Shades of Green: The Bidirectional Dynamic ESG-Return Relationship Across Global Markets

Marcel Philippe Eschwe

Dissertation written under the supervision of professor Zoë Venter.

Dissertation submitted in partial fulfilment of requirements for the MSc in Finance, at the Universidade Católica Portuguesa, 05.09.2025.

Abstract

This study empirically investigates the relationship between ESG scores and annual stock returns in a dynamic, bidirectional setting, addressing conflicting positive, negative, and insignificant links presented in prior literature. It presents robust evidence of a global negative ESG premium associated with the environmental and social dimensions. This effect is driven by highly liquid firms and those with low passive ownership, while it vanishes in emerging markets and in firms with low ESG scores. Further, a reverse impact of returns on ESG exists, which is driven by prior underperformance and materializes through the governance pillar. The results highlight the importance of modeling the ESG-return nexus as a dynamic endogenous system and indicate that the relationship is not uniform but bidirectional, dimension-specific, and mediated by the level of ESG, level of returns, liquidity, ownership, and firm-specific responses to past performance. Overall, the findings help to reconcile conflicting views on the ESG-return nexus.

Keywords: ESG, stock returns, bidirectional impacts, panel vector autoregression (PVAR), global panel

Resumo

Este estudo investiga empiricamente a relação entre os índices ESG e os retornos anuais das ações num ambiente dinâmico e bidirecional, abordando ligações contraditórias positivas, negativas e insignificantes apresentadas na literatura anterior. Apresenta evidências robustas de um prémio ESG negativo global associado às dimensões ambiental e social. Este efeito é impulsionado por empresas altamente líquidas e aquelas com baixa participação passiva, enquanto desaparece nos mercados emergentes e em empresas com baixos índices ESG. Além disso, existe um impacto reverso dos retornos sobre o ESG, impulsionado pelo desempenho inferior anterior e que se materializa através do pilar da governança. Os resultados destacam a importância de modelar a relação entre ESG e retorno como um sistema endógeno dinâmico e indicam que a relação não é uniforme, mas bidirecional, específica por dimensão e mediada pelo nível de ESG, nível de retornos, liquidez, propriedade e respostas específicas da empresa ao desempenho passado. No geral, as conclusões ajudam a conciliar visões contraditórias sobre a relação entre ESG e retorno.

Palavras-chave: ESG, retornos das ações, impactos bidirecionais, autoregressão vetorial de painel (PVAR), painel global

List of Contents

List of tables

List of appendices

List of abbreviations

CAPM - Capital Asset Pricing Model

CD - Coefficient of Determination

CFP - Corporate Financial Performance

CINS - CUSIP International Numbering System

CSP - Corporate Social Performance

CSR - Corporate Social Responsibility

CUSIP - Committee on Uniform Securities Identification Procedures

ESG - Environmental, Social, and Governance

ESGC - Environmental, Social, and Governance Combined Score

FD - First Differencing

FOD - Forward Orthogonal Deviation

GMM - Generalized Method of Moments

GSIA - Global Sustainable Investment Alliance

HQIC - Hannan-Quinn Information Criterion

ICB - Industry Classification Benchmark

IRF - Impulse Response Function

ISIN - International Securities Identification Number

KLD - Kinder, Lydenberg, and Domini (ESG rating provider)

LSEG - London Stock Exchange Group

MMSC - Moment and Model Selection Criteria

MSCI - Morgan Stanley Capital International

NRBV - Natural Resource-Based View

OLS - Ordinary Least Squares

PVAR - Panel Vector Autoregression

PVECM - Panel Vector Error Correction Model

ROW - Rest of World

SASB - Sustainability Accounting Standards Board

SRI - Socially Responsible Investment

VAR - Vector Autoregression

WRDS - Wharton Research Data Services

1 Introduction

Corporate social responsibility (CSR) and its relation to corporate financial performance (CFP) are extensively debated in academia. At the heart of this discussion lies the conflict between Milton Friedman's shareholder theory and Edward Freeman's stakeholder theory. Friedman argues that a firm's only responsibility is to maximize profits, while Freeman emphasizes the importance of considering the interests of all stakeholders for corporate success (Danielson et al., 2008; Freeman, 2010; Friedman, 1970). The clash between the two views intensified academic research on the impact of CSP on CFP (Freeman and McVea, 2001; Ruf et al., 2001).

This debate shaped the field of sustainable finance, which experienced strong growth both in research and practice, supported by global sustainable assets reaching a valuation of $30.3 trillion in 2022 (The Global Sustainable Investment Alliance, 2022). Nevertheless, the precise relationship between ESG factors and equity returns remains ambiguous, as empirical studies consistently report conflicting positive, negative, or insignificant links (e.g. Bolton and Kacperczyk, 2021; Khan et al., 2016; Renneboog et al., 2008).

One possible reason for the diversion of findings is the potential simultaneity of CSP and CRP (Lin et al., 2019). Existing literature primarily evaluates a static link between ESG and returns, often through cross-sectional regressions, event studies, or portfolio construction methods. However, such designs fail to incorporate complex and bidirectional relationships. Only a handful of recent single-country studies employ dynamic modelling frameworks such as vector autoregression (VAR) or panel vector autoregression (PVAR) methodologies, all observing significant bidirectional causality (Lin et al., 2019; Shackleton et al., 2022; van Emous et al., 2021). Still, global evidence remains scarce, sample sizes are often small and short, and CSP varies in definition and measurement. Furthermore, studies rarely explore the dynamics of the underlying ESG measurement itself (Erhart, 2022; Napier et al., 2023; Wedajo et al., 2024). Motivated by these findings, this study explicitly explores the nexus between ESG scores and returns in a dynamic and endogenous framework. The goal is to consider both directional impacts of CSP and CFP, extend the discussion to a global context, and explore common theories surrounding the relation between returns and ESG scores in a bidirectional dynamic setting. To the best of my knowledge, this thesis presents the first study on the dynamic relationship between composite ESG scores, all three ESG dimensions, and stock returns in a global setting, explicitly addressing bidirectional impacts using a PVAR model.

This thesis is based on a large global panel of more than 6,300 listed stocks across 52 countries over the period 2003 to 2024. I estimate PVAR models using annual return and ESG scores as key variables of interest. The results show that the relationship between ESG and returns is complex and highly context-dependent. A composite ESG score carries significant noise, signals of reverse causality, and a global negative impact on returns. This negative premium is driven by the E and S dimensions, while the reverse impact of returns runs through the G dimension. The dynamic is non-linear and asymmetric, as high ESG companies observe a negative premium, while low ESG assets show a positive premium. The reverse return impact runs from companies with prior bad performance. The negative ESG premium impact can be associated with ownership and liquidity, specifically low passive ownership and high stock turnover. The relationship between ESG and return is insignificant in emerging markets, while the return on ESG impact seems to be driven by the U.S.

The findings of this study speak to regulators, investors, and researchers. They highlight the need for more granular ESG data, as composite scores carry significant noise. Further, investors can use the underperformance of high E and S companies and the presence of an optimal ESG level to both boost portfolio ESG scores while allocating capital more efficiently. For research, this thesis presents a clear call towards more dynamic modelling techniques that consider the complex forward and reverse impacts of ESG and returns.

The rest of this thesis is organized as follows: Chapter 2 presents a systematic literature review, the motivational drivers, and the research hypotheses. Chapter 3 details the data sources, the final dataset construction, and presents key summary statistics. In Chapter 4, the methodology and assumptions of the empirical model are discussed and tested. Chapter 5 presents the main findings from the PVAR analysis along with a specific analysis of ESG dimensions and economic mediators. Chapter 6 proceeds to test the robustness of the main findings, while Chapter 7 discusses the study's key limitations and suggests directions for future research. Finally, Chapter 8 provides concluding remarks.

2 Literature review & hypotheses development

This section provides a systematic literature review of the relationship between corporate social responsibility and corporate financial performance, with a specific focus on ESG factors and stock returns. Chapter 2.1 traces the historical development of the field and introduces key terminology. Chapter 2.2 highlights findings from studies exploring the impact of ESG scores on returns. Based on previous findings, Chapter 2.3 presents the identified research gap and derives the research hypothesis driving this thesis.

2.1 History of CSR, CFP, and sustainable finance

Corporate social responsibility (CSR) formally emerged as a distinct field of study based on Howard Bowen's 1953 work, *Social Responsibility of the Businessman*, which explicitly connected business conduct to societal obligations (Bowen, 1953; Carroll, 2008). This evolved further into the concept of corporate social performance (CSP) in the 1970s, and by the early 2000s, the research focus had shifted to empirical links between CSP and corporate financial performance (CFP). A central example is Jones and Murrell (2001), who explored how a firm can use outstanding CSP to create a positive signal to shareholders (Carroll, 2008; Horn and Oehler, 2024).

Within the CSP-CFP discourse lies the field of sustainable finance, which has gained increased attention due to growing regulatory initiatives and climate change concerns. It is characterized by its focus on asset management and the adoption of the investor's perspective on sustainability. Corporate financial performance is consequently often approximated via returns or firm value. The European Union formally recognized sustainable finance as encompassing Environmental, Social, and Governance (ESG) considerations in investment decision-making processes (European Commission, 2021). Nowadays, the field addresses a wide array of topics ranging from carbon finance to ethical investing. However, the fundamental relationship between a firm's ESG characteristics and its financial performance remains among the most critical and contested areas (Bauer et al., 2022; Bolton and Kacperczyk, 2021; Hoffmann and Busch, 2008; Kumar et al., 2025; Lewandowski, 2017; Watson, 2011; Wilson, 1997).

2.2 Dynamics between ESG performance and stock returns

The relationship between ESG and equity returns still lacks a clear consensus. Research findings can be broadly classified into three perspectives: supporting positive, negative, or insignificant links.

The first strand of literature finds that ESG activities can have a positive impact on stock market performance, referred to as *doing well by doing good*. A notable example is Khan et al. (2016), who demonstrate that sustainability investments can be value-enhancing for shareholders if they are focused on material issues. They map MSCI KLD data for approximately 2,300 U.S. firms to construct materiality and immateriality scores and use the scores to form quintile portfolios.

They find that firms with superior materiality investment performance exhibit a 2%-6% top-minus-bottom decile spread after controlling for the Fama-French five-factor model.

Other portfolio-based studies confirm these positive findings. Nagy et al. (2016) apply two high-tracking-error portfolio strategies based on global equity data. They construct a tilt ESG and an ESG momentum portfolio and compare their performance to the MSCI World universe, recording a 1.4% annual outperformance.

Crisis evidence reinforces the positive ESG-return view, arguing that ESG offers downside-risk protection. Albuquerque et al. (2020) assign companies to quintile portfolios based on LSEG and MSCI ES scores to observe their performance in the first quarter of 2020. They show that during the COVID-19 market crash, US firms with high ES ratings observed significantly higher returns, reduced volatility, and higher operating profitability. The resilience of high ES firms is associated with higher customer and investor loyalty. Similarly, Broadstock et al. (2021) replicate the ESG-resilience phenomenon for China's CSI300 constituents, using ESG scores provided by Syntao. Their findings indicate that high ESG stocks experienced lower financial risk and return outperformance during the COVID-19 crisis, although this effect diminishes outside of crisis periods.

Escobar-Saldívar et al. (2025) further suggest that investors can capture a momentum-style ESG premium by targeting firms actively improving sustainability practices. They examine 3,856 U.S. equities over 2002-2022, combining quarterly ESG ratings with their year-over-year change in multi-factor panel regressions to assess impacts on returns and volatility. They find that higher static ESG scores correspond to lower long-term returns and higher risk, whereas positive ESG momentum delivers immediate excess returns and reduces volatility. However, in the long run, their findings indicate that higher ESG scores are associated with lower stock returns.

The finding that ESG efforts are associated with lower future stock returns presents the second view on CSR and its impact on financial performance, often framed as *underperforming by doing good*.

Hong and Kacperczyk (2009) were among the first to provide evidence for abnormal returns when investing in vice companies. They analyze alcohol, tobacco, and gambling stocks in the U.S. between 1965 and 2006, finding that these sin stocks observe higher expected returns when compared to stocks of similar characteristics, even after accounting for common risk factors. The authors show that sin stocks are held in smaller quantities by norm-constrained investors and receive less analyst coverage than comparable companies, presenting one possible explanation of the sin premium.

Going beyond a sin alpha, Bolton and Kacperczyk (2021) provide evidence of outperformance by stocks with high carbon emissions, as investors price their respective carbon risk. Using Trucost Scope 1-3 emissions for U.S. equities between 2005 and 2017, they construct monthly carbon-sorted portfolios and uncover a carbon-emission premium of roughly 3-4 % per year. Time-series regressions show that the premium persists after factor controls, disappears for emission intensity, and is stronger for firms with rising emissions. They conclude that limited risk sharing cannot explain the carbon premium, which stands in contrast to the previously discussed sin premium.

