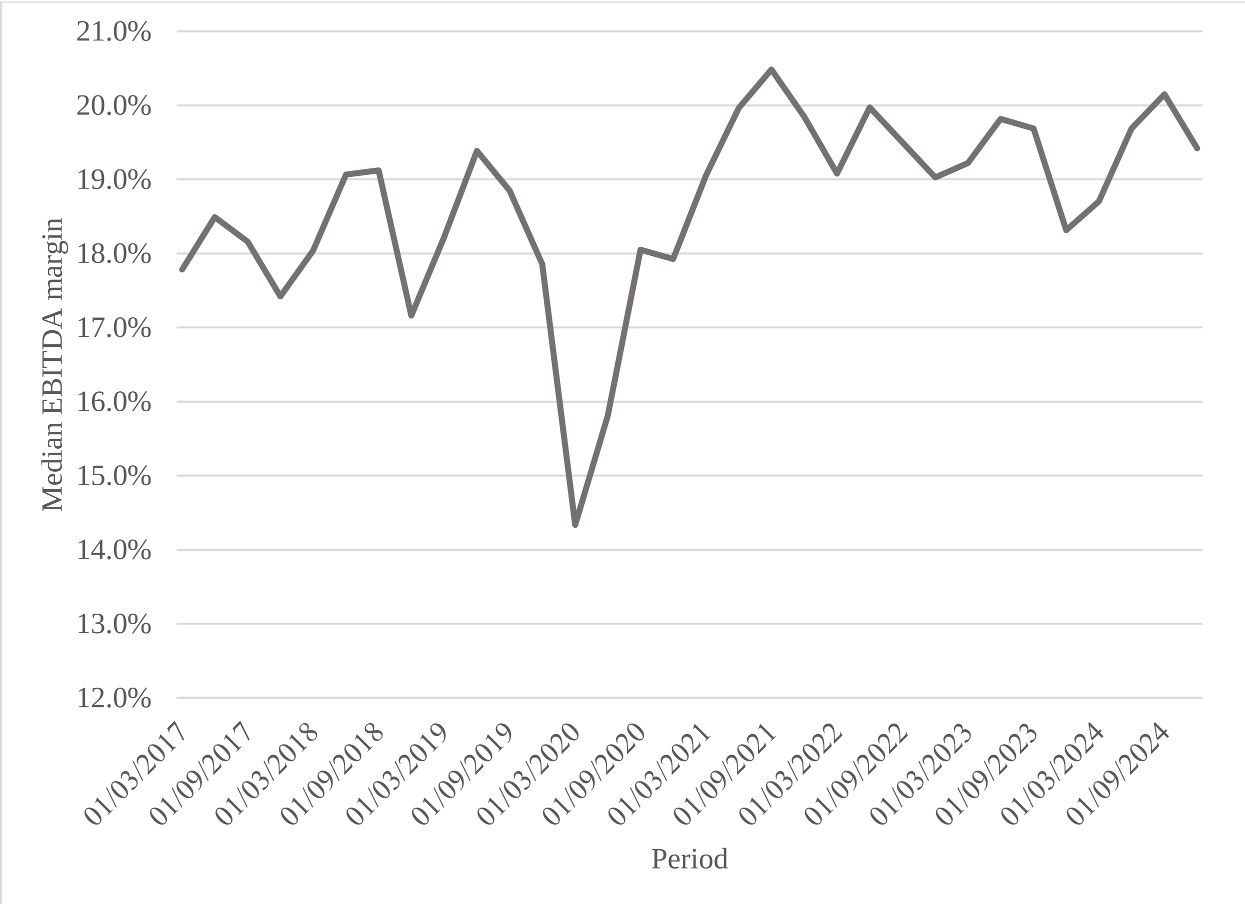
Additionally to that, after manually examining environmental, social, and governance scores, the chaotic nature of their time series became apparent – specifically, random and changing seasonal patterns, missing observations, non-stationarity, changing trends, and many more. In some cases, the scores were missing for over 30% of observations, as can be seen in Picture 1. A total of 13.2% of observations were missing.




Picture 1. A histogram of missing environmental, social and governance scores relative to the number of total observations for each company
Source: author calculations, based on the FMP data

Three approaches for solving the issue of missing data and imputing these values were considered, including linear interpolation, k-nearest neighbour imputation (kNN), and the Kalman filter. Considering ESG data complexities and the lack of significant correlation between the scores and other variables, a univariate Kalman filter was adopted as an imputation method. The Kalman filter allowed estimation of a true state for each ESG score, accounting for all the missing observations. A detailed comparison of the three imputation methods and the methodology behind the Kalman filter is presented in Appendix 2.

The next step was to consider seasonality in company profitability. Before that, however, EBITDA relative to total revenue ratios (EBITDA margins) were calculated. The problem of seasonality can be visually observed in a time series of median EBITDA margins, presented below.




Picture 2. Median EBITDA margin values across all companies
Source: author calculations, based on the FMP data

Seasonal trend decomposition using LOESS (STL) with robust weighting was systematically applied to companies exhibiting seasonal patterns of different frequencies. STL allowed deseasonalisation of data with as little effect on outliers as possible. A description of the exact methodology behind this process is provided in Appendix 3.

## 2.4. Descriptive statistics and data overview

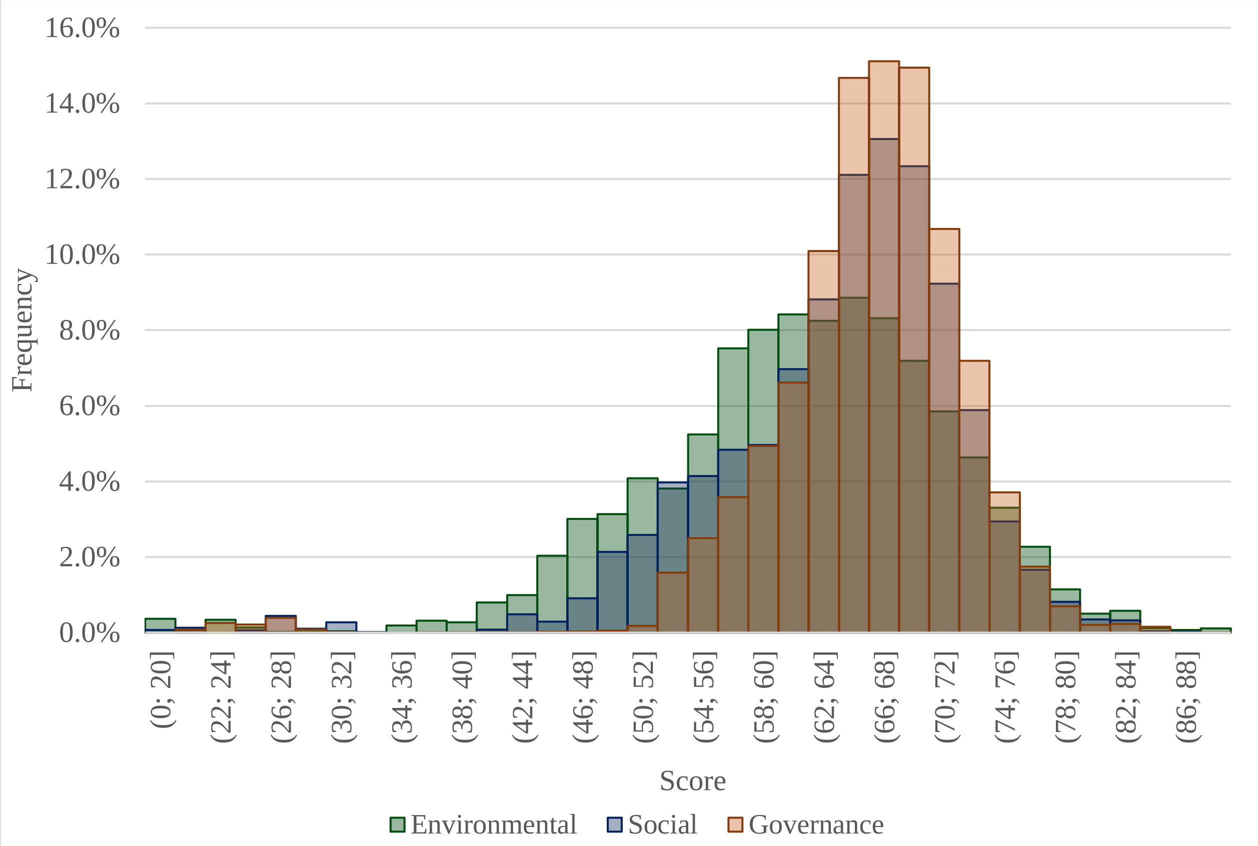
After data processing was completed, a balanced panel with 577 companies for the period 2017–2024 with quarterly frequency was achieved. The dataset included 11 different sectors, with the distribution of companies, average EBITDA margins, and average E, S, and G scores between sectors presented in Table 2. It can be seen that industrials clearly dominate the sample, representing 21% of total companies. Despite this, a variety of sectors remain represented, and considering random sampling, there is no reason to suspect any bias.

Table 2. Distribution of companies, average EBITDA margins (%), and average environmental, social and governance scores between sectors

| Sector | Total companies | EBITDA margin | Environm-ental score | Social score | Govern-ance score |
|---|---|---|---|---|---|
| Healthcare | 55 | 12.6 | 62.5 | 61.3 | 63.9 |
| Basic Materials | 35 | 14.0 | 58.1 | 62.9 | 66.6 |
| Industrials | 120 | 1.5 | 62.0 | 64.4 | 67.4 |
| Technology | 78 | 17.5 | 63.0 | 64.1 | 66.2 |
| Real Estate | 43 | 9.1 | 60.7 | 62.7 | 65.4 |
| Consumer Cyclical | 73 | 16.0 | 60.0 | 63.9 | 65.5 |
| Financial Services | 70 | 2.4 | 62.2 | 66.1 | 64.0 |
| Consumer Defensive | 38 | 6.3 | 62.9 | 64.9 | 67.2 |
| Utilities | 22 | 3.0 | 63.3 | 67.9 | 71.1 |
| Communication Services | 16 | 8.6 | 58.7 | 62.4 | 64.7 |
| Energy | 27 | 9.7 | 60.1 | 63.2 | 63.9 |

Source: author calculations, based on the FMP data (Haitin, Google Colab, 2025)

With regard to E, S, and G scores, clearly the average scores are all in the range of 60–70, while average governance scores are consistently higher than environmental and social ones across all sectors, with the exception of financial services. One possible explanation could be that companies give more attention to their governance practices. Taking into consideration that the scores are NLP-derived, another possible explanation could be that, on average, companies disclose better quality information with regard to their governance as compared to environmental and social aspects. This very pattern of higher governance scores can also be seen in Picture 3 below or by comparing the distributions of scores one by one in Appendices 6–8.




Picture 3. Distribution of environmental, social and governance scores frequency relative to total scores
Source: author calculations, based on the FMP data after applying the Kalman filter

From the distribution of ESG scores, high kurtosis and negative skewness also become rather apparent. A descriptive statistics table below confirms the visual conclusion and also sheds more light on the data differences between the three scores.