Pedersen et al. (2021) provide a formal theoretical framework for the ESG-return relationship by proposing an ESG-adjusted CAPM. They test their theory on constituents of the MSCI ACWI between 1991 and 2019, using pooled OLS, Fama-MacBeth regressions, and high-ESG minus low-ESG portfolio analysis. They find further support for a carbon and sin premium, while their evidence broadly suggests an inverse relationship between ESG scores and stock performance. Low-ESG portfolios tend to outperform high-ESG portfolios, consistent with their theory of an ESG-efficient frontier where investors trade off returns for high ESG exposure.

Further, studies confirm this relationship and associate it with liquidity and investor preference. Luo (2022) ranks equities from the FTSE All Share Index by LSEG ESG scores for the period 2003 to 2020 and shows that the lowest-ESG quintile outperforms the highest by roughly 0.5% per month. The effect is stable after adjusting for several risk factors, and the excess return is driven by the environment and social pillars. He further posits that low liquidity facilitates the ESG premium. Ciciretti et al. (2023) ask whether the negative global ESG premium is driven by the lower risk of high ESG companies or by investor preference for ESG top performers. Using worldwide equity data for 2004-2018, they separate firm-level ESG betas from ESG characteristics in Fama-MacBeth regressions and find that investor preference for high ESG assets can be associated with lower expected returns, while ESG betas explain almost none of the spread. They imply that investors are willing to accept lower returns to hold sustainable stocks.

The third cluster finds no systematic ESG return impact or relationship, which is referred to as *no effect by doing good*.

Renneboog et al. (2008) study almost the entire universe of ethical and socially responsible investment (SRI) funds across Europe, North America, and the Asia-Pacific. Relative to domestic Fama-French-Carhart benchmark portfolios, they find negative annual four-factor alphas of 2%-6%. However, when considering fees and matching performance with conventional funds, they find no statistically significant evidence for the underperformance of SRI funds. They argue that underperformance, if present, could be attributed to high SRI screening intensity constraining risk-return optimization.

Alves et al. (2023) analyze more than 16,000 stocks across 48 countries between 2001 and 2020. They use ESG data from several different rating providers and apply monthly Fama-MacBeth cross-sectional regressions on ESG ratings and returns. They find no consistent relation between ESG and global stock returns.

Halbritter and Dorfleitner (2015) revisit earlier U.S. evidence by extending the sample from 1991-2012 and comparing ESG scores by ASSET4, Bloomberg, and KLD. In high-minus-low portfolios, they initially observe significant abnormal returns, yet once the Carhart four-factor model and cross-sectional Fama-MacBeth regressions are applied, the difference between high- and low-ESG stocks turns insignificant. Moreover, significance flips across providers and sub-periods, leading the authors to conclude that the ESG premium evaporates when longer horizons and multiple data sources are considered.

As the preceding review demonstrates, academic literature remains fragmented and often ambiguous. Yet, ESG considerations are central to mainstream investment strategies and societal well-being, highlighting the importance of further research (The Global Sustainable Investment Alliance, 2022).

2.3 Research gaps and hypotheses development

One reason for the empirical fragmentation and the profound lack of consensus is the common usage of static, unidirectional methodologies, such as studies using simple OLS regressions or portfolio construction approaches as presented in Chapter 3.2. Most studies test whether ESG scores predict returns, a framework that can produce misleading results by ignoring the well-known issue of reverse causality, whereby financial performance itself may drive a firm's ESG activities (Lin et al., 2019; Wedajo et al., 2024). Overall, unidirectional models commonly employed by the presented studies may be insufficient to capture the full, complex, dynamic, and potentially bidirectional nature of the ESG-return relationship. Furthermore, the inherent divergence and reliability issues in ESG data itself necessitate research into how these metrics interact with stock returns in a dynamic, bidirectional setting. This is supported by more recent evidence of dynamic studies, which show significant bidirectional and reverse impact within the CSP-CFP relationship (e.g., Lin et al., 2019; Qureshi et al., 2021; Shackleton et al., 2022; Yu et al., 2022, 2024).

By employing a dynamic PVAR model, this study moves beyond static analysis to explicitly model feedback loops. This approach can help reconcile the conflicting positive, negative, and neutral findings by revealing a more complex, endogenous system. Furthermore, this analysis is conducted on a broad global sample, providing more generalizable evidence beyond the often US-centric focus of prior work. This approach leads to the first research hypothesis:

H1: A dynamic, intertemporal, bidirectional relationship exists between a firm's annual stock return and its ESG score proxy.

A second source of the conflicting findings is the use of aggregated ESG scores. A composite score can mask the distinct, and possibly diverging, relationships of its underlying pillars, introducing noise into the analysis. The literature already points to differing dimensional impacts, with studies like Luo (2022) or Pederson et al. (2021) finding different effects for the E, S, and G dimensions. While it is common to deconstruct ESG in static analysis, this study is among the first to apply a dimensional breakdown in a PVAR setting to better understand the ESG-return nexus in a more complex endogenous system. Therefore, the second hypothesis is:

H2: The bidirectional relationship varies in direction and magnitude across the individual Environmental (E), Social (S), and Governance (G) dimensions.

Finally, the literature often points to the investor demand channel as a key driver of ESG-related return patterns (Darolles et al., 2023; Luo, 2022; Pedersen et al., 2019). While investor preferences are generally hard to test, they can be approximated by equity characteristics like ownership structures. For instance, high institutional ownership may reflect the influence of norm-constrained investors as presented by Hong & Kacperczyk (2009). Additional studies argue that active institutional investors express their ESG preferences through engagement and portfolio allocation, suggesting that the premium should be stronger where passive ownership is lower (Dyck et al., 2019; Heath et al., 2022).

Further, liquidity has recently been found to be a mediating channel in the ESG-return relationship. Studies show that strong ESG performance can be associated with higher liquidity, which, according to classic asset-pricing theory, might lead to lower expected returns (Amihud, 2002; Wang et al., 2023). However, the evidence on liquidity is not uniform, with some studies finding the premium is more pronounced in illiquid stocks (Luo, 2022). To assess the influence of ownership structures and liquidity within a dynamic framework, the third hypothesis is:

H3: The dynamic ESG-return relationship is significantly mediated by stock liquidity and ownership structure.

3 Data and sample

The following chapter describes the selection, processing, and preparation of the data underlying this thesis. Chapter 3.1 describes the data sources, the merging procedures, and the data preprocessing to obtain a combined raw dataset for variable construction. Chapter 3.2 outlines the exact data items retrieved and presents how key variables were constructed. Further, it provides an overview of the data cleaning procedures applied to derive the final panel. Lastly, Chapter 3.3 presents key summary statistics of the final dataset.

3.1 Data sources and dataset construction

This study utilizes a survivorship-bias-free dataset covering the period from 2003 to 2024, constructed by merging three primary sources: Compustat for firm-level accounting data, Kenneth R. French Data Library for global risk-free rates and market risk premia, and LSEG Datastream for firm-level ESG and market data.

To overcome the small-sample-size and US-centric limitations of dynamic CSP-CFP studies, this research began by constructing a large, global dataset (Lin et al., 2019; Pastor et al., 2022; Useche et al., 2024). This required a targeted approach, as LSEG Datastream does not permit bulk downloads of its entire ESG universe and instead relies on pre-defined company identifiers for data extraction. Leveraging LSEG Workspace's Advanced Search app, I identified a universe of 11,903 stocks with at least one year of ESG data. The listings are restricted to primary issues and primary quotes to ensure data integrity and avoid cross-listings. This list of unique firms then served as the basis for downloading raw data from both LSEG and Compustat. After removing duplicates and missing data points, the preliminary LSEG dataset included 210,312 firm-year observations across 11,903 firms, while the preliminary Compustat dataset covered 212,230 firm-year observations for 11,222 firms.

The next critical step was to merge these two datasets. This presented a challenge, as Compustat accounting data is based on each company's fiscal year, while LSEG market data and ESG data are based on the specified database extraction data. Thus, merging accounting and market data requires fiscal year matching. Combining the standard approaches of Hayashi and Jagannathan (1990), Janakiraman et al. (1992), and Fama and French (1993) to handle calendar and fiscal year mismatches, fiscal closures in March, June, September, or December were retained, and LSEG data were realigned to the corresponding twelve-month accounting period. Roughly 95% of initially identified companies adhere to these criteria, while 550 companies with differing or changing fiscal year ends were dropped. Then, the Compustat and LSEG data are merged using ISIN and fiscal year-end variables as unique identifiers, resulting in a panel of 10,557 firms with 177,029 firm-year observations. Finally, Kenneth R. French data were merged with the combined Compustat-LSEG data, which did not reduce the sample size. The resulting database presents the foundation for all variable calculations and cleaning procedures outlined in the following chapter.

3.2 Data cleaning and variable construction

This section details the process of transforming the raw database into the final panel by describing the three key filters and cleaning procedures applied. This is followed by two subsections, 3.2.1 and 3.2.2, which describe the definitions and calculations for the variables employed in the analysis. An overview of all variables underlying this thesis, including their origin, construction, and academic reference, can be found in **Appendix 1**.

The first filter removed all panels missing any of the baseline variables (see **Appendix 1**, Baseline variables). Second, firms with fewer than three consecutive fiscal-year observations were removed to fulfill the requirements of the dynamic modeling approach underlying this thesis, outlined in Chapter 4. Third, to ensure that regional factor returns are not driven by small cross-sections, countries represented by fewer than ten distinct firms were excluded. Comparable minimum-size rules are common when constructing global samples (Fama and French, 1998; van Emous et al., 2021).

Finally, to mitigate the influence of extreme values, all variables except returns were winsorized at the 1st and 99th percentiles, a common practice in similar studies (Bolton and Kacperczyk, 2021; Drempetic et al., 2020; Trumpp and Guenther, 2017).

After applying all three filters, the final sample covers the period from 2003 to 2024 and includes 6,379 firms with 50,460 firm-year observations. The most significant drop in firm count resulted from missing fundamental data required to construct control variables and from sparse ESG coverage, especially in the early years of the sample. A summary of the dataset construction and cleaning steps is provided in **Appendix 2**.

3.2.1 Return and ESG measurement

Annual returns, further denoted as $R_{i,t}$, constitute returns during the fiscal year period, calculated as follows:

$$R_{i,t} = \frac{P_{i,t}}{P_{i,t-1}} - 1 \tag{1}$$

where $i = 1, 2, \ldots, N$ represents each firm in the panel at the fiscal year $t = 1, 2, \ldots, T$. T is defined as the total number of years, and N represents the total number of firms in the sample. Returns are raw returns, since this study does not rely on or benefit from the time-adding properties of logarithmic returns.

Abnormal returns, denoted as $AR_{i,t}$, are defined as market-adjusted returns estimated with the capital asset pricing model (CAPM) (Sharpe, 1964). Subsequently, they were determined as follows:

$$AR_{i,t} = R_{i,t} - (RF_t + \beta_{i,t} * MKT_{c,t}) \tag{2}$$

Where $R_{i,t}$ represents the return as defined in Equation (1), RF_t represents the one-year U.S. Treasury bill rate in year t, $\beta_{i,t}$ represents the leverage beta of firm i at time t, and $MKT_{c,i}$ represents the market factor of region c in year t. The variables are calculated as follows: $\beta_{i,t}$ is provided by LSEG and estimated as the slope coefficient from an OLS regression of monthly excess returns on the corresponding market excess returns over a rolling five-year window. This method aligns with common determinations of beta in a CAPM context (Jensen et al., 2006). Data for the market factor $MKT_{c,t}$, and global risk-free rates RF_t were obtained from the Kenneth R. French library. $MKT_{c,t}$ represents the value-weighted return of a regional portfolio minus the one-month U.S. Treasury bill rate and is aggregated over the calendar year. Here, $c = 1, 2, \ldots, C$ denotes the following five regions: North America, Europe, Japan, Asia-Pacific, and Emerging markets/RoW, as defined by Fama and French[1] (2012). To accurately match Kenneth R. French's global data, each firm in the sample was assigned to one of the five geographic regions, based on headquarters location, following the Fama-French classification. Whenever the classification does not specify a country included in the sample, the individual country is hand-mapped based on geographic location. **Appendix 3** outlines the country-region mapping underlying this thesis.

---

[1] Emerging market countries follow the classification presented in the Kenneth R. French Data Library: https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/f-f_5emerging.html

The final return metric used in this study is industry-adjusted returns. This is calculated for each firm-year return observation by subtracting the average annual return of its corresponding ICB industry.

ESG scores were obtained from LSEG. Their ESG database covers more than 90% of global market capitalization and offers historical data starting in 2002 for almost 16,000 public and private companies. Within this database, ESG performance is summarized through percentile rank scores ranging from 0 to 100, which are derived from over 870 underlying indicators across ten thematic categories. The 10 thematic areas are organized into separate pillar scores for Environment (E), Social (S), and Governance (G). These pillars are then aggregated to form a composite ESG score. In addition, an ESG Combined (ESGC) score is available, which adjusts the standard ESG evaluation for major corporate sustainability controversies (London Stock Exchange Group, 2025, 2024). Specifically, I obtained ESG, ESGC, E, S, and G scores for each firm. All the scores mentioned above were retained in their original levels, ranging from 1 to 100, without any alterations.

3.2.2 Market and accounting variables

Throughout this paragraph, Datastream/Worldscope mnemonics appear in **bold** within brackets, whereas Compustat codes drawn from WRDS appear in *italic* within brackets.

Control variables used in this study follow similar research papers on the CSP-CFP relationship, such as Bolton and Kacperczyk (2021) and Shackleton et al. (2022). $LEVERAGE_{i,t}$ is defined as long-term debt (*DLTT*) plus short-term debt (*DLC)* over total assets (*AT*). $PROFITABILITY_{i,t}$ is proxied by gross profit, calculated as sales (*REVT*) minus cost of goods sold (*COGS*), and scaled by total assets (*AT*) to account for differences in size. $LOGSIZE_{i,t}$ represents the natural logarithm of firm i's market capitalization in U.S. dollars (**MVC**). $LOGB/M_{i,t}$ is the natural logarithm of the book-to-market ratio, calculated by dividing a company's book value of equity by its market capitalization (**MVC**). The book value of equity follows the definition of Fama and French (1993). It is defined as the sum of the book value of shareholders' equity, deferred taxes, and investment tax credit, less the book value of preferred stock. Calculation of deferred taxes, investment tax credit, and preferred stock value follows the approach of Novy-Marx (2013). Deferred taxes include the total deferred tax balance plus investment tax credit (*TXDITC)*. If missing, it includes either the sum or single availability of deferred taxes (*TXDB*) and investment tax credit (*ITCB)*. Preferred stock is measured by redemption (*PSTKRV)*, liquidation (*PSTKL)*, or par value (*PSTK*), prioritized in that order based on availability.

$CASHRATIO_{i,t}$ includes cash and short-term investments (*CHE*) over total assets (*AT*). $INVESTMENT_{i,t}$ represents the growth rate of total assets (AT) in percent. $DIVRATIO_{i,t}$ are a firm's cash dividends (*DVC + DVP*) scaled by total assets (AT). $LOGAGE_{i,t}$ is the natural log of the difference of a firm's incorporation year (**WC18273**) and the fiscal year-end date.

Two additional metrics were used in this study to facilitate research on Hypothesis 3. $LIQUIDITY_{i,t}$ represents the turnover ratio following Gabrielsen et al. (2011), and is calculated as the total number of shares traded annually (**VO**) divided by the average number of shares outstanding (**NOSH**). Finally, ownership is measured using $PASSIVEHOLD_{i,t}$, defined as the percentage of shares held by passive owners (**NOSHSP**). Passive holdings represent ownership by funds that do not actively manage their portfolio allocation.

An overview of all variables can be found in **Appendix 1**.

3.3 Summary statistics and data description

Table 1 presents the descriptive statistics for the final dataset, which comprises 50,460 firm-year observations from 6,379 unique firms between 2003 and 2024.

The average ESG score is 46.95, while the average return is 12%. E, S, and G are in a similar range as expected. Notably, the Governance pillar has a comparatively lower correlation (0.66) compared to the other dimensional pillars. The correlations further indicate that ESG scores are negatively associated with returns, liquidity, cash holdings, and investment. Overall, the values for key financial variables are comparable in magnitude to those in recent literature (e.g. Bolton and Kacperczyk, 2021; Shackleton et al., 2022). The ownership variable *PASSIVEHOLD* is not available for the full sample, which reflects the limited availability of global ownership data. However, as it is reserved for subsequent sample splits rather than the primary analysis, this data limitation does not affect the main results.

The distribution of firms across industries, detailed in **Table 2** Panel A is well-diversified across all major ICB sectors, with the largest concentrations in Industrials, Consumer Discretionary, and Health Care. I include firms in the financial industry to maximize sample size. Chapter 8 will show that the results are robust to the exclusion of financial firms, which are often a concern in financial research (Deng et al., 2013; Lins et al., 2016).

Table 1: Summary statistics

This table reports descriptive statistics for all variables in the final dataset over the sample period from January 2003 to December 2024. Reported statistics include the number of observations, mean, standard deviation, minimum, maximum, and Pearson correlation coefficients. All variables are defined in **Appendix 1** and winsorized at the 1% and 99% levels, except returns. The sample comprises 50,460 firm-year observations from 6,379 unique firms.

Variable	N	Mean	Std	Min	Max	Correlation
ESG Score	50,460	46.95	20.53	0.08	95.57	
E Score	50,460	40.97	28.02	0.00	99.13	0.86***
S Score	50,460	47.24	24.15	0.05	98.51	0.90***
G Score	50,460	51.38	22.36	0.04	99.49	0.66***
ESGC Score	50,460	45.36	19.46	0.08	94.89	0.96***
Return	50,460	0.12	0.58	-0.99	35.65	-0.03***
LOGSIZE	50,460	8.17	1.54	3.15	12.50	0.43***
LOGB/M	50,460	1.49	0.97	-2.49	3.72	0.00
PROFITABILITY	50,460	0.26	0.19	-0.42	2.00	0.05***
INVESTMENT	50,460	0.10	0.25	-0.46	2.03	-0.07***
LEVERAGE	50,460	0.25	0.17	0.00	0.97	0.09***
CASHRATIO	50,460	0.16	0.16	0.00	0.93	-0.16***
DIVRATIO	50,460	0.02	0.03	0.00	0.18	0.07***
LOGAGE	50,460	3.32	0.86	0.00	4.90	0.21***
LIQUIDITY	50,460	2.02	5.25	0.00	160.39	-0.04***
PASSIVEHOLD	29,813	29.32	23.00	0.00	100.00	0.02***

Furthermore, **Table 2** Panel B shows that sample coverage expands significantly over time. The number of annual observations grows from just over 300 in the early 2000s to more than 5,000 in 2023, essentially driven by the availability of ESG data.

Finally, an overview of firm distribution across regions is presented in **Table 3**. In contrast to many U.S.-centric studies, this dataset is globally diversified, with North America and the Emerging Markets/RoW being the most represented regions. A country-level breakdown can be seen in **Appendix 3**, with the United States having the most significant representation (1,865 firms), followed by China (797 firms), India (513 firms), and Japan (379 firms).

Table 2: Sample distribution across industries and years

This table reports the distribution of firms and observations by industry and year for the final dataset. Panel A shows the number of firms by ICB industry classification. Panel B reports the yearly distribution of firm-year observations (N) from 2003 to 2024.

Panel A

ICB industry code	Industry name	Firms
10	Technology	671
15	Telecommunications	213
20	Health Care	864
30	Financials	125
35	Real Estate	69
40	Consumer Discretionary	1,144
45	Consumer Staples	490
50	Industrials	1,426
55	Basic Materials	685
60	Energy	378
65	Utilities	314

Panel B

Year	N	Year	N
2003	322	2014	1,974
2004	329	2015	2,032
2005	649	2016	2,384
2006	862	2017	2,737
2007	894	2018	3,254
2008	996	2019	3,670
2009	1217	2020	4,327
2010	1403	2021	4,967
2011	1720	2022	5,411
2012	1821	2023	5,623
2013	1907	2024	1,961

Table 3: Sample distribution across regions

This table reports the distribution of firms and firm-year observations (N) across regions. Regions are assigned following the Fama and French (2012) mapping, supplemented by the Kenneth R. French data library mapping for emerging markets. The detailed mapping is provided in **Appendix 3**.

Region	Firms	N
Japan	379	5,702
North-America	2,134	16,763
Europe	1,206	11,317
Asia-Pacific (ex. Japan)	274	2,696
Emerging/RoW	2,386	13,982

4. Methodology

This chapter presents the reasoning, methodology, and empirical tests used to derive and estimate the final PVAR system. Chapter 4.1 explains the economic rationale and motivation behind the modeling choices, Chapter 4.2 provides the numerical derivation of the final model specifications, and Chapter 4.3 presents the empirical test applied to evaluate and validate the appropriateness of the selected model specifications.

4.1 Econometric motivation

It has been established that most of the empirical work on the ESG-return nexus relies on static designs, such as cross-sectional portfolio sorts (Khan et al., 2016; Luo, 2022), Fama-MacBeth regressions (Pedersen et al., 2021) or long-short portfolio tests (Hong and Kacperczyk, 2009). To overcome previous limitations and empirically disentangle the effect of ESG and returns requires a framework that (i) treats ESG performance and stock-market performance as jointly endogenous, (ii) explicitly addresses reverse causality, (iii) can follow feedback across time, and (iv) handles heterogeneity in a cross-sectional setting. Historically, researchers relied on multivariate simultaneous-equation systems, and, later, on time-series vector autoregressions (VAR) as pioneered by Sims (1980). Yet both approaches struggle when the time dimension is short relative to the cross-section, thus failing requirement (iv) (Qureshi et al., 2021). Panel-VAR (PVAR) models overcome this limitation by allowing each variable in a multivariate system to be endogenous while adding a cross-sectional dimension to the representation. Thus, PVAR models are essentially VAR models that extend the standard VAR by incorporating several units, accounting for a panel data structure. Given that they are ideal to fulfill criteria (i)-(iv), PVAR modelling serves as the starting point of my empirical approach (Canova and Ciccarelli, 2013). Such designs require careful selection procedures and calibration to accurately fulfill data requirements. Thus, the following derives the final model specifications and the considerations made, closely following literature on PVAR methodologies as well as similar studies on the CSP-CFP nexus.

4.2 Baseline PVAR specification

Equation (3) represents the starting point of my model and shows the general k-variate panel-VAR specification of lag-order p, following the broad notation of Abrigo and Love (2016):

$$Y_{i,t} = \Sigma_{j=1}^{p} A_j Y_{i,t-j} + BX_{i,t-1} + \lambda_t + f_i + \varepsilon_{i,t} \tag{3}$$

where $i = 1, \ldots, N$ indexes firms and $t = 1, \ldots, T$ indexes years. $Y_{i,t}$ represents the vector of dependent variables with the dimensions 1 x k, with k representing the number of dependent variables. $X_{i,t}$ represents a vector including the number of exogenous control variables l, with the dimensions 1 x l. Each A_j is a k x k autoregressive coefficient matrix linking lag j of every endogenous variable to its current values. B represents a 1 x l matrix with the coefficients on the lagged exogenous controls $X_{i,t}$. Both B and A_j include the parameters of interest that need to be estimated for the PVAR equation system. Further, λ_t and f_i represent time and firm fixed effects, respectively. Lastly, $\varepsilon_{i,t}$ is a vector that includes error terms of dimensions 1 x k. The error vector is assumed to be a multivariate white noise process, characterized by a zero mean, a constant contemporaneous covariance matrix, and no serial correlation (Abrigo and Love, 2016).

Equation (3) is flexible. Its dimension expands with the number of variables *k,* and its dynamics deepen with additional lags *p*. Contrary to the parameters *i* and *t,* which are defined by the panel size, *k* and *p* must be chosen explicitly for each PVAR specification.

For the baseline analysis of this thesis, $k = 2$, and Y_{it} is denoted as[2]:

$$Y_{i,t} = \begin{pmatrix} R_{i,t} \\ ESG_{i,t} \end{pmatrix} \tag{4}$$

where $R_{i,t}$ represents the buy-and-hold return during fiscal year *t* of firm *i*, as defined by Equation (1), while $ESG_{i,t}$ denotes the ESG score of firm *i* in fiscal year *t*. Further, as will be demonstrated in Chapter 4.3, the favored lag order for my PVAR specification is $p = 1$. Thus, Equation (3) can be unfolded into a two-equation system as follows:

$$\mathrm{R_{i,t}} = \alpha_1 \mathrm{R_{i,t-1}} + \beta_1 \mathrm{ESG_{i,t-1}} + \gamma_1' \mathrm{X_{i,t-1}} + \lambda_{1_t} + f_{1_i} + \varepsilon_{1_{i,t}} \tag{5}$$

$$\mathrm{ESG_{i,t}} = \mu_1 \mathrm{R_{i,t-1}} + \omega_1 \mathrm{ESG_{i,t-1}} + \gamma_2' \mathrm{X_{i,t-1}} + \lambda_{2_t} + f_{2_i} + \varepsilon_{2_{i,t}} \tag{6}$$

In this PVAR system, the coefficients α_1, β_1, and the vector of coefficients γ_1' are obtained by regressing current returns $R_{i,t}$ on the one-year-lagged values of return ($R_{i,t-1}$), the ESG score ($ESG_{i,t-1}$)*,* and the exogenous controls ($X_{i,t-1}$), while accounting for both time (λ_{1_t}) and firm fixed effects (f_{1_i}). Symmetrically, the coefficients μ_1*,* ω_1*,* and the coefficient vector γ_2' are obtained by regressing the current ESG score ($ESG_{i,t}$) on one-year-lagged values of returns ($R_{i,t-1}$), the ESG score ($ESG_{i,t-1}$), and the same set of controls ($X_{i,t-1}$), again accounting for time and firm fixed effects. Finally, $\varepsilon_{1_{i,t}}$ and $\varepsilon_{2_{i,t}}$ represent the system's white noise disturbances. These are assumed to satisfy the orthogonality condition of strict exogeneity (Shackleton et al., 2022). This means the error terms are uncorrelated with all past values of the endogenous variables in the system, or explicitly, for all $s < t$:

$$\mathrm{E}(\varepsilon_{1_{i,t}} \cdot Y_{i,s}') = 0 \tag{7}$$

[2] Note that both the return proxy and the ESG proxy will later be replaced by other measures for returns and ESG. All measures were defined previously in Chapter 3, and the choice of variables will be explicitly mentioned in the respective result section. For ease of notation, I only show the baseline specification in Equation (4).

$$E(\varepsilon_{2_{i,t}} \cdot Y'_{i,s}) = 0 \tag{8}$$

While these orthogonality conditions are foundational, they also introduce a significant econometric challenge due to the model's structure. The inclusion of lagged dependent variables as regressors alongside unobserved firm-fixed effects creates a dynamic panel bias, also known as Nickell bias (Nickell, 1981). This renders standard OLS estimates inconsistent, especially for large N and small T samples (Abrigo and Love, 2016; Lin et al., 2019). To address this, a two-step approach is employed, consistent with existing literature: initially, a data transformation is performed to remove firm-fixed effects, and subsequently, the transformed model is estimated using the Generalized Method of Moments (GMM).