Firstly, environmental scores are both more volatile, show the lowest between and sample kurtosis, and have the most within variation. In contrast, governance scores show the least variation on any chosen metric – in-sample, between, and within entities – highlighting their stability and giving more weight to the assumption that companies tend to disclose more governance-related information in their filings, leading to more consistent NLP estimates. Also, interestingly, social scores have the highest kurtosis and are most right-skewed within entities, suggesting that over time the social aspects are kept rather stable.

Table 3. Sample, between and within summary statistics for environmental, social and governance scores

| Statistic | Environmental score | Social score | Governance score |
|---|---|---|---|
| sample statistics | | | |
| Mean | 61.5 | 64.0 | 66.0 |
| Standard deviation | 9.7 | 8.2 | 6.9 |
| Minimum | 0.4 | 14.5 | 20.2 |
| Maximum | 99.8 | 86.6 | 88.6 |
| Skewness | -0.8 | -1.3 | -2.2 |
| Kurtosis | 2.2 | 3.7 | 11.3 |
| between statistics | | | |
| CV (%) | 14.8 | 12.2 | 9.9 |
| Minimum | 12.1 | 20.5 | 21.9 |
| Maximum | 83.7 | 82.1 | 83.9 |
| Skewness | -1.1 | -1.5 | -2.5 |
| Kurtosis | 3.0 | 4.5 | 14.3 |
| within statistics | | | |
| CV (%) | 6.1 | 4.0 | 3.3 |
| Minimum | 8.3 | 20.7 | 46.7 |
| Maximum | 94.5 | 83.7 | 83.4 |
| Skewness | -0.5 | -1.3 | -0.5 |
| Kurtosis | 13.1 | 22.0 | 6.6 |

Source: author calculations, using Stata xtsum function in Python (Haitin, Google Colab, 2025)

Turning to financials, the data looks less stable, with extreme coefficients of variation (CV), high kurtosis values, and high within variation. Table 4 presents summary statistics for financial variables. As mentioned, CVs stand out, especially in sample statistics. However, between CVs are below 100 for ratios and even smaller for within statistics, meaning that the data is more unstable when taken as a whole. Total assets statistics, instead of their natural logarithms, are shown intentionally, as this is more interpretable. Extremely high kurtosis and positive skewness paint a clear picture of the distribution of companies in the sample – many large corporations, with outliers being mostly smaller companies rather than bigger ones, and the smallest value being just 6.6 million U.S. dollars. Another characteristic that draws attention is the highly negative minimum values of EBITDA margins. In fact, during data processing, even more extreme values were found, but the sample was restricted to EBITDA exceeding total revenues by only 5x, which is described in Appendix 1. Also, EBITDA margins are the only right-skewed variable in the dataset—it is the only variable where the outliers lie mostly to the left of the mean.

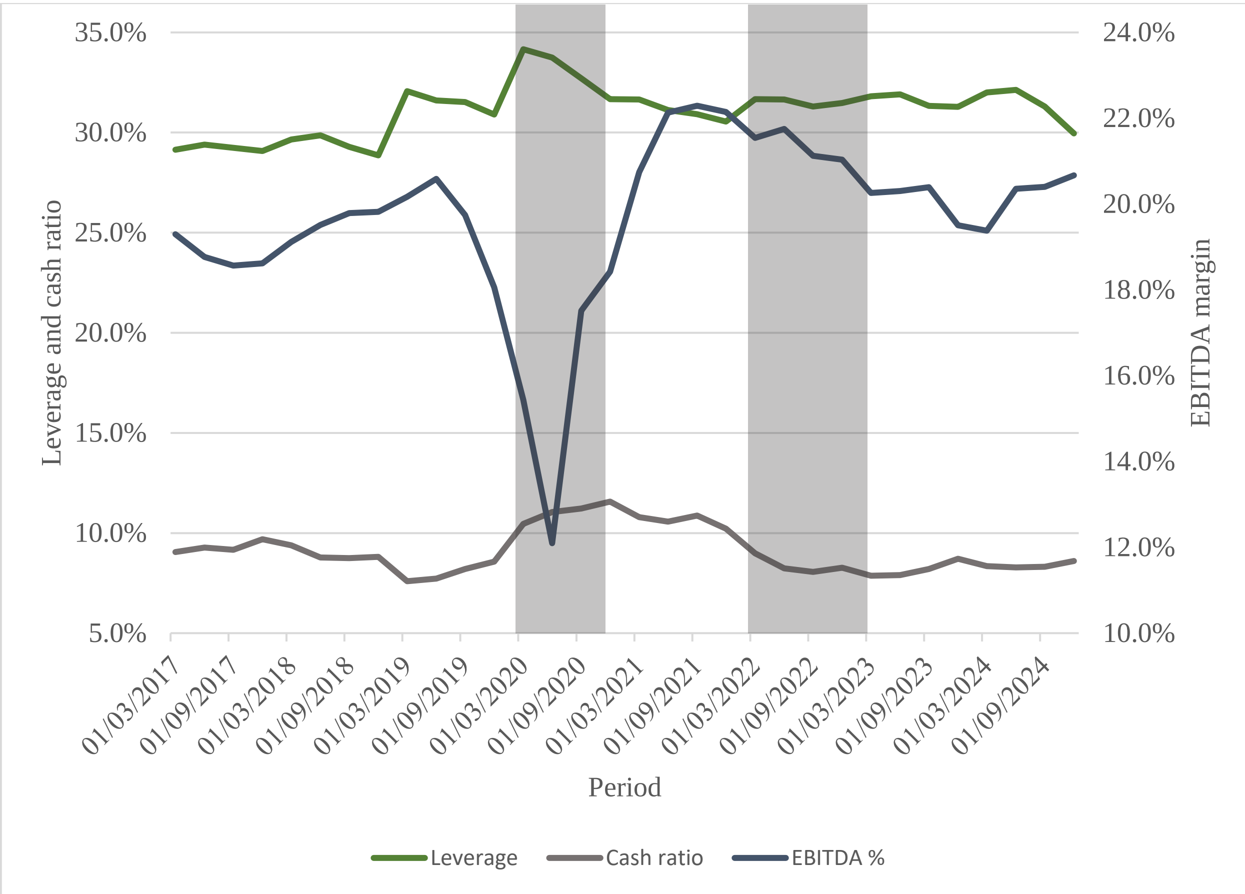
Table 4. Sample, between and within summary statistics for leverage, cash ratio, EBITDA margins (%) and total assets in billions U.S. dollars

| Statistic | EBITDA (%) | Leverage (%) | Cash ratio (%) | Total assets |
|---|---|---|---|---|
| sample statistics | | | | |
| Mean | 19.7 | 31.1 | 9.1 | 25.9 |
| St. dev. | 23.4 | 18.7 | 10.1 | 118.3 |
| CV (%) | 119.1 | 60.0 | 110.5 | 456.7 |
| Minimum | -485.4 | 0.0 | 0.0 | 6.6e-3 |
| Maximum | 88.8 | 89.7 | 83.6 | 2,455.1 |
| Skewness | -3.7 | 0.4 | 2.4 | 13.9 |
| Kurtosis | 53.6 | -0.3 | 8.2 | 234.5 |
| between statistics | | | | |
| CV (%) | 96.4 | 54.6 | 92.4 | 450.1 |
| Minimum | -88.5 | 0.6 | 0.0 | 1.2e-1 |
| Maximum | 79.5 | 77.9 | 56.2 | 2,190.7 |
| Skewness | -0.6 | 0.3 | 2.0 | 13.8 |
| Kurtosis | 5.9 | -0.3 | 5.6 | 230.2 |
| within statistics | | | | |
| CV (%) | 70.1 | 25.0 | 60.7 | 79.5 |
| Minimum | -417.2 | -17.7 | -32.6 | -353.4 |
| Maximum | 168.4 | 81.7 | 79.1 | 480.6 |
| Skewness | -8.8 | -0.2 | 1.5 | -0.7 |
| Kurtosis | 222.7 | 4.4 | 13.3 | 166.0 |

Source: author calculations, using Stata xtsum function in Python (Haitin, Google Colab, 2025)

While summary statistics provide a good understanding of the underlying data structure, visualisations are often useful to better understand data dynamics over time.

Picture 4 shows time series of mean EBITDA margins, leverage, and cash ratio. From the chart, it can be seen that profitability has indeed declined in both shock periods previously defined, while the debt-to-assets ratio has largely remained the same, and the cash-to-assets ratio increased during the COVID-19 pandemic, as companies increased their cash reserves in light of uncertainty.




Picture 4. Time-series of EBITDA margins, leverage and cash ratio averaged across entities. Shaded zones represent chosen macro-shock periods
Source: author calculations, based on the FMP data

It must be noted, however, that an evident downward trend in profitability starts from 2022. Considering that, besides the Russia–Ukraine war, other economic developments took place after the pandemic with arguably more impact on U.S. companies, it could be questioned whether the decline is related to the invasion or something else. This is further discussed in the results of the analysis.

Finally, correlation matrices are due to understand the relationships between variables and not only their univariate descriptive properties. Table 5 presents the between correlation matrix for the dependent variable and the regressors, except shocks and interactions, while Table 6 presents the within correlation matrix. It must be noted that the tables are presented purely for descriptive purposes, and statistical significance of coefficients is not provided, since to arrive at adequate estimates, one should account for cross-dependence of entities.

Table 5. Between entities correlation matrix for EBITDA margins (%), min-max normalised E, S and G scores, leverage, cash ratio in percentages and size as a natural logarithm

| Variable | EBITDA % | env_min max | soc_min max | gov_min max | leverage | cash ratio | size |
|---|---|---|---|---|---|---|---|
| EBITDA % | 1.000 | – | – | – | – | – | – |
| env_minmax | -0.018 | 1.000 | – | – | – | – | – |
| soc_minmax | -0.079 | 0.684 | 1.000 | – | – | – | – |
| gov_minmax | -0.042 | 0.574 | 0.656 | 1.000 | – | – | – |
| leverage | 0.120 | -0.050 | -0.099 | 0.035 | 1.000 | – | – |
| cash ratio | -0.369 | 0.021 | -0.037 | -0.046 | -0.160 | 1.000 | – |
| size | 0.420 | -0.108 | -0.173 | -0.086 | 0.090 | -0.310 | 1.000 |

Source: author calculations, based on the FMP data (Haitin, Google Colab, 2025)

The highest correlation coefficients are unsurprisingly between the normalised ESG scores themselves, with coefficients over 0.55 for all three variables. It is also noteworthy that each of the scores is barely correlated with any of the control variables, with the exception of social scores and size, where the relationship is negative. In fact, all three scores have negative correlation coefficients with the company size variable, which might suggest that for bigger companies, it is harder to achieve above-average performance.