The first step transforms Equations (5) and (6) by applying either First Differencing (FD) or Forward Orthogonal Deviation (FOD). FD, proposed by Anderson and Hsiao (1982), transforms each variable in the system of equations by subtracting its previous value from its current one. For the ESG score, this would result in $ESG_{i,t}$ - $ESG_{i,t-1}$. While this effectively removes firm fixed effects, it has a significant drawback in panels with gaps, as it magnifies the impact of missing observations and leads to a reduction in sample size. Consequently, since this study relies on an unbalanced panel, this thesis uses the FOD transformation as proposed by Arellano and Bover (1995). FOD subtracts the firm-specific average of all available future observations from each variable. This approach is robust to gaps in the panel and minimizes data loss (Roodman, 2009; Shackleton et al., 2022).

After Equations (5) and (6) are transformed using FOD, the standard two-step GMM approach is used to estimate the coefficients. Conceptually, GMM works by using a set of instrumental variables that are assumed to be uncorrelated with the error term. In this framework, the instruments used for the transformed equation are the lagged levels of the dependent variables (e.g., $R_{i,t-2}$, $ESG_{i,t-2}$). The validity of these instruments rests on the orthogonal assumptions stated in Equations (7) and (8). By using these internal instruments drawn from the panel's history, GMM can eliminate endogeneity issues and generate consistent coefficient estimates in a PVAR setting (Holtz-Eakin et al., 1988; Roodman, 2009; Yu et al., 2022).

4.3 Diagnostic testing

The model specifications outlined in Chapter 4.2 were derived by several testing procedures. First, I test if the key dependent variables, ESG score and annual return, contain a unit root. If these series contain a unit root, that is, they are non-stationary and exhibit a random-walk drift, then the system should be specified as a panel vector error-correction model (PVECM), as this can correct for non-stationarity and cointegration (Engle and Granger, 2015). The Phillips-Perron test is employed to examine panel unit-root properties, as it accommodates unbalanced panels with data gaps. The null hypothesis of the Phillips-Perron test states that all panels contain unit roots. As reported in **Table 4**, the null hypothesis for both the ESG score and the annual return is strongly rejected at 1% significance, implying that at least one panel is stationary. If non-stationary data were found, then the correct way to proceed is to run panel cointegration tests and apply a PVECM. Since this is not the case, variables are not cointegrated by definition, and the PVAR model remains the right choice for this thesis (Binder et al., 2005; Kumar Mandal and Madheswaran, 2010; Yu et al., 2022).

Table 4: Stationarity test of ESG and Return

This table reports Fisher-type Phillips-Perron panel unit root tests for the variables ESG and Return over the full sample period (2003-2024). The test evaluates the null hypothesis that all panels contain a unit root. Reported statistics are based on four transformations: inverse chi-squared, inverse normal, inverse logit t, and the modified inverse chi-squared. Statistical significance is indicated by ***, **, *, denoting the 1%, 5%, and 10% levels, respectively.

Methodology	Statistic	*ESG*	*Return*
Phillips-Perron	Inverse	31,400***	69,900***
	Inverse normal	-34.15***	-171.40***
	Inverse logit t	-78.81***	-266.64***
	Modified inv. chi-squared	135.70***	393.89***

Second, selecting an appropriate lag order is crucial in GMM estimated PVARs, because every additional lag expands the number of instruments, a proliferation that can dilute Hansen's J test and overfit endogenous dynamics. To keep the instrument set parsimonious and choose the correct PVAR lag order, I follow Andrews and Lu (2001) and compute three moment and model selection criteria (MMSC): an AIC-type measure that balances fit and complexity with a constant penalty, a BIC-type measure that applies a heavier penalty that grows with the logarithm of the panel size and therefore favors the most compact model, and an HQIC-type measure based on a log-log penalty. Each statistic is derived from Hansen's J value, which assesses the overall validity of the over-identifying restrictions and thus tests for the appropriateness of instruments (Abrigo and Love, 2016). Alongside the MMSCs, **Table 5** reports the coefficient of determination (CD), a goodness-of-fit measure indicating the variation captured by the PVAR model, and the J statistic with its p-value, where a high p-value indicates that the instrument set is valid. For the PVAR model specified in Equations (5) and (6), I estimate the system allowing up to three lags.

Table 5: Lag order selection for PVAR

This table reports lag order selection statistics for the PVAR model specified in Equations (5) and (6), using ESG scores and annual stock returns over the sample period 2003-2024. Reported measures include the coefficient of determination (CD), Hansen's J statistic with its corresponding p-value, and the moment and model selection criteria (MMSC) proposed by Andrews and Lu (2001): the modified Bayesian information criterion (MBIC), the modified Akaike information criterion (MAIC), and the modified Hannan-Quinn information criterion (MQIC). Results are presented for specifications with one, two, and three lags.

Lag Order	CD	J	J p-value	MBIC	MAIC	MQIC
Lag 1	0.979	20.08	0.068	-135.14	-11.92	-52.64
Lag 2	0.971	41.73	0.003	-152.29	1.73	-49.16
Lag 3	0.971	79.61	0.000	-153.22	31.61	-29.46

All three MMSCs reach their minimum at a single lag, indicating that the first-order PVAR model is the preferred option. The Hansen J test p-value increases as lag depth decreases, and the overall model fit, proxied by the CD statistic, also improves with shorter lags. By contrast, the fit of the ESG scores deteriorates with deeper lag structures, hinting at the limited quality of the underlying ESG data. The Hansen J test is not rejected at the 1% and 5% significance levels for lag order 1. Nonetheless, to ensure robustness, I estimated the PVAR with longer instrument lags, which produced qualitatively similar results. However, consistent with the Hansen J outcomes, deeper lags introduce additional noise in the return and ESG estimates. I therefore adopt a just-identified system of lag order 1 as the preferred specification. Given these econometric limitations, this thesis emphasizes the consistency of the results, the theoretical framework, alignment with the literature, and multiple robustness checks as further bases for its conclusions.

Finally, the stability of the one-lag PVAR specification is assessed by inspecting the eigenvalues of the companion matrix. Following Lütkepohl (2005) and Hamilton (1994), the model is stable if every eigenvalue modulus lies strictly inside the unit circle, meaning their values are lower than one. Under this condition, each observable variable can be expressed as a sum of past structural shocks whose effects decay over time. Stability, therefore, ensures the validity of long-run dynamics and forms the basis for impulse response analysis. As reported in **Table 6**, Panel A, both eigenvalues are below 1 and hence lie within the unit circle, confirming model stability, as plotted in Panel B.

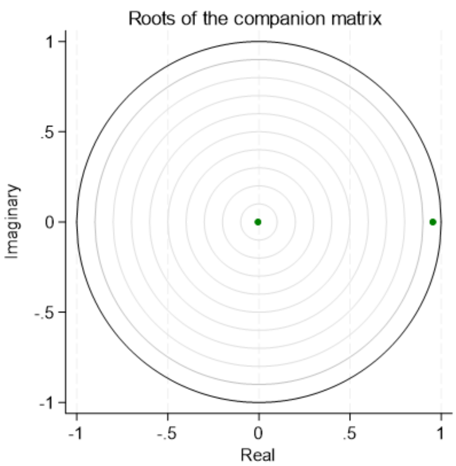
Table 6: Stability of first-order ESG-return PVAR

This table reports the stability diagnostics of the one-lag PVAR model specified in Equations (5) and (6), using ESG scores and annual stock returns for the sample 2003-2024. Panel A lists the real and imaginary parts of the eigenvalues of the companion matrix and their corresponding moduli. Panel B plots the roots of the companion matrix, where stability requires all points to lie strictly within the unit circle.

Panel A

Specification	Eigenvalue Real	Eigenvalue Imaginary	Modulus
ESG-Return	0.95	0	0.95
	-0.01	0	0.01

Panel B



5. Results and discussion

Chapter 5 presents and discusses the main findings of this thesis and is structured as follows: Chapter 5.1 provides the baseline results on the dynamic relationship of a firm's ESG score and its annual return. Chapter 5.2 assesses the relationship between the E, S, and G dimensions and their impact on stock returns individually. Chapter 5.3 further examines the dimensional relationships based on impulse response functions and Granger causality tests. Lastly, Chapter 5.4 tests key theories surrounding the presence of an ESG premium.

5.1 Results on the relationship between ESG scores and annual stock returns

This section presents the results of running the empirical regression as specified in Equations (5) and (6). The baseline PVAR results in **Table 7** are estimated in three stages, each incorporating an increasingly rich set of control variables. Model (1) is estimated based on the controls $LOGSIZE_{t-1}$ and $LOGB/M_{t-1}$, Model (2) adds $PROFITABILITY_{t-1}$ and $INVESTMENT_{t-1}$, and Model (3) adds $LEVERAGE_{t-1}$, $CASH_{t-1}$, $DIVIDEND_{t-1}$, and $LOGAGE_{t-1}$, following Shackleton et al. (2022). This approach ensures that the observed ESG-return dynamic is not an artifact of omitted variables and remains consistent across different control specifications.

The baseline results reveal a bidirectional relationship between a firm's ESG score and its 12-month return. Consistent with the *doing bad by doing good* perspective, the estimated coefficient for the ESG score is negative and significant at the 5% and 1% levels, indicating that a firm's past ESG score has a negative relationship with subsequent year returns (Columns (1), (3), and (5)). Consistent with the presence of reverse causality between ESG and return, a firm's past return exhibits a similar negative impact on subsequent ESG scores, significant at the 5% and 10% level (Columns (3) and (5)). Thus, a positive/negative shock in one is associated with a subsequent negative/positive shock in the other, holding firm and time constant.

On a first glance, the magnitude of the return coefficient estimates in Columns (3) and (5) indicates that past stock performance has a minor effect on subsequent ESG scores: a coefficient of -0.161 (Column (5)) corresponds to a change of about 0.1 points in ESG, given a one standard deviation change in returns. This mitigated unscaled shock aligns with the high persistence of ESG ratings, with autoregressive coefficients around 0.95, consistent with prior findings that CSP measures only evolve gradually over time (Lin et al., 2019; Qureshi et al., 2021; Shackleton et al., 2022; Yu et al., 2022). However, considering that a firm in my sample has an average change of 10 points in its ESG score, while the average period is approximately 13 years, a one-standard-deviation change in prior returns is associated with a 13% change in the theoretical annual variation of the ESG score.

Looking at the effect of past ESG on subsequent returns, a stable estimated coefficient of -0.002 can be observed across all models. Theoretically, a one standard deviation increase in the ESG score reduces subsequent annual returns by roughly 4 percentage points. The indicated size aligns well with the literature on negative ESG premia, which consistently are in the annual range of 2%-6% (Ciciretti et al., 2023; Hong and Kacperczyk, 2009; Luo, 2022). The observed negative ESG on return impact will subsequently be referred to as (ESG) premium.