Table 6. Within entities correlation matrix for EBITDA margins, min-max normalised E, S and G scores, leverage, cash ratio and size

| Variable | EBITDA % | env_min max | soc_min max | gov_min max | leverage | cash ratio | size |
|---|---|---|---|---|---|---|---|
| EBITDA % | 1.000 | – | – | – | – | – | – |
| env_minmax | 0.059 | 1.000 | – | – | – | – | – |
| soc_minmax | -0.013 | 0.339 | 1.000 | – | – | – | – |
| gov_minmax | 0.019 | 0.334 | 0.527 | 1.000 | – | – | – |
| leverage | -0.125 | -0.052 | -0.016 | -0.053 | 1.000 | – | – |
| cash ratio | 0.066 | -0.028 | 0.013 | 0.012 | -0.070 | 1.000 | – |
| size | 0.054 | 0.006 | -0.181 | -0.103 | 0.118 | -0.127 | 1.000 |

Source: author calculations, based on the FMP data (Haitin, Google Colab, 2025)

Within correlations show a different picture, where ESG scores have lower coefficients between themselves but are more correlated with control variables. Also, profitability is much less correlated with all other variables in the within statistics, as compared to between, suggesting that it is more affected by common factors rather than firm-specific ones. From both tables, it could also be concluded that there is no reason to suspect serious collinearity problems between the regressors, as most of the coefficients are rather small.

Now that a detailed overview of input data, sources, and processing methods has been described, and a specific modelling approach presented, the following chapter shows the results of the estimated models and their discussion.

# 3. ANALYSIS AND DISCUSSION

The current chapter is split into two sub-chapters. In the first one, the results of the estimated models are presented and their statistical properties described. The second section follows with a discussion of these results and their alignment or misalignment with other studies.

## 3.1. Analysis and results

Considering system GMM limitations in regard to panel data with large $T$, two models were estimated to capture firm performance before, during, and after the shock. Model 1 included the years 2018–2021 and was formulated as:

$$y_{it} = \alpha \times y_{it-1} + \sum_{i=1}^{3} \beta_i \times ESG_{it-1}^{i} + SHK + \sum_{j=1}^{3} \gamma_j S\_ESG_{it-1}^{j} + \sum_{k=1}^{3} \theta_k CNTR_{it-1}^{k} + \varepsilon_{it} \quad (5)$$

where
$y_{it}$ represents EBTIDA margin,
$ESG$ represent E, S, and G scores,
$SHK$ is a macro-shock dummy variable,
$S\_ESG$ are interactions between min-max normalised E, S, G scores and the shock dummy,
$CNTR$ represents control variables.

For Model 1, the shock was *covid* and $SHK$ was defined according to the Table 1 specification. Model 2 included the years 2021–2024 and was formulated using Equation (5), where $SHK$ also followed the respective Table 1 specification for the war variable. The decision to use first-order lags of independent variables follows Dwibedi et al (2024).

All of the ESG scores, as well as control variables, were instrumented, as they were both expected to be endogenous and model specifications improved with their instrumentation. The choice of instrument lags was based on improvements or deteriorations of model specification tests: AR(2) and the Hansen J test.

Both models were estimated for a panel with dimensions 577 × 20. A constant term was excluded due to collinearity issues between it and the size variable. Below, a summary table of Model 1 is presented.

Table 7. Model 1 summary results

| Regressor | Coefficient (s.e.) | p-value | Instrumented | Instrument lags |
|---|---|---|---|---|
| ebitda %, t-1 | 0.433*** (0.064) | 1.578e-11 | Yes | 3:4 |
| env_minmax | -0.023 (0.038) | 0.549 | Yes | 1:2 |
| soc_minmax | -0.081 (0.054) | 0.137 | Yes | 1:2 |
| gov_minmax | 0.029 (0.057) | 0.612 | Yes | 1:2 |
| covid | -0.019 *** (0.003) | 6.231e-8 | No | – |
| cov_env | -0.008 (0.031) | 0.806 | No | – |
| cov_soc | 0.023 (0.039) | 0.558 | No | – |
| cov_gov | -0.012 (0.028) | 0.666 | No | – |
| size | 0.009 *** (0.002) | 2.044e-6 | Yes | 2:3 |
| leverage | -0.092 * (0.053) | 0.082 | Yes | 2:3 |
| cash_ratio | -0.126 *** (0.046) | 0.006 | Yes | 2:3 |
| Arellano-Bond AR (1) | p-value = 0.007 | | | |
| Arellano-Bond AR (2) | p-value = 0.085 | | | |
| Hansen J test | p-value = 0.230, DoF = 223 | | | |
| Wald test | p-value = 2.22e-16 | | | |
| * p < 0.1, ** p < 0.05, *** p < 0.01 | | | | |

Source: author calculations using „plm“ package in R (Haitin, Google Docs, 2025)

The Arellano-Bond test statistic rejects the null hypothesis of no first-order autocorrelation and fails to reject the null of no second-order autocorrelation at the 5% significance level, suggesting serial correlation in first-differenced residuals and supporting the validity of instruments, as second-order differences are serially uncorrelated (Bond & Arellano, 1991). The Hansen J test shows a p-value of 0.230, failing to reject the null hypothesis of valid instruments and confirming that the moment conditions of GMM hold. Simultaneously, the Hansen J test is quite far from 1, meaning there is no reason to suspect model overfitting or instrument proliferation (Roodman, 2009). Model 1 included 230 instruments, resulting in 223 degrees of freedom in the Hansen J test, given 7 endogenous variables (Hansen L. P., 1982). The Wald test for joint significance of regressors strongly rejects the null, confirming the explanatory power of the model as a whole. The COVID dummy is statistically significant, reducing EBITDA margins by 1.9%, meaning that the shock successfully captures a meaningful structural break in profitability. Finally, all of the ESG variables, including their interactions, are statistically insignificant by themselves. After re-estimating the model without ESG variables and interactions, the Wald test showed the same p-value of 2.22e-16, indicating that ESG variables do not contribute meaningfully to the model's explanatory power and are likely redundant in this specification.

Model 2 shows largely contrasting results, especially in signifance of regressors. A summary table of the estimated model is presented below.

Table 8. Model 2 summary results

| Regressor | Coefficient (s.e.) | p-value | Instrumented | Instrument lags |
|---|---|---|---|---|
| ebitda %, t-1 | 0.744*** (0.054) | 2.000e-16 | Yes | 4:5 |
| env_minmax | 0.006 (0.014) | 0.654 | Yes | 4:5 |
| soc_minmax | -0.041* (0.022) | 0.063 | Yes | 4:5 |
| gov_minmax | 0.014 (0.023) | 0.546 | Yes | 4:5 |
| war | 0.001 (0.001) | 0.470 | No | – |
| war_env | 0.025 ** (0.012) | 0.030 | No | – |
| war_soc | -0.031 ** (0.014) | 0.033 | No | – |
| war_gov | 0.002 (0.011) | 0.864 | No | – |
| size | 0.004 ** (0.001) | 0.001 | Yes | 3:5 |
| leverage | -0.050 (0.033) | 0.131 | Yes | 3:5 |
| cash_ratio | -0.018 (0.033) | 0.577 | Yes | 3:5 |
| Arellano-Bond AR (1) | p-value = 0.004 | | | |
| Arellano-Bond AR (2) | p-value = 0.133 | | | |
| Hansen J test | p-value = 0.085, DoF = 299 | | | |
| Wald test | p-value = 2.22e-16 | | | |
| * p < 0.1, ** p < 0.05, *** p < 0.01 | | | | |

Source: author calculations using „plm" package in R (Haitin, Google Docs, 2025)

All three model specification tests – AR(1), AR(2), and the Hansen J test – once again support valid instruments and correct model specification. The Wald test p-value is also close to zero, suggesting joint significance of regressors. Model 2 included a total of 306 instruments, resulting in 299 degrees of freedom for the Hansen J test, given 7 endogenous variables. Regressor significance has changed, as both leverage and the cash ratio failed to reject the null hypothesis, while interaction variables of environmental and social scores, as well as the social score itself, are all significant at the 5% and 10% confidence levels, respectively. Interestingly, the war shock has both a coefficient of small magnitude and is found to be insignificant, with a p-value of 0.470 – meaning that the macro-shock either had low impact or was misspecified. Model 2 was also re-estimated without the shock and interactions, but no material improvements to the significance of coefficients were found.

Both models show interesting findings, and several possible explanations could be provided for them. The following section compares the results with other studies and touches on some important limitations of this analysis.

## 3.2. Discussion of the results

Firstly, it is important to reiterate the three main ways in which the analysis in this paper differs from the majority of other studies:

1. Raw ESG scores are replaced with min-max normalised values, serving as a proxy for relative ESG performance, less subject to rater bias.
2. The relationship is modelled for EBITDA margins, which represent operational performance, unaffected by stock price movements.
3. System GMM is adopted to account for endogeneity concerns.

This approach is inherently more skeptical, and lower coefficients could be expected; however, it is also more robust to numerous potential biases that can mistakenly find significant relationships, but for the wrong reasons.

Arguably, the biggest deviation from the majority of other studies is related to the insignificance of the governance aspect of ESG – in both models, neither the normalised score nor its interactions rejected the null hypothesis of a zero coefficient, or even had a borderline p-value. This contrasts heavily with the findings of Saygili et al. (2022) Juhee, et al. (2021) Moodhi, et al. (2023), and numerous others, who have found that corporate governance has not only a significant, but also the most substantial impact of the three ESG aspects on firm performance. In the case of this analysis, a possible explanation for the absence of a relationship can be purely mathematical, since governance scores in the sample are highly centered around the mean and have little variance, as was shown in the summary statistics. On the other hand, this could also suggest that relative governance performance is not necessarily related to financial performance, and the benefits become immaterial beyond a certain threshold. This explanation aligns with the findings of asymmetrical ESG benefits by Bax et al (2023).

Turning to the results of the first model, clear and definitive answers for all three questions posed in the introduction – although not too satisfactory – are achieved. Namely, relative ESG performance is not associated with better financial performance, nor does it provide protection against macro-shocks. This aligns with Bae et al (2021), who found no evidence of CSR affecting stock returns during the pandemic – a rather rare result. Considering that the majority of pandemic studies focused on market-based performance measures, including returns, risk-adjusted returns, stock price volatility, and others (Whelan, Atz, Van Holt, & Clark, 2021), the results in this paper could suggest that the positive relationship is indeed caused by investor preferences and capital inflows, rather than fundamental improvements in performance.

Regarding Model 2, the first thing to consider is the positive, although insignificant, coefficient of the macro-shock dummy – *war*. The reason could be that the shock is either incorrectly defined or its effects vary largely by sector, and an aggregate estimate becomes inadequate. This is in line with the contrasting findings of French et al (2023), who show negative impacts of the war on cumulative abnormal returns, and Nguyen & Khominich (2024), who provide evidence of exceptional profitability for fossil fuel companies in the United States. Nevertheless, environmental and social interaction variables are significant in Model 2, with respectively positive and negative effects on firm profitability following the first year after the invasion. This answers research question two: environmental aspects can provide protection in crises, social aspects don't improve but rather deteriorate resilience, and governance aspects remain immaterial. Taking into account the effects of the conflict on global energy markets (Liadze, Macchiarelli, Mortimer-Lee, & Sanchez, 2023), it is only logical that better environmental practices resulted in higher EBITDA margins, as companies relying more on carbon-effective energy sources were less affected by the energy price disruptions. This aligns with the findings of Shen et al. (2023). The negative and material impact of relative social scores – reducing margins by 4.1% during the period from 2021 to 2024, and by an additional 3.1% in 2022 – is in line with Dwibedi et al (2024), who similarly arrived at a significant and negative relation with ROE and ROA metrics, though their analysis focused on the pandemic years. This also provides an answer to research question three: social aspects have the most impact on financial performance, and the impact is highly negative.

More generally, the mixed results of this study with regard to the significance and direction of the relationships are aligned with two important findings of Whelan et al. (2021), from their meta-analysis of ESG studies. Firstly, ESG works best as a long-term strategy, while the analysis in this paper was performed over a shorter time period of four years to allow for system GMM as the analysis tool. For example, one of the explanations for the negative coefficients of ESG scores – also mentioned in several of the above-listed studies – is that firms aiming for better ESG performance are willing to sacrifice short-term profitability to retain their image and reputation, but such behaviour increases their long-term value nonetheless. This is also complemented by the theoretical view of ESG, as discussed in Chapter 1. Secondly, ESG disclosure does not drive financial performance. The ESG scores used in the analysis are NLP-derived from keywords in company filings, meaning that it is the very quality of disclosure that can have a significant impact on the scores. Furthermore, the more information a firm provides about itself, the more its ESG scores can vary, and the Kalman filter imputation and smoothing technique used in this paper can lead to even more discrepancies in the results – two issues that, among others, were discussed in detail by Serafeim & Kotsantonis (2019).

Finally, the author feels that, while it was not the aim of the study, one additional conclusion can be drawn from the results of the analysis. The NLP-derived scores – arguably the biggest limitation of this research – require a revision of methodology. The problem is also not exclusive to the database used in this analysis, as Standard & Poor's, MSCI, and other rating agencies also use NLP techniques as one of their ESG evaluation tools. Besides several studies already mentioned in this thesis, many others bring attention to the very issue of subjectivity and ESG score or rating divergence and inconsistency between different data providers, including Agosto et al. (2023), Christensen et al. (2022), Cesarone et al. (2024), and many others, who analyse the differences between the scoring systems, propose solutions, and examine the issue from the standpoint of investors, researchers, and companies.

# CONCLUSIONS

This thesis set out to examine the relationship between ESG performance, firm financial outcomes, and corporate resilience to macroeconomic shocks. Three core research questions guided the inquiry: (1) Is relative ESG performance, measured through ESG scores, associated with better firm performance? (2) Does relative ESG performance, measured through ESG scores, provide protection to firms when exposed to external shocks? (3) Is there a significant relationship between relative ESG scores and firm performance, and if there is, which aspects have the most effect, and in what direction? While these questions are common in ESG literature, this thesis applied a more rigorous and skeptical analytical approach, aiming specifically to address three main biases prevalent in prior research: rater bias, the concerns of market-based financial metrics, and simultaneity bias.

The theoretical foundation rests on stakeholder theory, which states that firms satisfying the needs of a broad set of stakeholders – not only their shareholders – can achieve higher value. This can happen due to improved employee productivity, more favorable supplier relations, increased customer loyalty and other reasons. Accordingly, the thesis explored whether such stakeholder-oriented behavior enhances firms' resilience to external shocks. These shocks can have both immediate (direct) effects – such as disruptions to supply and demand – and longer-term (indirect) effects via altered macroeconomic conditions, which in turn affect firm performance. On this basis, stakeholder theory would predict that firms with superior ESG performance should exhibit better financial performance during crises.

Empirical analysis was conducted using a system GMM approach across two recent periods characterized by distinct macroeconomic shocks: the 2018–2021 window encompassing the Covid-19 pandemic, and the 2021–2024 period, capturing the economic consequences of the Russian invasion of Ukraine. In the first period, no statistically significant relationship was found between ESG scores and financial performance, nor was there evidence of protective effects during the pandemic. In contrast, the second period revealed a short-term negative association between social ESG components and firm performance, alongside a positive influence of environmental factors in 2022 – a year marked by heightened energy market volatility due to the war in Ukraine. These mixed results suggest that, once key biases are accounted for, ESG scores do not consistently offer protection from external shocks or lead to improved financial outcomes.

Several limitations constrained the analysis. Arguably the biggest among them were the reliability of NLP-derived ESG scores and the relatively short time period analysed, which was necessary to apply a system GMM approach. Both limitations could have affected the underlying relationships. Nevertheless, the study underscores the necessity for future ESG research to explicitly confront the methodological biases identified here.

This thesis proposes a robust analytical approach to ESG analysis, which can shed more light on the direction of casuality between ESG scores and financial performance, the significance of their relationships and the effects. It was shown, that considering often omitted factors, results can divergy significantly from other studies and ESG scores might not be a good indicator of financial performance or corporate resilience to shocks.

Future work could extend this analysis by applying similar GMM methodologies across alternative ESG data providers, constructing composite or averaged ESG metrics, expanding the geographical scope, or incorporating a broader array of accounting-based performance indicators beyond EBITDA margins, used in this study, or ROE and ROA, most commonly analysed in other research. Such approaches would further illuminate whether ESG characteristics are linked to the fundamental financial resilience of firms.

# KOKKUVÕTE

## ESG, KASUM JA MAAVÄRINAD: MAKROMAJANDUSLIKU KAOSE VASTUPIDAVUSE SKEPTILINE ANALÜÜS

Mikael Haitin

See lõputöö uuris seoseid ESG tulemuslikkuse, ettevõtete finantstulemuste ja nende vastupanuvõime vahel makromajanduslikele šokkidele. Uuring keskendus kolmele peamisele uurimisküsimusele: (1) Kas ESG skooridega mõõdetav suhteline ESG tulemuslikkus on seotud paremate finantstulemustega? (2) Kas ESG skoorid pakuvad ettevõtetele kaitset majanduslike šokkide eest? (3) Kui ESG tulemuslikkus ja ettevõtte finantstulemused on omavahel seotud, siis millised aspektid mõjutavad seda seost kõige rohkem ja mis suunas? Kuigi need küsimused on ESG-teemalises kirjanduses levinud, rakendas see lõputöö rangemat ja kriitilisemat analüüsi, mille eesmärk oli käsitleda kolme peamist metoodilist kallutatust varasemas uurimistöös: ESG hindajate subjektiivsust, turupõhiste finantsnäitajate ebausaldusväärsust ning samaaegsuse põhjustatud seoseprobleeme.

Teoreetiline alus tugineb sidusrühmade teooriale, mille kohaselt ettevõtted, kes arvestavad lisaks aktsionäridele ka teiste osapoolte huvidega, loovad rohkem väärtust. Selle põhjuseks võib olla töötajate suurem pühendumus, paremad suhted tarnijatega, lojaalsemad kliendid jms. Seetõttu uuriti, kas sidusrühmadele keskenduv juhtimine suurendab ettevõtte vastupanuvõimet välistele šokkidele. Sellised šokid võivad avaldada kohest mõju (nt nõudluse ja pakkumise häired) või pikaajalist mõju muutunud makromajandusliku keskkonna kaudu, mis omakorda mõjutab ettevõtete finantstulemusi. Teooria kohaselt peaksid ettevõtted, kellel on tugevad ESG näitajad, kriiside ajal paremini toime tulema.

Empiiriline analüüs viidi läbi süsteemse GMM-meetodiga, hõlmates kahte hiljutist perioodi, mida iseloomustasid erinevad šokid: 2018–2021, mil toimus Covid-19 pandeemia, ja 2021–2024, mil oli mõju avaldanud Venemaa sissetung Ukrainasse. Esimesel perioodil ei ilmnenud ESG skooride ja finantstulemuste vahel olulist seost ega kaitsvat mõju pandeemia ajal. Teisel perioodil ilmnes lühiajaline negatiivne seos sotsiaalsete ESG komponentide ja finantstulemuste vahel, samas kui

keskkonnafaktoritel oli 2022. aastal positiivne mõju – perioodil, mida iseloomustas energiaturu suur volatiilsus seoses Ukraina sõjaga. Need vastuolulised tulemused viitavad, et kui metoodilised kallutused arvesse võtta, ei paku ESG skoorid järjepidevat kaitset šokkide vastu ega taga paremaid finantstulemusi.

Analüüsi piirasid mitmed tegurid. Kõige olulisemad olid NLP-põhiste ESG skooride usaldusväärsus ja suhteliselt lühike vaatlusperiood, mis oli vajalik süsteemse GMM-meetodi rakendamiseks. Mõlemad piirangud võisid mõjutada leitud seoseid. Siiski toob uurimus esile vajaduse, et tulevased ESG-alased uuringud käsitleksid neid metoodilisi puudujääke otseselt.

Lõputöö pakub tugeva analüütilise lähenemise ESG teemade uurimiseks, mis aitab paremini mõista ESG skooride ja finantstulemuste vahelist põhjuslikku seost, selle tähtsust ja mõju. Tulemused näitavad, et kui arvestada sageli tähelepanuta jäetud tegureid, võivad tulemused erineda märgatavalt varasemast teaduskirjandusest ning ESG skoorid ei pruugi olla usaldusväärne näitaja ettevõtte finantstulemustele või šokikindlusele.

Tulevikus võiks analüüsi laiendada, kasutades sarnaseid metoodikaid teiste ESG andmepakkujate puhul, luues liit- või keskmistatud ESG mõõdikuid, laiendades geograafilist haaret või kaasates rohkem raamatupidamispõhiseid finantsnäitajaid, näiteks lisaks EBITDA marginaalile ka ROE ja ROA. Selline lähenemine aitaks paremini mõista, kas ESG omadused on seotud ettevõtete finantsilise vastupidavusega.

# REFERENCES


Abaidoo, R. (2019). Corporate Performance Volatility and Adverse Macroeconomic Conditions: A Causal Interaction Perspective. *Journal of Financial Economic Policy, 11*(4), 533-547. doi:https://doi.org/10.1108/JFEP-11-2018-0158

Agosto, A., Giudici, P., & Tanda, A. (2023). How to combine ESG scores? A proposal based on credit rating prediction. *Corporate Social Responsibility and Environmental Management, 30*(6), 3222-3230. doi:https://doi.org/10.1002/csr.2548

Alphabet. (2025). *Google Colab*. Allikas: https://colab.research.google.com/

Anthropic. (2025). *Claude AI*. Allikas: https://claude.ai/new

Antwi, S., & Issah, M. (2017). Role of macroeconomic variables on firms' performance: Evidence from the UK. *Cogent Economics & Finance, 56*(1). doi:https://doi.org/10.1080/23322039.2017.1405581

Arellano, M., & Bover, O. (1995). Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics, 68*(1), 29-51.