Overall, the stability of the estimated coefficients in terms of magnitude and direction, as well as the absence of significant return autocorrelation across Models (1) to (3), confirms the PVAR

Table 7: Baseline PVAR results (ESG-Return)

This table reports the results from the one-lag PVAR model specified in Equations (5) and (6), using ESG_t and $Return_t$ as dependent variables over the full sample period 2003-2024. The regressions include lagged values of both *ESG* and *Return*, alongside lagged firm-level controls. Models (1) to (3) are identical specifications but expand the set of included controls. Columns (1), (3), and (5) present regressions with $Return_t$ as the dependent variable, while Columns (2), (4), and (6) present regressions with ESG_t as the dependent variable. Control variables are $LOGSIZE_{t-1}$, $LOGB/M_{t-1}$, $PROFITABILITY_{t-1}$, $INVESTMENT_{t-1}$, $LEVERAGE_{t-1}$, $CASHRATIO_{t-1}$, $DIVRATIO_{t-1}$, and $LOGAGE_{t-1}$, as defined in **Appendix 1**. All regressions are estimated via GMM and include firm and time fixed effects. Z-statistics are presented in parentheses below each coefficient. Statistical significance of the estimates is indicated by ***, **, and * for the 1%, 5%, and 10% levels, respectively.

	Model (1)		Model (2)		Model (3)	
	(1) Return	(2) ESG Score	(3) Return	(4) ESG Score	(5) Return	(6) ESG Score
$Return_{t-1}$	-0.009	-0.149	-0.005	**-0.192****	-0.005	**-0.161***
	(-1.19)	(-1.61)	(-0.64)	(-2.04)	(-0.62)	(-1.75)
$ESG\ Score_{t-1}$	**-0.002****	0.948***	**-0.002*****	0.952***	**-0.002*****	0.954***
	(-2.29)	(94.54)	(-2.71)	(94.71)	(-2.76)	(93.11)
$LOGSIZE_{t-1}$	0.020*	-0.731***	0.019*	-0.724***	0.050***	-0.515**
	(1.89)	(-4.77)	(1.81)	(-4.73)	(3.19)	(-2.54)
$LOGB/M_{t-1}$	0.036	-1.577***	0.072**	-2.000***	0.107***	-1.881***
	(1.17)	(-4.34)	(2.24)	(-5.37)	(3.39)	(-5.16)
$PROFITABILITY_{t-1}$			0.294***	-3.431***	0.414***	-3.808***
			(4.10)	(-3.31)	(5.06)	(-3.36)
$INVESTMENT_{t-1}$			0.004	-0.234	-0.014	-0.198
			(0.18)	(-1.12)	(-0.61)	(-0.94)
$LEVERAGE_{t-1}$					0.621***	-1.641
					(5.24)	(-1.12)
$CASHRATIO_{t-1}$					0.950***	-3.388*
					(5.18)	(-1.95)
$DIVRATIO_{t-1}$					0.724**	13.008**
					(2.38)	(2.47)
$LOGAGE_{t-1}$					-0.015	-0.831***
					(-0.67)	(-2.65)
Firm fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
N	37,347	37,347	37,347	37,347	37,347	37,347

specification choices. Moreover, control variables also display stable magnitudes and directions across Models (1) to (3), whereas the impacts reconcile with previous literature. PROFITABILITY, LEVERAGE, CASHRATIO, DIVRATIO, and LOGB/M significantly increase future returns, while INVESTMENT and AGE show no effect. Regarding ESG scores, INVESTMENT and LEVERAGE are insignificant, whereas PROFITABILITY is negatively related at the 1% level, supporting the *doing bad by doing good* narrative. DIVRATIO exhibits a positive association, while other controls are negatively linked to ESG.

Overall, the baseline PVAR results uncover a dynamic, negative bidirectional relationship between a firm's ESG score and its return. Thus, they conclusively confirm Hypothesis 1.

The key to understanding and interpreting this bidirectional relationship is to separate the adverse shocks of prior ESG and prior returns. Precisely, it is unclear whether a firm's previous poor performance leads to an ESG increase or if prior good performance leads to an ESG decrease, or both. A similar question to interpret my findings is whether firms with high versus low ESG scores observe different impacts on their subsequent return. To answer this question, I disentangle both the ESG and return variables by introducing high and low dummies. *High* dummies are assigned a value of 1 if the underlying firm is in the top 30^{th} percentile, and 0 otherwise. *Low* dummies are assigned a value of 1 if the firm is in the bottom 30^{th} percentile, and 0 otherwise. **Table 8** Panel A reports the results for the high and low return dummies, and Panel B presents the corresponding results for the ESG dummies. All subsequent models are estimated with the complete set of control variables listed in **Table 7**, Model (3).

Table 8 Panel A reveals an asymmetric dynamic in how past returns influence future ESG scores. The estimated coefficient for the *High Return* dummy in Column (2) is insignificant, while for the *Low Return* dummy in Column (4), a negative and significant coefficient (-0.263 at the 5% level) is observed. This indicates that the impact of prior return on subsequent ESG is driven by underperformance: firms do not reduce ESG efforts after strong performance, but they increase their ESG engagement following periods of weak performance. This finding supports the evidence presented by Shackleton et al. (2022), confirming that improvements in corporate social performance are often preceded by poor stock returns.

Panel B reveals a similarly asymmetric dynamic in how past ESG scores influence future returns. The estimated coefficient for the *High ESG Score* dummy in Column (1) is negative and significant at the 1% level, indicating that firms with high prior ESG scores experience an adverse effect on subsequent returns. Conversely, the *Low ESG Score* dummy in Column (3) is

Table 8: Results of high vs. low return and ESG dummies

This table reports PVAR estimates from Equations (5) and (6) for the period 2003-2024, where *ESG* and *Return* are replaced with high and low dummy variables. High (low) dummies have a value of 1 (0) if the underlying datapoint exceeds the annual 70th (30th) percentile, and 0 (1) otherwise. Panel A presents results for *High-* and *Low Return* dummies, while Panel B reports estimates for *High-* and *Low ESG* dummies. All specifications include the full set of lagged control variables presented in **Table 7**, as well as firm and time fixed effects. Estimates are obtained via GMM. Z-statistics are presented in parentheses. Statistical significance of the estimates is indicated by ***, **, and * for the 1%, 5%, and 10% levels, respectively.

Panel A

	Model (1)		Model (2)	
	(1)	(2)	(3)	(4)
	High Return	ESG Score	Low Return	ESG Score
High $Return_{t-1}$	0.005	-0.103		
	(0.82)	(-0.99)		
ESG $Score_{t-1}$	**-0.003*** **	0.953***		
	(-3.57)	(93.37)		
Low $Return_{t-1}$			0.011	**-0.263** **
			(1.55)	(-2.47)
ESG $Score_{t-1}$			0.001	0.952***
			(1.33)	(93.59)
Controls	Yes	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes
N	37,347	37,347	37,347	37,347

Panel B

	Model (3)		Model (4)	
	(1)	(2)	(3)	(4)
	Return	High ESG Score	Return	Low ESG Score
$Return_{t-1}$	-0.004	-0.005		
	(-0.53)	(-1.48)		
High ESG $Score_{t-1}$	**-0.063*** **	0.587***		
	(-3.24)	(52.99)		
$Return_{t-1}$			-0.005	0.005
			(-0.62)	(1.47)
Low ESG $Score_{t-1}$			**0.042*** **	0.612***
			(2.63)	(64.25)
Controls	Yes	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes
N	37,347	37,347	37,347	37,347

positive and significant at the 1% level, showing that firms with low prior ratings observe a positive impact on subsequent returns. This non-linear dynamic, where high-ESG firms are penalized and low-ESG firms are rewarded, is highly consistent with several foundational theories on how ESG is priced in capital markets. The asymmetric dynamic provides empirical evidence for the investor demand theory, whereby investors cluster in high-ESG assets and accept lower returns in exchange for non-financial value. The ownership concentration in high ESG assets subsequently drives the lower returns of such assets. The consequent reverse effect is that low ESG assets face reduced demand and thus observe positive return impacts, represented by the significant positive low ESG score coefficient of Column (4) (Darolles et al., 2023; Pastor et al., 2022; Pedersen et al., 2021).

The observed pattern further aligns with the existence of a non-linear relationship between CSP and CFP, documented in prior studies (Lewandowski, 2017; Trumpp and Guenther, 2017). My study provides additional evidence for a global inverted U-shaped nexus between returns and ESG. Investors seem to price an optimal level of ESG, as returns are dynamically adjusted in the opposite direction, driven by high or low ESG levels. A deviation from the mean ESG score might change the investor's marginal benefit or alter the underlying business impact of ESG. The positive impact of the low ESG dummy on subsequent returns is further evidence for the existence of a sin-premium, which could be driven by limited risk sharing and the existence of alleviated risk factors of low ESG assets (Hong and Kacperczyk, 2009).

The results demonstrate that the ESG-return nexus is characterized by complex, non-linear, and bidirectional dynamics. This inherent complexity helps explain the persistent lack of consensus in the broader CSP-CFP literature, with the asymmetric and U-shaped nexus providing a basis for differing positive and negative findings.

5.2 Results on the relationship between E, S, and G scores and annual stock returns

ESG scores are a composite measure that aggregates heterogeneous indicators into a single index. To gain more granular insights into the drivers of the ESG premium and the positive ESG response to prior underperformance, the aggregate ESG variable is deconstructed into its three pillars (E, S, and G), and the baseline PVAR framework is re-estimated for each pillar. Results are presented in **Table 9**.

Interestingly, the pillar-level analysis reveals that past stock returns do not significantly influence the environmental or social dimensions, as the estimated coefficients of prior *Return* are insignificant (Columns (2) and (4)). By contrast, governance responds strongly to prior

Table 9: Results of E, S, G, and Return

This table reports PVAR estimates from Equations (5) and (6), where the aggregate ESG score is decomposed into its environmental (E), social (S), and governance (G) pillars. The dependent variables are pillar scores (E_t, S_t, G_t) and 12-month stock returns ($Return_t$) over the period 2003-2024. Model (1) presents the system for *E* and *Return*, Model (2) for *S* and *Return*, and Model (3) for *G* and *Return*. All specifications include the full set of lagged control variables presented in **Table 7**, as well as firm and time fixed effects. Estimates are obtained via GMM. Z-statistics are presented in parentheses. Statistical significance of the estimates is indicated by ***, **, and * for the 1%, 5%, and 10% levels, respectively.

	Model (1)		Model (2)		Model (3)	
	(1)	(2)	(3)	(4)	(5)	(6)
	Return	E Score	Return	S Score	Return	G Score
$Return_{t-1}$	-0.004	-0.154				
	(-0.52)	(-1.25)				
$E\ Score_{t-1}$	**-0.002*****	0.970***				
	(-4.10)	(-109.37)				
$Return_{t-1}$			-0.005	0.091		
			(-0.62)	(-0.79)		
$S\ Score_{t-1}$			**-0.001****	0.960***		
			(-2.04)	(-97.59)		
$Return_{t-1}$					-0.005	**-0.432*****
					(-0.56)	(-2.69)
$G\ Score_{t-1}$					-0.001	0.646***
					(-1.47)	(-73.42)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
N	37,347	37,347	37,347	37,347	37,347	37,347

performance. The estimated coefficient of lagged *Return* on the *G Score* is -0.432 and significant at the 1% level (Column (6)). The magnitude implies that a one-standard-deviation change in return shifts the governance score by approximately 0.24 points. Given an average within-firm variation of 12.6 governance points over 13 years, this corresponds to roughly 26% of the annual variation in the G score at the firm level. Combining the findings with the high and low dynamics before, the insights suggest that prior adverse return shocks trigger positive governance-related adjustments. The relationship between *G* and prior stock returns supports evidence on increased shareholder activism, shareholder campaigns, and shareholder pressure after poor stock performance (Easterwood et al., 2012; Hermalin and Weisbach, 2005; Karpoff, 2001).

Turning to the *E*, *S*, and *G* on *Return* impact, *E* and *S* scores exhibit significant negative coefficients (-0.002 and -0.001 at the 1% and 5% levels, respectively), with the effect most pronounced for the E pillar. In contrast, governance does not show a significant premium. This refines the results of Chapter 5.1, showing that the composite score premium is mainly driven by the E and S dimensions. This finding again hints at the investor preference view, suggesting that investors with ethical or climate-related preferences accept lower returns in exchange for high E and S performance. The premium for the environmental score is consistent with a carbon-risk interpretation and could reflect exposure to corporate carbon emissions (Bolton and Kacperczyk, 2021), while the S score premium points to a broader ESG frontier nexus, as formalized by Pedersen et al. (2021), who likewise document a premium for E and S but not for G.

Overall, the results highlight the heterogeneity, noise, and converging effects embedded in composite ESG data. Investors clearly differentiate between dimensions, while a composite index nets effects across multiple pillars and thus obscures dimension-specific dynamics. This further explains the mixed evidence reported in the ESG-CFP literature.

5.3 Impulse response functions and Granger causality tests on the E, S, G, and annual stock return link

The results in Sections 5.1 and 5.2 point to a bidirectional asymmetric ESG-return nexus that differs by dimension and is partially non-linear. To confirm whether a persistent return discount follows environmental and social upgrades and whether governance adjusts endogenously after performance shocks, impulse response functions (IRFs) and Granger causality tests are subsequently estimated.