Bae, K.-H., Ghoul, S. E., Gong, Z. J., & Guedhami, O. (2021). Does CSR matter in times of crisis? Evidence from the COIVD-19 pandemic. *Journal of Corporate Finance, 67*. doi:https://doi.org/10.1016/j.jcorpfin.2020.101876

Bai, J., & Ng, S. (2002). Determining the Number of Factors in Approximate Factor Models. *Econometrica, 70*(1), 191-221. doi:https://doi.org/10.1111/1468-0262.00273

Bai, J., & Ng, S. (2004). A Panic Attack on Unit Roots and Cointegration. *Econometrica, 72*(4), 1127-1177. Allikas: https://www.jstor.org/stable/3598781

Bai, J., & Ng, S. (2010). PANEL UNIT ROOT TESTS WITH CROSS-SECTION DEPENDENCE: A FURTHER INVES. *Econometric Theory, 26*(4), 1088-1114. Allikas: https://www.jstor.org/stable/40800875

Berg, F., Kölbel, J. F., & Rigobon, R. (2019). Aggregate Confusion: The Divergence of ESG Ratings. *Forthcoming Review of Finance*. doi:http://dx.doi.org/10.2139/ssrn.3438533

BlackRock. (2025). *iShares*. Allikas: iShares Russell 3000 ETF: https://www.ishares.com/us/products/239714/ishares-russell-3000-etf

Blanchard, O., & Fischer, S. (1989). *Lectures on macroeconomics*. Cambridge, Massachussets: MIT Press.

Bloom, N., Fletcher, R. S., & Yeh, E. (2021). THE IMPACT OF COVID-19 ON US FIRMS. *NBER WORKING PAPER SERIES, 28314*. doi:10.3386/w28314

Bond, S., & Arellano, M. (1991). Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *The Review of Economic Studies, 58*(2), 277–297. doi:https://doi.org/10.2307/2297968

Chen, N.-F., Roll, R., & Ross, S. A. (1986). Economic Forces and the Stock Market. *The Journal of Business, 59*(3), 383-403. Allikas: https://www.jstor.org/stable/2352710

Chiaramonte, L., Dreassi, A., Girardone, C., & Piserà, S. (2022). Do ESG strategies enhance bank stability during financial turmoil?Evidence from Europe. *The European Journal of Finance, 28*(12), 1173-1211. doi:https://doi.org/10.1080/1351847X.2021.1964556

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

Chytis, E., Eriotis, N., & Mitroulia, M. (June 2023. a.). Hotspots and trends of environmental, social and governance (ESG) research: a bibliometric analysis. *Data Science and Management, 6*(2), 65-75. doi:https://doi.org/10.1016/j.dsm.2023.03.001

Clément, A., Robinot, É., & Trespeuch, L. (March 2025. a.). The use of ESG scores in academic literature: a systematic literature review. *Journal of Enterprising Communities: People and Places in the Global Economy, 19*, 92-110. doi:https://doi.org/10.1108/JEC-10-2022-0147

Cornett, M. M., Erhemjamts, O., & Tehranian, H. (2016). Greed or good deeds: An examination of the relation between corporate social responsibility and the financial performance of U.S. commercial banks around the financial crisis. *Journal of Banking & Finance, 70*, 137-159. doi:https://doi.org/10.1016/j.jbankfin.2016.04.024

Demers, E., Hendrikse, J., Joos, P., & Lev, B. (2021). ESG Didn't Immunize Stocks During the COVID-19 Crisis, But Investments in Intangible Assets Did. *Journal of Business Finance & Accounting, 48*, 433-462. doi:doi.org/10.1111/jbfa.12523

Dickey, D. A., & Fuller, W. A. (1979). Distribution of the Estimators for Autoregressive Time Series With a Unit Root. *Journal of the American Statistical Association, 73*(366), 427-431. doi:https://doi.org/10.2307/2286348

Ding, W., Levine, R., Lin, C., & Xie, W. (2021). Corporate immunity to the COVID-19 pandemic. *Journal of Financial Economics, 141*, 802-830. doi:https://doi.org/10.1016/j.jfineco.2021.03.005

Dongjin, J., Juhee, H., & Hyuna, K. (2021). The Effect of ESG Activities on Financial Performance during the COVID-19 Pandemic—Evidence from Korea. *Sustainability, 13*(20), 11362. doi:https://doi.org/10.3390/su132011362

Dwibedi, P., Pahi, D., & Sahu, A. (2024). Unveiling the Relationship Between ESG Scores and Firm. *Australasian Accounting, Business and Finance Journal, 18*(3).

Egorov, A. (2023). ESG as an innovative tool to improve the efficiency and financial stability of financial organizations. *Procedia Computer Science, 221*, 705-709. doi:https://doi.org/10.1016/j.procs.2023.08.041

Fernández, M. S., Abu-Alkheil, A., & Khartabiel, G. M. (2019). Do German Green Mutual Funds Perform Better Than Their. *10*(2), 297-312. doi:doi: 10.20409/berj.2019.169

FMP. (2025). *Financial Modelling Prep*. Allikas: https://site.financialmodelingprep.com/

Freeman, R. E. (1984). *Strategic management: A stakeholder approach.* Cambridge, UK: Cambridge University Press.

Freeman, R. E. (1994). The Politics of Stakeholder Theory: Some Future Directions. *Business Ethics Quarterly, 4*(4), 409-421. doi:https://doi.org/10.2307/3857340

Friedman, M. (1962). *Capitalism and Freedom.* Chicago, IL: University of Chicago Press.

Giakoumelou, A., Salvi, A., Bertinetti, G. S., & Micheli, A. P. (2022). 2008's mistrust vs 2020's panic: can ESG hold your institutional investors? *Management Decision, 60*(10), 2770-2785. doi:https://doi.org/10.1108/MD-12-2021-1669

Haitin, M. (2025). *Google Colab*. Allikas: https://colab.research.google.com/drive/1E50Qn8NUkMCrDalEoGrkfKR5-RyKpvzB?usp=sharing

Haitin, M. (2025). *Google Docs*. Allikas: https://docs.google.com/document/d/1_4yC7kRByFlFxn_c4Gl7FQe0NjjSE_wRuCQie7mGCfo/edit?usp=sharing

Haitin, M. (2025). *Google Drive*. Allikas: https://drive.google.com/file/d/1WgGXhdwy3cnEgXt3YIuGe6NEqq3dIkG_/view?usp=sharing

Hansen, G. S., & Wernerfelt, B. (1989). Determinants of Firm Performance: The Relative Importance of Economic and Organizational Factors. *Strategic Management Journal, 10*(5), 399-411. Allikas: https://www.jstor.org/stable/2486469
Hansen, L. P. (1982). Large Sample Properties of Generalized Method of Moments Estimators. *Econometrica, 50*(4), 1029-1054. doi:https://doi.org/10.2307/1912775
Harvey, A. C. (1990). *Forecasting, Structural Time Series Models and the Kalman Filter.* Cambridge: Cambridge University Press.
Heese, J., Wang, C. C., & Labruyere, T. (March 2021. a.). *Wirecard: The Downfall of a German Fintech Star.* Harvard Business School Case.
Holton, G. A. (2004). Defining Risk. *Financial analysts journal, 60*(6), 19-25. doi:https://doi.org/10.2469/faj.v60.n6.2669
Hong, H., & Kacperczyk, M. (July 2009. a.). The price of sin: The effects of social norms on markets. *Journal of Financial Economics, 93*(1), 15-36. doi:https://doi.org/10.1016/j.jfineco.2008.09.001
Ilhan, E., Krueger, P., Sautner, Z., & Starks, L. T. (July 2023. a.). Climate Risk Disclosure and Institutional Investors. *The Review of Financial Studies, 36*(7), 2617–2650. doi:https://doi.org/10.1093/rfs/hhad002
Ilhan, E., Sautner, Z., & Vilkov, G. (2021). Carbon Tail Risk. *The Review of Financial Studies, 34*(3), 777-799. doi:http://dx.doi.org/10.2139/ssrn.3204420
International Organization for Standardization. (2010). *ISO 26000 Social responsibility.* Allikas: ISO Standards: https://www.iso.org/iso-26000-social-responsibility.html
Jones, T. M. (April 1995. a.). Instrumental Stakeholder Theory: A Synthesis of Ethics and Economics. *The Academy of Management Review, 20*(2), 404-437.
Kalman, R. E. (1960). A New Approach to Linear Filtering.
Kellstedt, P. M., & Boef, S. D. (2004). The Political (and Economic) Origins of Consumer Confidence. *American Journal of Political Science, 48*(4), 633-649. doi:https://doi.org/10.1111/j.0092-5853.2004.00092.x
Koopman, S. J., & Durbin, J. (2013). *Time Series Analysis by State Space Methods, 2nd edn.* Oxford: Oxford Statistical Science Series.
Kou, S., & Peng, X. (2016). On the Measurement of Economic Tail Risk. *Operations Research, 64*(5), 1053-1176. doi:https://doi.org/10.1287/opre.2016.1539
Li, T.-T., Wang, K., Sueyoshi, T., & Wang, D. D. (2021). ESG: Research Progress and Future Prospects. *Sustainability, 13*(21). doi:https://doi.org/10.3390/su132111663
Li, Y. (October 2010. a.). The Case Analysis of the Scandal of Enron. *International Journal of Business and Management, 5*(10).
Liadze, I., Macchiarelli, C., Mortimer-Lee, P., & Sanchez, P. J. (2023). Economic costs of the Russia-Ukraine war. *The World Economy, 46*(4), 874-886. doi:https://doi.org/10.1111/twec.13336
Lin, X., Liu, M., So, S., & Yuen, D. (2019). Corporate social responsibility, firm performance and tax risk. *Managerial Auditing Journal, 34*(9), 1101-1130. doi:https://doi.org/10.1108/MAJ-04-2018-1868
Liu, X., Zhao, T., & Li, R. (2023). Studying the green economic growth with clean energy and green finance: The role of financial policy. *Renewable Energy, 215*. doi:https://doi.org/10.1016/j.renene.2023.118971
Mahajan, R., Lim, M. W., Sareen, M., Kumar, S., & Panwar, R. (November 2023. a.). Stakeholder theory. *Journal of Business Research, 166*. doi:https://doi.org/10.1016/j.jbusres.2023.114104
Monard, M. C., & Gustavo, B. E. (2002). A Study of K-Nearest Neighbour as an. *His, 87*, 251-260.

Moodhi, R., Asma, M., & Nisar, A. (2023). Impact of global financial crisis on firm performance in UK: Moderating role of ESG, corporate governance and firm size. *Cogent Business & Management, 10*(1), 2167548. doi:https://doi.org/10.1080/23311975.2023.2167548

Morgan Stanley Capital International. (2024). *ESG Ratings Methodology.* MSCI ESG Research LLC .

Netter, J. M., Linck, J. S., & Wintoki, B. (2012). Endogeneity and the dynamics of internal corporate governance. *Journal of Financial Economics, 105*(3), 581-606. doi:https://doi.org/10.1016/j.jfineco.2012.03.005

Nguyen, D. H., & Khominich, I. P. (2024). Financial performance of EU-27 fossil fuel companies and their counterparts after imposing energy sanctions on Russia: A comparative analysis. *10*, 190-210. doi: 10.32609/j.ruje.10.124364

Nofsinger, J. R., Sulaeman, J., & Varma, A. (October 2019. a.). Institutional investors and corporate social responsibility. *Journal of Corporate Finance, 58*, 700-725. doi:https://doi.org/10.1016/j.jcorpfin.2019.07.012

O'Hare, J. (2022). Don't forget the "g" in esg: the sec and corporate governance disclosure. *Arizona Law Review, 64*(2), 417-462.

OpenAI. (2025). *ChatGPT*. Allikas: https://chatgpt.com/

Oxelheim, L. (2003). Macroeconomic Variables and Corporate Performance. *Financial Analysts Journal*, *59*(4), 36-50.

Paterlini, S., Czado, C., Sahin, Ö., & Bax, K. (2023). ESG, risk, and (tail) dependence. *International Review of Financial Analysis, 87*, 102513. doi:https://doi.org/10.1016/j.irfa.2023.102513

Pérignon, C., Hurlin, C., Colliard, J.-E., & Benoit, S. (2017). Where the Risks Lie: A Survey on Systemic Risk. *Review of Finance, 21*(1), 109–152. doi:https://doi.org/10.1093/rof/rfw026

Pesaran, M. H. (2004). General Diagnostic Tests for Cross. *CESifo Working Paper Series No. 1229*.

Pesaran, M. H. (2007). A simple panel unit root test in the presence of cross-section dependence. *Journal of Applied Econometrics, 22*(2), 265-312. doi:https://doi.org/10.1002/jae.951

Pollman, E. (2024). The Making and Meaning of ESG. *Harvard Business Law, 14*(2), 403-454.

Posit. (2025). *Posit Cloud*. Allikas: https://posit.cloud.com/

Ramey, V. A. (2016). Macroeconomic Shocks and Their Propagation. *National Bureau of Economic Research*(21978). doi:10.3386/w21978

Roodman, D. (2009). A Note on the Theme of Too Many Instruments. *Oxford Bulletin of Economics and Statistics, 71*(1), 135–158. doi:https://doi.org/10.1111/j.1468-0084.2008.00542.x

S&P. (2025). *S&P Global ESGScores Methodology.* S&P Global Inc.

Saygili, E., Birkan, A. O., & Arslan, S. (2022). ESG practices and corporate financial performance: Evidence from Borsa Istanbul. *Borsa Istanbul Review, 22*(3), 525-533. doi:https://doi.org/10.1016/j.bir.2021.07.001

Schoenmaker, S. D., Cauwenberge, P. V., & Bauwhede, H. V. (2014). Effects of local fiscal policy on firm profitability. *The Service Industries Journal, 34*(16), 1289-1306. doi:https://doi.org/10.1080/02642069.2014.942654

Scozzari, A., Ricca, F., Martino, M. L., & Cesarone, F. (2024). Managing ESG ratings disagreement in sustainable portfolio selection. *Computers & Operations Research, 170*, 106766. doi:https://doi.org/10.1016/j.cor.2024.106766

Senadheera, S. S., Gregory, R., Rinklebe, J., Farrukh, M., & Rhee, J. H. (2022). The development of research on environmental, social, and governance (ESG): A bibliometric analysis. *Sustainable Environment, 8*(1). doi:https://doi.org/10.1080/27658511.2022.2125869

Serafeim, G., & Kotsantonis, S. (2019). Four Things No One Will Tell You About ESG Data. *Journal of Applied Corporate Finance, 31*(2), 50-58. doi:https://doi.org/10.1111/jacf.12346

Sheehy, B. (2015). Defining CSR: Problems and Solutions. *Journal of Business Ethics, 131*, 625-648. doi:10.1007/s10551-014-2281-x

Shen, J., Chen, C., & Liu, Z. (2023). Does environmental investment pay off?—portfolio analyses of the E in ESG during political conflicts and public health crises. *Journal of Financial Research, 46*, S107-S131. doi:https://doi.org/10.1111/jfir.12357

Shin, S., Gurdgiev, C., & French, J. J. (2023). When doing the right thing doesn't pay: Impact of corporate decisions on Russian market participation in the wake of the Ukraine-Russia war. *Finance Research Letters, 58*(B), 104468. doi:https://doi.org/10.1016/j.frl.2023.104468

Smith, D. G. (1998). The Shareholder Primacy Norm. *Journal of Corporation Law, 23*(2), 277-324.

Stout, L. A. (2002). BAD AND NOT-SO-BAD ARGUMENTS FOR SHAREHOLDER PRIMACY. *Southern, 75*(5), 1189-1210.

Townsend, B. (2020). From SRI to ESG: The origins of socially responsible and sustainable investing. *The Journal of Impact and ESG Investing, 1*(1), 10-25.

Tripopsakul, S., & Puriwat, W. (2022). Understanding the Impact of ESG on Brand Trust and Customer Engagement. *Journal of Human, Earth and Future, 3*(4), 430-440. doi:10.28991/HEF-2022-03-04-03

Tse, T. (2011). Shareholder and stakeholder theory: after the financial crisis. *Qualitative Research in Financial Markets, 3*(1), 51-63. doi:https://doi.org/10.1108/17554171111124612

United Nations Global Compact. (2004). *Who Cares Wins: Connecting Financial Markets to a Changing World.* United Nations Department of Public Information.

van der Beck, P. (September 2024. a.). Flow-Driven ESG Returns. *Swiss Finance Institute Research Paper No. 21-71*. doi:http://dx.doi.org/10.2139/ssrn.3929359

Wang, H., Jiao, S., & Ma, C. (2024). The impact of ESG responsibility performance on. *International Review of Economics and Finance, 93*, 1115-1129. doi:http://doi.org/10.1016/j.iref.2024.05.033

Whelan, T., Atz, U., Van Holt, T., & Clark, C. (2021). *ESG AND FINANCIAL PERFORMANCE.* NYU Stern Center for Sustainable Business.

Whited, T. M., & Roberts, M. R. (2013). Chapter 7 - Endogeneity in Empirical Corporate Finance. *Handbook of the Economics of Finance, 2*, 493-572. doi:https://doi.org/10.1016/B978-0-44-453594-8.00007-0

Wu, J., Lodorfos, G., Dean, A., & Gioulmpaxiotis, G. (2015). The Market Performance of Socially Responsible Investment during Periods of the Economic Cycle – Illustrated Using the Case of FTSE. *Managerial and Decision Economics, 38*(2), 238-251. doi:doi/10.1002/mde.2772

Youngs, R., & Goldthau, A. C. (2023). The EU Energy Crisis and a New Geopolitics of Climate. *Journal of Common Market Studies*, 1-10. doi:10.1111/jcms.13539

Zhou, D., & Zhou, R. (2021). ESG Performance and Stock Price Volatility in Public Health Crisis: Evidence from COVID-19 Pandemic. *International Journal of Environmental Research and Public Health, 19*(1), 202. doi:https://doi.org/10.3390/ijerph19010202

Zingales, L., & Hart, O. (2022). The New Corporate Governance. *University of Chicago, 1*(1), 195-216.

# APPENDICES

## Appendix 1. A detailed overview of the data cleaning process

Data cleaning was performed in multiple steps. Several times the methodology had to be revised as new issues became evident. During the cleaning process data quality was prioritised over the number of cross-sectional units in the dataset.

The first step included ensuring panel data balance, by excluding any companies lacking complete quarterly data for the period 2017–2024. Duplicate period entries within cross-sections (that is, individual companies) were also removed, leaving only the latest available information. Additionally, companies with more than eight consecutive missing ESG scores were excluded, as imputing nine or more consecutive observations was considered methodologically unreliable.

The second step included scanning for companies with over three consecutively missing financial data entries. If they had less, the values were imputed using averages of the nearest available observations. Companies that had more were excluded from the sample.

For step three, several anomalous data patterns were manually identified in EBITDA and total revenue values. Such anomalies are henceforth referred to as *clowns* and the process of cleaning the panel from outliers is referred to as *declowning*. Companies exhibiting *clown-type* patterns were systematically flagged and, where feasible, corrected; otherwise, they were removed from the dataset. The following categories of *clowns* were identified:

1. *Reversed EBTIDA* – a pattern in which EBITDA values remain relatively stable over five consecutive quarters, except for the middle observation, which exhibits a sign reversal while maintaining a similar magnitude. Similar magnitude was defined as a value within 1.75 times the size of its immediate neighbours. A total of 86 EBITDA signs were reversed.
2. *Magnitude typo* – a case where an EBITDA value is approximately an order of magnitude (that is, 10x) larger or smaller than its two adjacent values, suggesting a likely data entry error. If the outlier's ratio relative to both neighbors was between 9 and 11, the value was

corrected by multiplying or dividing by 10, accordingly. A total of 10 such values were adjusted.

3. *Revenue confusion* – a scenario in which total revenue values have been replaced by gross profit figures, and in some instances, later reverted back to actual revenue values. This pattern was found to be common in the insurance and banking sectors, where revenue reporting conventions can resemble gross profit in company filings. A total of 25 companies exhibiting this pattern were excluded from the dataset.

For step three, additional likely erroneous extreme outliers in financials were treated in the following way:

- Identifying extreme outliers as values beyond four standard deviations from the mean,
- Excluding the first halves of 2020 and 2022, assuming outliers could occur naturally in these periods,
- The value was considered a curable data entry mistake, if it was a single extreme value in-between five consecutive observations that remained within the limits of two standard deviations from the mean.

Curable entry mistakes were corrected by replacing them with averages for a total of 183 cured datapoints. Other companies were removed from the sample. After the above procedures, 1244 companies remained in the dataset, resulting in a balanced panel with dimensions 1244 × 32.

After cacluating control variables and imputing ESG scores, two more issues came to light. First, some companies exhibited fixed ESG scores, or minimal variance scores for the whole period. Second, some companies had inadequate financial ratios exceding 1 or extremely lower than 0. The following restrictions on financial ratios were defined:

- EBITDA margins in between -5 and 1,
- Debt-to-Assets ratio in between 0 and 1,
- Cash-to-Assets ratio in between 0 and 1.

Any companies that did not satisfy the above criteria were excluded from the dataset, removing a total of 242 companies. Comanies with low ESG variance were also excluded from the dataset, where low variance was defined as less then 0.01. A total of 425 companies were removed during this process. After all the cleaning procedures, a total of 577 entities were left, on which the analysis was performed (Haitin, Google Colab, 2025).