IRFs provide further insights by showing how an innovation (shock) in one variable affects the other variables in the system over a specified horizon, holding other shocks constant. Based on the coefficients of Models (1) to (3) as reported in **Table 9** of Chapter 5.2, I estimate the IRFs following the methodology of Chang and Zhang (2015). The shocks are orthogonalized using the inverse Cholesky decomposition of the residual covariance matrices of each model. This imposes a causal ordering on the endogenous variables in the system and isolates each shock and its respective response. To assess the statistical significance of the IRFs, confidence intervals and standard errors are estimated via Monte Carlo simulations with 500 repetitions. In total, six IRFs are estimated to trace the system's dynamics over a five-year horizon. These functions map the response of *Return* to a one-standard-deviation shock in each *E*, *S*, and *G*, as well as the reverse response of each pillar to a shock in *Return*. **Table 10** provides an overview of the results.

The IRFs clearly confirm the main pattern from Chapters 5.1 and 5.2. A positive shock to the environmental pillar results in a persistent and statistically significant decline in returns. The social pillar also produces a negative but less pronounced return response. In both cases, the cumulative five-year return path remains below 0, consistent with a negative premium, significant at the 5% level. By contrast, governance shocks do not generate a response in returns, as the 95% confidence bands cross the zero line.

In the reverse direction, the *G Score* is significantly impacted by subsequent returns. Theoretically, a negative return shock leads to a positive subsequent *G* shock. The *E* and *S* responses are not significantly different from 0 during the 5 years, as the confidence interval of the shocks crosses the 0 line.

The IRF dynamics are confirmed by the results of the Granger causality tests presented in **Table 11**. The null hypothesis states that the variable does not Granger-cause the other variable in the system. As can be seen, both *E* and *S* Granger cause *Return* (1% and 5% significance), while *Return* does not Granger cause *E* or *S*. For *G*, *Return* does Granger cause *G* (1% significance), while *G* does not Granger cause *Return*.

Accordingly, Hypothesis 2 can be confirmed: the bidirectional relationship varies in both direction and magnitude across the individual environmental, social, and governance dimensions. In my sample, the negative premium associated with high ESG is primarily driven by the environmental and social dimensions. At the same time, the governance dimension adjusts in response to past performance while showing no negative return impact. Consequently, the governance dimension drives the reverse causality between overall ESG and returns in my sample.

Table 10: Impulse response functions of E, S, G, and Return

This table presents orthogonalized impulse response functions (IRFs) based on the estimated PVAR models in **Table 9** for the sample period 2003-2024. The graphs display six cases: Model (1) shows the responses of *Return* to shocks in the *E Score* and of the *E Score* to shocks in *Return*, Model (2) reports the corresponding responses for the *S Score* and *Return*, and Model (3) presents the responses for the *G Score* and *Return*. Shocks are orthogonalized using the inverse Cholesky decomposition of the residual covariance matrix. Confidence intervals at the 95% level are estimated based on 500 Monte Carlo simulations. The dynamics are shown over a five-year horizon, where shocks occur at time 0 and responses unfold in subsequent periods.

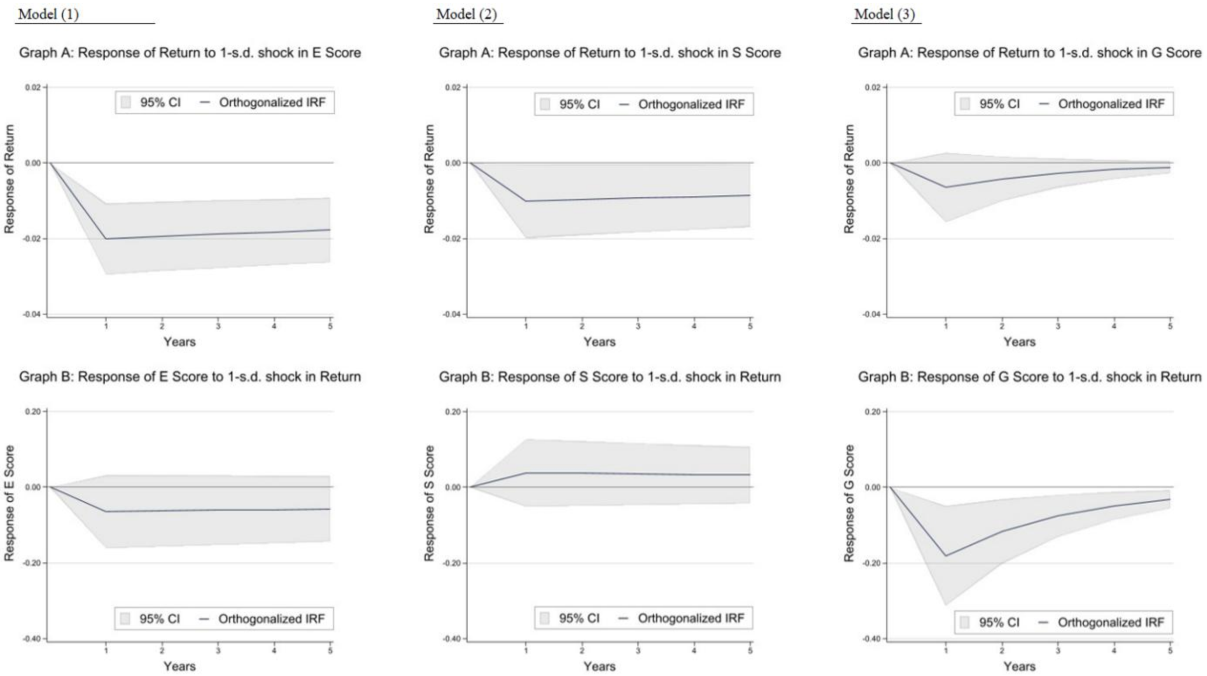


Table 11: PVAR Granger causality tests for E, S, G, and Return

This table reports panel Granger causality tests based on the one-lag PVAR models specified in Equations (5) and (6) over the sample period 2003-2024. The null hypothesis states that past values of one variable do not Granger-cause the other variable in the system. Reported are the number of observations, the Wald test chi-squared statistic, and the associated p-value.

Model	Null hypothesis	Obs.	X^2	p value
E Score - Return	E Score does not Granger cause Return	37,347	16.811	0.000
	Return does not Granger cause E Score	37,347	1.554	0.213
S Score - Return	S Score does not Granger cause Return	37,347	4.144	0.042
	Return does not Granger cause S Score	37,347	0.631	0.427
G Score - Return	G Score does not Granger cause Return	37,347	2.166	0.141
	Return does not Granger cause G Score	37,347	7.223	0.007

5.4 Mediating channels of the negative ESG premium

The preceding analysis successfully isolated two distinct dynamics within the ESG-return nexus: a reverse-causality channel driven by governance adjusting to past underperformance, and a negative premium driven by the environmental and social pillars. While the adjustment of the G Score is consistent with established theories on shareholder pressure and lies outside of traditional CSP, the more disputed finding concerns the persistent negative E&S premium. The remainder of this chapter, therefore, focuses on exploring its origins by testing two prominent theories: investor preferences and market frictions.

First, I treat ownership composition as revealed preference, building on evidence on the ownership ESG relationship. Long-horizon institutions tilt toward higher-ESG firms and rebalance around credible ESG inclusions, implying that owners actively embed ESG preferences. Cross-country evidence shows that institutions actively push portfolio firms toward stronger E&S, again linking active owner identity to ESG preferences (Dyck et al., 2019). Further, passive index funds are less likely to vote against management and are weaker monitors than active funds, so a lower passive share should correlate with more expressed preferences through trading and ESG engagement (Heath et al., 2022). While large passive owners can still influence policies via engagement, the split into active and passive ownership captures marginal expression intensity rather than a total absence of preferences among passive holders (Azar et al., 2021). If the premium is preference-driven, it should intensify or cluster where passive ownership is relatively low.

Second, I test whether liquidity serves as a mediator through which the premium materializes. A growing body of evidence shows that better ESG performance and disclosure positively impact stock liquidity. (He et al., 2023; Krueger et al., 2024; Meng-tao et al., 2023; Zhang et al., 2024). Following asset-pricing intuition, the reverse of the illiquidity premium indicates that greater liquidity can imply a lower required return, so if ESG improves liquidity, the associated negative valuation premium might be most apparent in the high-liquidity segment (Amihud, 2002). Investor-recognition models reinforce this mechanism. As ESG disclosure and visibility broaden the investor base, prices would rise and expected returns fall, an effect that is more powerful where trading is deep and friction is low (Lehavy and Sloan, 2005; Merton, 1987). Parallel evidence from sustainable debt markets points the same way. Corporate green bonds exhibit a liquidity Greenium alongside price premia, indicating that greener, more liquid issues might imply higher valuations (Molino et al., 2023). Interestingly, evidence is not uniform. Luo (2022) finds that the premium is present for low-liquidity UK stocks.

Building on these ownership- and liquidity-based findings, the PVAR model is estimated for split sample groups. Specifically, I split the sample by passive ownership (*PASSIVEHOLD)* and by the turnover ratio (*LIQUIDITY*), defined in Chapter 3, using the below the 30^{th} and above the 70^{th} percentile high and low cutoffs.

The results presented in **Table 12** support both the investor preference and liquidity channels. For the liquidity split, the negative premium on both the *E* and *S* scores is statistically significant and economically meaningful for firms in the high-liquidity segment. The estimated coefficient of the *E Score* on subsequent *Return* is negative and significant at the 1% level for high-liquidity firms (-0.002, Column (1)) but insignificant for low-liquidity firms (Column (3)). A similar pattern holds for the *S Score*, as the negative premium, significant at the 1% level, is only present in the high-liquidity group and vanishes in the low-liquidity subsample. This finding might suggest that the ESG valuation effect materializes through a broader investor base and reduced trading frictions, which are characteristic of liquid stocks. The premium is absent where trading is thin, confirming that, in a global sample, liquidity acts as a key mediator for the pricing of ESG attributes.

The ownership split provides evidence for the investor preference channel. The analysis reveals that the negative E&S premium is only present in firms with low passive ownership. For the E score, the coefficient is -0.007 and significant at the 1% level in the low-passive group, while it is not significant for firms with high passive ownership. The S score exhibits the same dynamic, with a significant (10% level) coefficient of -0.004 for the low-passive segment, which disappears in the high-passive group. This result provides strong support for the hypothesis that the premium is driven by the preferences of active, engaged investors. As argued by Heath et al. (2022), a lower passive share implies a greater influence of active managers who express their ESG preference through investment and engagement.

In conclusion, the CSP premium is not uniform; rather, it is conditional on firm characteristics like passive ownership and liquidity. The premium is present where passive ownership is low and where stocks exhibit high liquidity. These findings support Hypothesis 3, confirming that the dynamic ESG-return relationship is significantly mediated by both firm liquidity and ownership structure. Overall, the results can provide evidence for the investor preference theory.

Table 12: Sample splits for the E and S dimension

This table reports PVAR estimates from Equations (5) and (6) over the sample period 2003-2024, where the analysis focuses on the environmental (E) and social (S) dimensions. Panel A presents results for the *E Score*, and Panel B reports results for the *S Score*. Within each panel, the sample is split by liquidity and by passive ownership, based on the 30th and 70th percentiles of the respective distributions. All specifications include the full set of lagged control variables presented in **Table 7**, as well as firm and time fixed effects. Estimates are obtained via GMM. Z-statistics are presented in parentheses. Statistical significance of the estimates is indicated by ***, **, and * for the 1%, 5%, and 10% levels, respectively.

Panel A

	Model (1) - Liquidity				Model (2) - Passive Ownership			
	High		Low		High		Low	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Return	E Score	Return	E Score	Return	E Score	Return	E Score
$Return_{t-1}$	0.033*	-0.316	-0.048**	-0.306	-0.040	-0.042	-0.007	-1.059*
	(1.85)	(-1.18)	(-2.54)	(-0.97)	(-1.57)	(-0.09)	(-0.20)	(-1.90)
$E\ Score_{t-1}$	**-0.002*****	0.966***	**0.000**	0.996***	**-0.000**	0.929***	**-0.007*****	0.929***
	(-3.08)	(57.56)	**(0.24)**	(36.13)	**(-0.23)**	(37.16)	**(-2.76)**	(37.16)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	10,061	10,061	9,058	9,058	5,662	5,662	4,393	4,393

Panel B

	Model (3) - Liquidity				Model (4) - Passive Ownership			
	High		Low		High		Low	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Return	S Score	Return	S Score	Return	S Score	Return	S Score
$Return_{t-1}$	0.031*	0.436*	-0.047**	0.138	-0.006	0.336	-0.040	0.048
	(1.73)	(1.77)	(-2.51)	(0.48)	(-0.17)	(0.87)	(-1.56)	(0.1)
$S\ Score_{t-1}$	**-0.003*****	0.978***	**0.003**	0.977***	**-0.001**	0.964***	**-0.004***	0.918***
	(-2.98)	(48.64)	**(1.55)**	(35.09)	**(-0.42)**	(33.57)	**(-1.68)**	(23.99)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	10,061	10,061	9,058	9,058	5,662	5,662	4,393	4,393

6. Further analysis and robustness

This chapter assesses the robustness of the results presented in Chapter 5. Chapter 6.1 explores regional heterogeneity, Chapter 6.2 assesses robustness to alternative variable definitions, and Chapter 6.3 presents results for varying sample restrictions.