## Appendix 2. ESG scores imputation methodology

As shown in Picture 1, ESG scores had a lot of missing data points. Additionally to that, they exhibited chaotic behaviour. Three approaches for solving the issue of missing data were considered:

1. Linear interpolation,
2. k-nearest neighbours imputation,
3. The Kalman filter.

Linear interpolation uses linear polynomials to impute missing values between the two known observations. Despite it's simplicity, however, considering that companies could have up to eight consecutively missing values, linear interpolation would not be able to yield reliable estimates. Additionally, it could potentially alter the underlying trends in the data, introducing linearity, where it may not exist.

k-nearest neighbors (kNN) imputation is a non-parametric data processing technique that fills in missing observations using the values from the k most similar rows in the dataset. Although this method could produce more reliable estimates and was found to handle large amounts of missing data well (Monard & Gustavo, 2002), kNN is a multivariate imputation technique – it relies on the assumption that variables are correlated. Since there is no *a priori* justification to assume a strong correlation between financial variables and NLP-derived ESG scores, the nearest neighbours identified by kNN may not share similarities with the missing observations. To test this, simulations were conducted on a sample of companies with complete ESG datasets, where values were artificially removed at random. Based on those simulations, univariate temporal kNN imputation consistently outperformed multivariate kNN, suggesting that no strong multivariate dependencies exist between ESG scores and other data.

Given the underlying complexity of ESG scores, Kalman filter was finally adopted as a processing tool, as it allows simultaneous treatment of the problem of missing data and accounts for their chaotic nature. The Kalman filter is an optimal recursive algorithm that uses a dynamic model and noisy observations to estimate the state of a system. In his seminal paper, Kalman (1960) proposes a solution to the problem of optimally estimating and predicting random signals from observable noisy data. Following his work, the problem of estimating a trend for ESG scores could be formulated as a *smoothing* problem with the goal of computing an optimal (mean squared error minimising) true-state of $ESG(t_1|t)$, where $t_1 < t$, given the observable values

$$ESG(t) = ESG_1(t) + ESG_2(t), \tag{6}$$

where
$ESG_1(t)$ is the signal,
$ESG_2(t)$ is noise.

Missing values are naturally handled by the Kalman filter, since it is a recursive process – it first estimates the state equation and observation equation based on the existing data. Both equations are later adjusted in the Kalman gain based on the actual observed value; as such, if no value exists, the process simply continues on, skipping the update step. However, to correctly estimate the true state of the system, the Kalman filter requires accurate specification of both state noise ($\sigma^2_{state}$) and observation noise ($\sigma^2_{obs}$) parameters, which are necessary for calculating the state disturbance covariance matrix (Q) and observation disturbance covariance matrix (R). It must be noted here that, while ideally ESG scores for each company would be processed using company-specific parameters, due to computational infeasibility, a simplification is required. Hence, it is assumed that each of the three scores exhibits similar behaviour between companies, leaving the need to estimate a total of six optimal parameters – two for each score. Following Harvey (1990), a log-likelihood function is formulated for a case of a univariate space-time model:

$$log\ L = -\frac{T}{2}\log(2\pi) - \frac{1}{2}\sum_{t=1}^{T}\log F_t - \frac{1}{2}\sum_{t=1}^{T}\frac{v_t^2}{F_t} \tag{7}$$

where
$L$ is the likelihood function,
$T$ is the number of time periods,
$F_t$ is the forecast error variance at time t, where $F_t = f(\sigma^2_{state}, \sigma^2_{obs})$,
$v_t$ is the innovation term at time t, that is $ESG(t) - E[ESG(t)|ESG(t-1)]$ .

Given the log-likelihood function a maximisation problem has been formulated:

$$\max_{\sigma^2_{state},\ \sigma^2_{obs}} \sum_{m=1}^{M} log\ L_m(\sigma^2_{state},\ \sigma^2_{obs}) \tag{8}$$

where
$L$ is the likelihood function,
$M$ is the number of companies,
$\sigma^2_{state}$ is the state noise parameter,
$\sigma^2_{obs}$ is the observation noise parameter.

To arrive at the optimal parameters, Optuna – a Bayesian hyperparameter optimisation software in Python – was used. First, companies with at least 30 consecutive ESG scores were identified, resulting in a sub-sample of 113 companies. For each of the environmental, social, and governance scores, 30 Optuna trials were performed with an objective presented in equation (8). Additionally, the Ljung-Box test was carried out with up to four lags on the noise component of the process, and for each score, the percentage of companies that failed to reject the null hypothesis of no autocorrelation at the 95% confidence level was calculated. The results are presented in the table below:

Table 9. Results of Kalman filter parameter optimisation using Optuna

| Parameter | Environmental | Social | Governance |
|---|---:|---:|---:|
| State noise | 0.006218 | 0.007944 | 0.002869 |
| Observation noise | 0.850505 | 0.037205 | 0.023230 |
| Log-likelihood value | -15,676 | -13,626 | -13,442 |
| Ljung-Box percentage | 97.7% | 100% | 100% |

Source: author calculations using Optuna library in Python (Haitin, Google Colab, 2025)

## Appendix 3. Adjustment for seasonality

With regard to financial data, the main concern was seasonal patterns in EBITDA figures. Through manual examination of autocorrelation function (ACF) with up to 16 lags on several arbitrarily selected companies, two seasonal patterns were identified:

1. ACF lag spikes – a pattern, where a time-series with seasonal frequency $f$ showed higher values than it's immediate neighbours each $f \times k$ lags.
2. Alternating pattern – a pattern, where a time-series with seasonality frequency = 2, would have sharper drops (at least twice as fast) in ACF values for each alternating value. A pattern was deemed existing if at least half of ACF value changes exhibited such behaviour.

For companies exhibiting the identified patterns, seasonal-trend decomposition using LOESS (STL) was used to decompose time-series into trend, seasonal, and residual components. Robust weighting was applied with the aim of minimising the effect of outliers. The seasonal frequency component was applied according to the frequency identified using ACF analysis. Trend and residual components were later kept for companies with seasonal patterns (Haitin, Google Colab, 2025).

## Appendix 4. A note on unit-roots and non-stationarity

As mentioned in sub-chapter 2.1, performing FE or RE regression on non-stationary dataset can result in spurious regression, leading to statitstically significant, but meaningless coefficients. For the sakes of academical rigour, unit-root test were performed on all the regressors. To test for stationarity, unit root tests are usually conducted under the null hypothesis that a unit root is present, that is, $X_t \sim I(1)$. This means that the series is non-stationary in levels, but becomes stationary after first differencing (Dickey & Fuller, 1979). $I(1)$ processes are allowed for system GMM, but not for standard regression approaches.

Pesaran CD test (2004) was performed to test the data for cross-sectional dependence. The null hypothesis of weak cross-sectional dependence was rejected for all variables. Based on diagnostics of residuals from CIPS unit-root test (Pesaran, 2007), serial dependence of data was also confirmed.

Given these two problems, panel analysis of nonstationarity in idiosyncratic and common components or PANIC, developed by Bai & Ng (2004), was chosen as a method to reliably estimate the presence of a unit-root in the data. PANIC applies principal component analysis (PCA) to decompose panel data into common factors and idiosyncratic factors, both of which are then tested using standard ADF approach allowing to distinguish between systematic and entity-specific non-stationarity. PANIC relies on a large dimensional framework, that is, $N \to \infty,\ T \to \infty$, and while in this study $N$ is indeed large, $T$ is relatively limited by the original dataset. Adopting PANIC for a small $T$ panel can compromise the reliability of PCA-based factor extraction, leading to misestimated common components. Consequently, the residuals of the ADF test may exhibit serial correlation.

The extended PANIC framework (Bai & Ng, 2010) proposes a solution to this very issue by applying moving block bootstrap (MBB) to the idiosyncratic components, while holding estimated common factors fixed. The procedure then re-estimates PCA based on the bootstrapped distribution and computes an adjusted p-value for the subsequent ADF test. This methodoly was applied to all of the variables used in the model, except normalised and demeaned ESG scores. PCA was adopted with $r$ components, where $r$ denotes the number of retained principal components. Following Bai-Ng (2002), optimal $r$ value was identified using IC1 and IC2 criteria. In cases where an optimal point could not be determined, a cumulative variance threshold of 95%

was used as a fallback method. A total of 499 bootstraps for each variable were estimated. Unit-root testing results are presented below.

Table 10. Results of PANIC analysis on unit-roots in control variables and dependent variable

| Variable | Number of factors (*r*) | Total explained variance (%) | Trend | Common component p-value | Idiosyncratic component p-value |
|---|---|---|---|---|---|
| EBITDA margin | 15 | 95.56 | No | 0.0009 | 1.0000 |
| Debt-to-Assets | 16 | 95.54 | No | 0.0000 | 1.0000 |
| Cash-to-Assets | 19 | 95.16 | No | 0.0000 | 1.0000 |
| ln(totalAssets) | 6 | 95.65 | Yes | 0.0064 | 0.8441 |

Source: author calculations in Python (Haitin, Google Colab, 2025)

Results of PANIC anaysis provide additional ground for the adoption of system GMM, instead of standard regression approach like FE or RE models, as unit-roots have been found in idiosyncratic components of all the variables. ESG scores did not require unit-root testing, since Kalman filter-estimated state processes are usually stationary, or at most trend stationary (Koopman & Durbin, 2013).

## Appendix 5. An honest commentary on the usage of LLMs

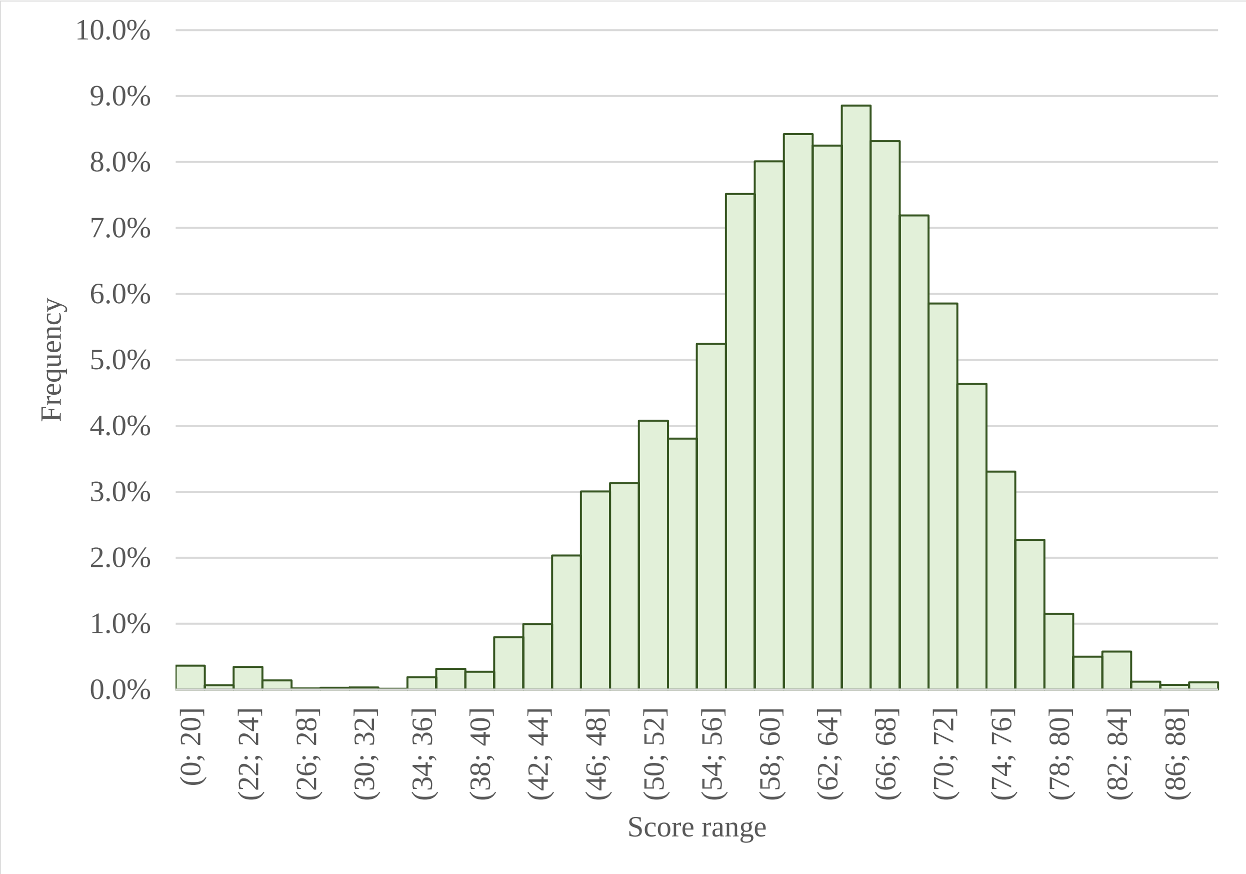
Given the rapid adoption of large language models (LLMs) and their expanding capabilities, the author considers it necessary to transparently disclose their role in the preparation of this thesis. Two LLMs were used: ChatGPT by OpenAI and Claude by Anthropic. While the content, structure, methodology, and interpretations in this work were developed independently by the author, the models were used to assist with technical implementation and language refinement.

Claude was primarily used for writing Python and R code required for the analysis. This enabled the author to apply complex analytical techniques without needing prior coding knowledge, thereby broadening the scope of the analysis. However, despite its usefulness, Claude introduced numerous challenges, such as hallucinating unrequested functions, incorporating fallback methods that delayed problem comprehension, and generating code with errors. After many hours spent revising the code used in this thesis, the limitations of LLMs became apparent. It must be emphasized that without a solid understanding of the underlying logic and reasoning, behind the requests, it becomes much harder to achieve accurate results. Claude often applies general modelling specifications, rules of thumb, or other approaches without considering or attempting to take context into account. As such, it was no exception that a formula needed to be unambiguously defined in order to get code that actually does what is requested.

ChatGPT was used primarily as a text corrector, phrasing assistant, search engine or other similar purposes. Once again, while the benefits of using ChatGPT are substantial, the current model often fails to provide relevant and accurate responses when the prompt is vague or when the user lacks a strong understanding of the subject. Misleading answers, conclusions and ignored context were not exceptions.

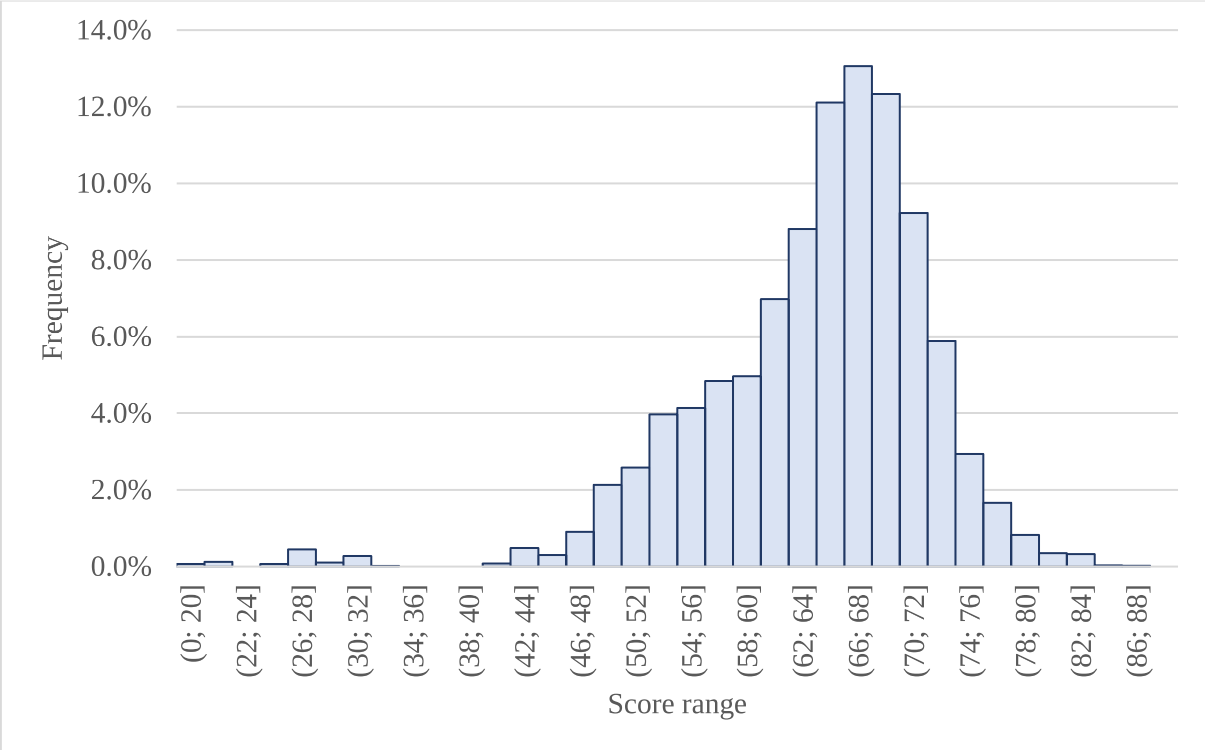
The author is firmly of the opinion that LLMs have a valuable role in academic research. When used thoughtfully, they can expand a researcher's capabilities and save valuable time for researchers. However, caution must be exercised in their use: without a thorough grasp of the topic, the answers provided by LLMs can be highly misleading and without critical evaluation of their answers, academical rigour can be eroded.

## Appendix 6. Distribution of environmental scores



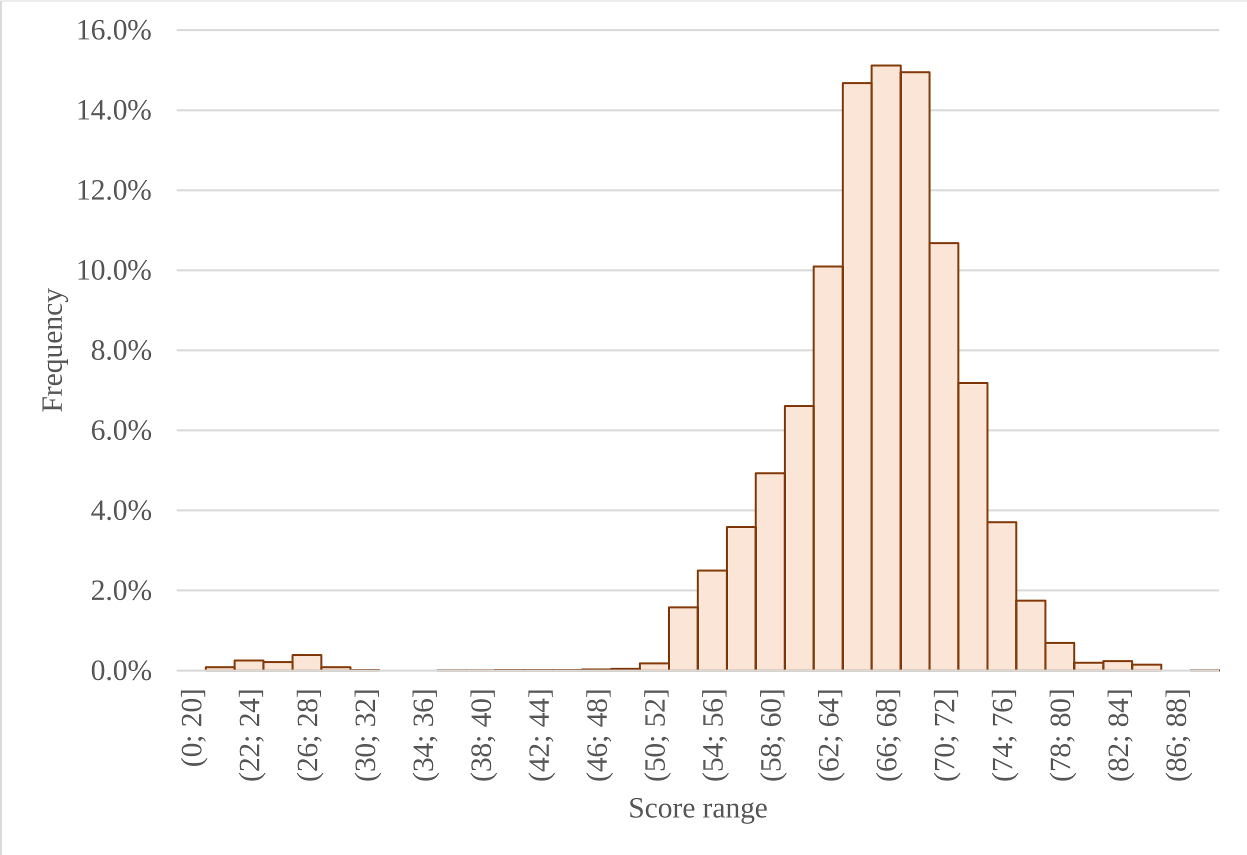

Source: author calculations

## Appendix 7. Distribution of social scores




Source: author calculations

## Appendix 8. Distribution of governance scores




Source: author calculations

# Appendix 9. Non-Exclusive Licence

**Non-Exclusive Licence for the Reproduction and Public Availability of the Thesis**[1]

I, Mikael Haitin

1. Grant Tallinn University of Technology a free licence (non-exclusive licence) for the work I have created, ESG, EARNINGS, AND EARTHQUAKES: A SKEPTICAL ANALYSIS OF RESILIENCE IN MACROECONOMIC CHAOS, supervised by Kaido Kepp,

1.1 for the purposes of reproducing the thesis for archival and electronic publication, including adding it to the digital collection of the Tallinn University of Technology Library, until the expiration of the copyright term;

1.2 for making the thesis publicly accessible via the Tallinn University of Technology's online environment, including the digital collection of the University Library, until the expiration of the copyright term.

2. I acknowledge that the rights specified in clause 1 of this non-exclusive licence remain with the author.

3. I confirm that the granting of this non-exclusive licence does not infringe upon the intellectual property rights or data protection rights of any third parties as provided by law or other legal acts.

---

*/signed digitally/* (15/05/2025)

[1] *The non-exclusive licence does not apply during the period of access restriction established at the student's request and signed by the dean of the faculty, except for the university's right to reproduce the thesis solely for preservation purposes. If the thesis has been created by two or more individuals through joint creative effort, and the co-author(s) or joint author(s) have not granted the defending student permission to reproduce and publish the thesis by a specified deadline in accordance with clauses 1.1 and 1.2 of the non-exclusive licence, then the licence shall not be valid during that specified period.*