6.1 Regional dynamics

The ESG-return relationship is expected to exhibit regional heterogeneity because institutional environments, ownership structures, and the intensity of shareholder activism differ across countries. This can impact firms' sustainability policies and investors' reactions (Dyck et al., 2019; Griffin et al., 2021). To investigate this, the baseline analysis between ESG and returns is re-estimated individually for each major geographic region.

Panels A and B of **Table 13** show that the negative premium remains stable and significant across all developed regions. However, the reverse-causality channel is statistically significant only in North America. Furthermore, in emerging markets, no significant relationship between ESG and returns is observed.

The concentration of the reverse-causality channel in North America might reflect the stronger culture of shareholder activism in the United States (Becht et al., 2017; Hartmann et al., 2021). Crucially, the U.S. seems to be a dominant driver of the return on ESG impact associated with the G dimension in the overall sample. The absence of a significant ESG-return relationship in emerging markets could be linked to weaker institutional frameworks, noisier ESG data, and shorter time series of ESG scores in the emerging countries.

6.2 Robustness to variable specifications

To mitigate concerns that the results are driven by idiosyncratic shocks, systematic risk, and noise contained in raw returns, the main ESG-return analysis is repeated using two alternative return metrics (Black, 1986). Abnormal returns, approximated by the CAPM model, are used to ensure that market co-movement, market shocks, and firm-level risk characteristics are not driving the explored return-ESG relationship. This addresses previous findings that CSR can be used to decrease systematic risk and offer downside protection (Albuquerque et al., 2020; Shackleton et al., 2022).

Next, industry-adjusted returns are used to account for sector-specific performance patterns, as firms within the same industry often share common characteristics and are exposed to the same economic shocks (Bolton and Kacperczyk, 2021; Fama and French, 1998).

Finally, to validate the results against an alternative composite ESG metric, the analysis is replicated using the ESGC score from LSEG.

The main findings are robust across all three alternative specifications, presented in **Table 14**. The statistical and economic significance of the key coefficients is maintained, and in some cases, strengthened.

Table 13: Relationship of ESG-Return across different regions

This table reports PVAR estimates from Equations (5) and (6) for the sample period 2003-2024, split by region. Panel A presents results for Japan, North America, and Europe, while Panel B reports results for Asia-Pacific (excluding Japan) and Emerging Markets/Rest of World. Firms are assigned to regions following the Fama-French (2012) mapping, supplemented by Kenneth-French data library mapping for emerging markets/RoW. See **Appendix 3** for the detailed mapping. All specifications include the full set of lagged control variables presented in **Table 7**, as well as firm and time fixed effects. Estimates are obtained via GMM. Z-statistics are presented in parentheses. Statistical significance of the estimates is indicated by ***, **, and * for the 1%, 5%, and 10% levels, respectively.

Panel A

	Model (1) - Japan		Model (2) - N. America		Model (3) - Europe	
	(1)	(2)	(3)	(4)	(5)	(6)
	Return	ESG Score	Return	ESG Score	Return	ESG Score
$Return_{t-1}$	-0.032	**-0.999**	-0.003	**-0.219***	0.029	**-0.053**
	(-0.92)	**(-1.37)**	(-0.22)	(-1.69)	(1.65)	**(-0.26)**
$ESG\ Score_{t-1}$	**-0.003****	0.959***	**-0.005****	0.954***	**-0.005*****	0.865***
	(-1.98)	(32.51)	(-2.11)	(40.36)	(-4.55)	(49.95)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
N	4,913	4,913	12,374	12,374	8,814	8,814

Panel B	Model (4) - APAC		Model (5) - Emerging	
	(7)	(8)	(9)	(10)
	Return	ESG Score	Return	ESG Score
$Return_{t-1}$	0.059	**-0.331**	-0.018	**-0.035**
	(1.56)	**(-0.60)**	(-0.92)	**(-0.17)**
$ESG\ Score_{t-1}$	**-0.005****	0.931***	**-0.000**	0.994***
	(-2.27)	(25.05)	**(-0.26)**	(56.02)
Controls	Yes	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes
N	2,114	2,114	9,132	9,132

Table 14: Alternative return and ESG specifications

This table reports PVAR estimates from Equations (5) and (6) for the sample period 2003-2024, using alternative measures of returns and ESG. Model (1) replaces returns with industry-adjusted returns, Model (2) uses abnormal returns, and Model (3) substitutes the ESG score with the ESGC score. All specifications include the full set of lagged control variables presented in **Table 7**, as well as firm and time fixed effects. Estimates are obtained via GMM. Z-statistics are presented in parentheses. Statistical significance of the estimates is indicated by ***, **, and * for the 1%, 5%, and 10% levels, respectively.

	Model (1)		Model (2)		Model (3)	
	(1)	(2)	(3)	(4)	(5)	(6)
	Ind. Adjusted Return	ESG Score	Abnormal Return	ESG Score	Return	ESGC Score
Ind. Adj. $Return_{t-1}$	-0.01	**-0.196****				
	(-1.19)	(-2.11)				
ESG $Score_{t-1}$	**-0.002****	0.954***				
	(-2.31)	(93.08)				
Abnormal $Return_{t-1}$			-0.006	**-0.262*****		
			(-0.64)	(-2.66)		
ESG $Score_{t-1}$			**-0.002****	0.955***		
			(-2.29)	(92.46)		
$Return_{t-1}$					-0.005	**-0.384*****
					(-0.66)	(-2.95)
ESGC $Score_{t-1}$					**-0.002*****	0.871***
					(-2.78)	(73.88)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
N	37,347	37,347	37,347	37,347	37,347	37,347

6.3 Robustness to sample specifications

A final set of robustness checks addresses potential bias from the sample composition by re-estimating the baseline model on two subsamples. First, financial firms are excluded from the sample. This is a common practice in the literature, as their regulatory environment and balance-sheet structure differ substantially from other sectors (Hoepner et al., 2024; Lins et al., 2016). Second, the COVID-19 crisis period is excluded, using the U.S. recession indicator to identify the main crash period.[3] This test is motivated by findings that the ESG-return relationship was

[3] The recession indicator published by the Federal Reserve Bank of ST. Louis can be found here: https://fred.stlouisfed.org/series/JHDUSRGDPBR

anomalous during this period, with high-ESG firms exhibiting abnormal resilience (Albuquerque et al., 2020).

The results in **Table 15** remain consistent across both sample restrictions. The negative ESG premium and the reverse impact persist after excluding financial firms. Removing the COVID-19 period makes the return impact on ESG insignificant, however, the relationship remains negative and significant for the G dimension even when excluding the COVID-19 impact (Model (3)).

Table 15: Robustness to sample specifications

This table reports PVAR estimates from Equations (5) and (6) for the sample period 2003-2024 under alternative sample restrictions. Model (1) excludes financial firms (ICB code 30), Model (2) excludes observations from the Covid-19 period, and Model (3) applies both restrictions and replaces the ESG score with the governance (G) score. All specifications include the control variables defined in Table 2, as well as firm and time fixed effects. Estimates are obtained via system GMM; z-statistics are reported in parentheses. Statistical significance is indicated by ***, **, and * for the 1%, 5%, and 10% levels, respectively.

	Model (1) - No Financials		Model (2) - No Covid		Model (3) - G & No Covid	
	(1)	(2)	(3)	(4)	(5)	(6)
	Return	ESG Score	Return	ESG Score	Return	G Score
Return	-0.005	**-0.165***	-0.001	-0.142		
	(-0.59)	**(-1.77)**	(-0.10)	(-1.10)		
ESG Score$_{t-1}$	**-0.002*****	0.955***	**-0.001***	0.947***		
	(-2.83)	-93.02	**(-1.85)**	-88.85		
Return$_{t-1}$					-0.001	**-0.635*****
					(-0.09)	**(-3.03)**
G Score$_{t-1}$					-0.001	0.649***
					(-1.52)	-69.89
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
N	36,922	36,922	29,618	29,618	29,618	29,618

7. Limitations and future research

This chapter addresses the study's limitations and outlines potential directions for future research. Chapter 7.1 highlights limitations of the analysis, while Chapter 7.2 proposes avenues for further research.

7.1 Limitations

The first limitation is the quality and specification of the ESG data. Historical Refinitiv/LSEG scores have been retroactively revised, and methodology changes are implemented periodically, which challenges time-series consistency in dynamic models (Berg et al., 2021). Additionally, rating disagreements among providers can be substantial. Varying methodologies and weighting schemes can result in findings that are provider-specific. The divergence itself has previously been found to have an impact on the relation between the ESG score and other variables, further supporting that it is challenging to generalize results across rating providers (Berg et al., 2022; Horn and Oehler, 2024). Coverage and score-construction also matter. LSEG scores can rely on self-reported data and percentile benchmarking, which can correlate with firm fundamentals and disclosure frequency, leading to selection and disclosure bias (Drempetic et al., 2020; London Stock Exchange Group, 2024). Furthermore, the findings show the inherent noise and slow-moving nature of ESG scores as well as the problem of indicator aggregation. Changes in composite scores may be triggered by a narrow subfactor rather than a broad dimensional change (Benuzzi et al., 2025).

Secondly, the GMM diagnostics, particularly the marginal significance of the Hansen J test, necessitate a careful interpretation of the findings. This result highlights well-known challenges of GMM in short dynamic panels. Especially, the trade-off between using a parsimonious instrument set to ensure test validity and using deeper lags to identify long-term effects is relevant for my sample. This is compounded by the high persistence of the ESG variable, which can lead to weak instrument issues and make the coefficient estimates sensitive to model specification. While the findings are robust across several checks, the results still rely on key model assumptions.

Third, results of split-sample analysis generally need to be interpreted with caution. The splitting variables themselves can be endogenous and correlate with other firm fundamentals, which means the subsamples may differ systematically in their composition. Further, proxies for passive ownership and turnover ratio might not accurately signal ownership preference or liquidity, respectively.

Finally, the observed dynamics may be influenced by omitted risk factors that are correlated with ESG performance. For instance, the negative premium on the E and S pillars could be driven by unobserved reduced risk factors of ESG firms. While the inclusion of control variables, the usage of fixed effects, and the robust results for abnormal returns do mitigate this concern, the possibility that unobserved risk factors are driving the results cannot be fully ruled out.

7.2 Future research

These limitations open several avenues for further research. First, replicating the results using multiple ESG providers would enable research to explicitly model rating disagreement as a distinct variable and assess its impact on the CSP-CFP nexus in a dynamic setting, as well as evaluate whether the findings are generalizable across rating providers.

Secondly, technological advances now make it feasible to develop more informative, higher-frequency ESG signals. Using modern AI, researchers can analyze public news and disclosures to construct daily or weekly indices. Such measures could further help to overcome the noise and persistence of traditional ESG measures.

Lastly, future research should expand the analysis further to individual industries and countries. For example, by including country and industry-specific factors, the impact of global and industry differences on the dynamic ESG-return nexus can be better assessed. Such factors might include country development or equality indicators, as well as industry competitiveness and cyclicity.

8. Conclusion

This thesis aims to examine the relationship between firms' ESG scores and stock returns, addressing the limitations of static models and US-centric samples in prior research. Consequently, a dynamic PVAR framework was applied to a global panel dataset covering more than 6,000 firms across 52 countries between 2003 and 2024.

Three central findings emerge. First, ESG scores and returns are linked through an asymmetric negative two-way relationship: higher ESG scores are associated with lower subsequent returns, driven by high ESG firms. Low ESG firms are associated with a positive (sin) premium, while weak past returns precede ESG improvements. Second, these dynamics differ across dimensions. The environmental and social pillars drive the negative return premium, consistent with investor-preference and carbon-risk channels. Governance, however, carries no premium, primarily adjusts after poor performance, and consequently drives the return on ESG impact. Third, the ESG premium is conditional on stock characteristics. It materializes in highly liquid stocks and in those with low passive ownership.

These findings explain the mixed evidence in the literature. They highlight that composite ESG scores mask dimension-specific dynamics and that static frameworks cannot capture the endogenous feedback loops between ESG and financial outcomes. For practitioners, the results imply that ESG premia are context-dependent, shaped by investor preferences and market structures. For policymakers, the study underlines the importance of transparent, detailed, and consistent ESG measurement to reduce noise and improve comparability. Finally, for researchers, the results strongly emphasize the need for dynamic modeling when exploring the ESG-return relationship.

In conclusion, this thesis demonstrates that the ESG-return relation cannot be reduced to a simple, unidirectional effect. Instead, it represents a dynamic and shifting nexus shaped by factors such as ownership, firm responses, and market factors. Overall, I suggest that neither Freeman's nor Friedman's theoretical frameworks fully capture the realities of modern capital markets and ESG impacts. The results of this thesis, therefore, move beyond the simplistic dichotomy of Friedman and Freeman. They demonstrate that the financial implications of ESG are not a binary outcome but are highly conditional, which reconciles the conflicting positive, negative, and neutral findings that characterize the existing literature.

Appendices

Appendix 1: Variable definitions, sources, and references

This table defines all variables used in the analysis, including their source, identifiers, description, and academic references. Panel A covers the key PVAR model variables: returns, ESG scores, and firm-level controls, while Panel B summarizes additional liquidity and ownership variables.

Variable name	Source	Description	References
Baseline Variables			
Return (R)	LSEG	Annual return derived from the stock price (P)	Shackleton et al. (2022)
Abnormal return (AR)	LSEG, Kenneth French Data Library	CAPM-adjusted annual return	Sharpe (1964), Jensen et al. (2006)
Industry-adjusted return	LSEG	Annual return (P) minus ICB industry mean return	Shackleton et al. (2022)
ESG	LSEG	Composite ESG percentile score (0-100)	Luo (2022)
ESGC	LSEG	ESG score adjusted for controversies (0-100)	n.a.
E	LSEG	Environmental pillar score (0-100)	Luo (2022)
S	LSEG	Social pillar score (0-100)	Luo (2022)
G	LSEG	Governance pillar score (0-100)	Luo (2022)
LOGSIZE	LSEG	Natural logarithm of market capitalization (MVC)	Bolton & Kacperczyk (2021)
LOGB/M	Compustat	Natural logarithm of book equity incl. deferred taxes (TXDITC) minus preferred stock (PSTKRV/PSTKL/PSTK) over market cap (MVC)	Fama & French (1993); Novy-Marx (2013)
LEVERAGE	Compustat	Long-term debt (DLTT) + short-term debt (DLC) over total assets (AT)	Shackleton et al. (2022)

PROFITABILITY	Compustat	Sales (REVT) - cost of goods sold (COGS), scaled by total assets (AT)	Shackleton et al. (2022)
INVESTMENT	Compustat	Change in total assets (ATt - ATt-1) divided by lagged assets (ATt-1)	Shackleton et al. (2022)
CASHRATIO	Compustat	Cash and short-term investments (CHE) over total assets (AT)	Shackleton et al. (2022)
DIVRATIO	Compustat	Cash dividends (DVC + DVP) scaled by total assets (AT)	Shackleton et al. (2022)
LOGAGE	LSEG	Natural logarithm of years since incorporation (WC18273)	Shackleton et al. (2022)
Additional Variables			
LIQUIDITY	LSEG	Annual trading volume (VO) divided by outstanding shares (NOSH)	Gabrielsen et al. (2011)
PASSIVEHOLD	LSEG	Percentage of shares held by passive owners (NOSHSP)	Heath et al. (2022)

Appendix 2: Data construction and cleaning steps

This table provides an overview of the raw LSEG and Compustat datasets, the merged panel, and the three key filters applied to construct the final sample. The sample covers the period from January 2003 to December 2024. The cleaning steps consist of removing firms with missing values for baseline variables defined in **Appendix 1**, restricting the dataset to firms with at least three consecutive fiscal-year observations, and excluding countries represented by fewer than ten firms.

Data construction step	Firms	Observations
LSEG raw dataset	11,903	210,312
Compustat raw dataset	11,222	212,230
Raw merged dataset	10,557	177,029
Data after removal of missing values for key variables	7,558	52,724
Data restricted to firms with ≥ 3 consecutive observations	6,437	50,788
Data restricted to countries with ≥ 10 firms	6,379	50,460

Appendix 3: Overview of countries, regions, and firm counts

This table reports the distribution of firms in the sample by country and region, as well as the total number of firms in each country. It lists ISO country codes, country names, and their classification into the five geographic regions employed in this study: North America, Europe, Japan, Asia-Pacific (excluding Japan), and Emerging Markets/Rest of World. The regional mapping follows Fama and French (2012), except for Emerging countries, which are classified based on the Kenneth R. French data library. Additionally, manual assignment based on headquarters location is conducted for countries not explicitly covered by either classification.

ISO country code	Country name	Region name	Firms
AE	United Arab Emirates	Emerging/RoW	17
AR	Argentina	Emerging/RoW	30
AT	Austria	Europe	22
AU	Australia	Asia-Pacific (ex. Japan)	171
BE	Belgium	Europe	27
BM	Bermuda	Europe	63
BR	Brazil	Emerging/RoW	81
CA	Canada	North-America	269
CH	Switzerland	Europe	116
CL	Chile	Emerging/RoW	29
CN	China	Emerging/RoW	797
CO	Colombia	Emerging/RoW	12
DE	Germany	Europe	84
DK	Denmark	Europe	45
ES	Spain	Europe	34
FI	Finland	Europe	65
FR	France	Europe	121
GB	United Kingdom	Europe	303
GR	Greece	Emerging/RoW	11
HK	Hong Kong	Asia-Pacific (ex. Japan)	37
ID	Indonesia	Emerging/RoW	51
IE	Ireland	Europe	29
IL	Israel	Emerging/RoW	17
IN	India	Emerging/RoW	513
IT	Italy	Europe	11
JE	Jersey Channel Islands	Europe	12
JP	Japan	Japan	379
KR	South Korea	Emerging/RoW	134
KY	Cayman Islands	Europe	121
LU	Luxembourg	Europe	24
MA	Morocco	Emerging/RoW	15
MH	Marshall Islands	North-America	12

MX	Mexico	Emerging/RoW	36
MY	Malaysia	Emerging/RoW	98
NL	Netherlands	Europe	31
NO	Norway	Europe	59
NZ	New Zealand	Asia-Pacific (ex. Japan)	33
PE	Peru	Emerging/RoW	16
PH	Philippines	Emerging/RoW	25
PL	Poland	Emerging/RoW	26
PT	Portugal	Europe	10
QA	Qatar	Emerging/RoW	19
RU	Russian Federation	Emerging/RoW	30
SA	Saudi Arabia	Emerging/RoW	34
SE	Sweden	Europe	29
SG	Singapore	Asia-Pacific (ex. Japan)	33
TH	Thailand	Emerging/RoW	123
TR	Turkey	Emerging/RoW	57
TW	Taiwan	Emerging/RoW	140
US	United States	North-America	1,853
VN	Vietnam	Emerging/RoW	10
ZA	South Africa	Emerging/RoW	65

Bibliography

Abrigo, M.R.M., Love, I., 2016. Estimation of Panel Vector Autoregression in Stata. Stata J. Promot. Commun. Stat. Stata 16, 778–804. https://doi.org/10.1177/1536867x1601600314

Albuquerque, R.A., Koskinen, Y., Yang, S., Zhang, C., 2020. Resiliency of Environmental and Social Stocks: An Analysis of the Exogenous COVID-19 Market Crash. https://doi.org/10.2139/ssrn.3583611

Alves, R., Krueger, P., van Dijk, M.A., 2023. Drawing Up the Bill: Are ESG Ratings Related to Stock Returns Around the World? https://doi.org/10.2139/ssrn.4674146

Amihud, Y., 2002. Illiquidity and stock returns: cross-section and time-series effects. J. Financ. Mark. 5, 31–56. https://doi.org/10.1016/S1386-4181(01)00024-6

Anderson, T.W., Hsiao, C., 1982. Formulation and estimation of dynamic models using panel data. J. Econom. 18, 47–82. https://doi.org/10.1016/0304-4076(82)90095-1

Andrews, D.W.K., Lu, B., 2001. Consistent model and moment selection procedures for GMM estimation with application to dynamic panel data models. J. Econom. 101, 123–164. https://doi.org/10.1016/S0304-4076(00)00077-4

Arellano, M., Bover, O., 1995. Another look at the instrumental variable estimation of error-components models. J. Econom. 68, 29–51. https://doi.org/10.1016/0304-4076(94)01642-D

Azar, J., Duro, M., Kadach, I., Ormazabal, G., 2021. The Big Three and corporate carbon emissions around the world. J. Financ. Econ. 142, 674–696. https://doi.org/10.1016/j.jfineco.2021.05.007

Bauer, M.D., Huber, D., Rudebusch, G.D., Wilms, O., 2022. Where is the carbon premium? Global performance of green and brown stocks. J. Clim. Finance 1, 100006. https://doi.org/10.1016/j.jclimf.2023.100006

Becht, M., Franks, J.R., Grant, J., Wagner, H.F., 2017. Returns to Hedge Fund Activism: An International Study. https://doi.org/10.2139/ssrn.2376271

Benuzzi, M., Sahin, Ö., Paterlini, S., 2025. From KPIs to ESG: Addressing Redundancy and Distortions in ESG Scores. https://doi.org/10.2139/ssrn.5136069

Berg, F., Fabisik, K., Sautner, Z., 2021. Is History Repeating Itself? The (Un)Predictable Past of ESG Ratings. https://doi.org/10.2139/ssrn.3722087

Berg, F., Koelbel, J.F., Pavlova, A., Rigobon, R., 2022. ESG Confusion and Stock Returns: Tackling the Problem of Noise. Working Paper Series. https://doi.org/10.3386/w30562

Binder, M., Hsiao, C., Pesaran, M.H., 2005. Estimation and Inference in Short Panel Vector Autoregressions with Unit Roots and Cointegration. Econom. Theory 21, 795–837.

Black, F., 1986. Noise. J. Finance 41, 528–543. https://doi.org/10.1111/j.1540-6261.1986.tb04513.x

Bolton, P., Kacperczyk, M., 2021. Do investors care about carbon risk? J. Financ. Econ. 142, 517–549. https://doi.org/10.1016/j.jfineco.2021.05.008

Bowen, H.R., 1953. Social Responsibilities of the Businessman. University of Iowa Press. https://doi.org/10.2307/j.ctt20q1w8f

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

Canova, F., Ciccarelli, M., 2013. Panel Vector Autoregressive Models: A Survey. https://doi.org/10.2139/ssrn.2201610

Carroll, A.B., 2008. A History of Corporate Social Responsibility: Concepts and Practices, in: Crane, A., Matten, D., McWilliams, A., Moon, J., Siegel, D.S. (Eds.), The Oxford Handbook of Corporate Social Responsibility. Oxford University Press, p. 0. https://doi.org/10.1093/oxfordhb/9780199211593.003.0002

Chang, X., Zhang, H.F., 2015. Managerial Entrenchment and Firm Value: A Dynamic Perspective. J. Financ. Quant. Anal. 50, 1083–1103. https://doi.org/10.1017/S0022109015000423

Ciciretti, R., Dalò, A., Dam, L., 2023. The contributions of betas versus characteristics to the ESG premium. J. Empir. Finance 71, 104–124. https://doi.org/10.1016/j.jempfin.2023.01.004

Danielson, M.G., Heck, J.L., Shaffer, D., 2008. Shareholder Theory - How Opponents and Proponents Both Get it Wrong. https://doi.org/10.2139/ssrn.1309066

Darolles, S., Le Fol, G., He, Y., 2023. Understanding the Effect of ESG Scores on Stock Returns Using Mediation Theory. https://doi.org/10.2139/ssrn.4634699

Deng, X., Kang, J., Low, B.S., 2013. Corporate social responsibility and stakeholder value maximization: Evidence from mergers. J. Financ. Econ. 110, 87–109. https://doi.org/10.1016/j.jfineco.2013.04.014

Drempetic, S., Klein, C., Zwergel, B., 2020. The Influence of Firm Size on the ESG Score: Corporate Sustainability Ratings Under Review. J. Bus. Ethics 167, 333–360. https://doi.org/10.1007/s10551-019-04164-1

Dyck, A., Lins, K.V., Roth, L., Wagner, H.F., 2019. Do institutional investors drive corporate social responsibility? International evidence. J. Financ. Econ. 131, 693–714. https://doi.org/10.1016/j.jfineco.2018.08.013

Easterwood, J.C., İnce, Ö.Ş., Raheja, C.G., 2012. The evolution of boards and CEOs following performance declines. J. Corp. Finance, Special Section: Contemporary corporate finance research on South America 18, 727–744. https://doi.org/10.1016/j.jcorpfin.2012.05.006

Engle, R., Granger, C., 2015. Co-integration and error correction: Representation, estimation, and testing. Appl. Econom. 39, 106–135.

Erhart, S., 2022. Take it with a pinch of salt—ESG rating of stocks and stock indices. Int. Rev. Financ. Anal. 83, None. https://doi.org/10.1016/j.irfa.2022.102308

Escobar-Saldívar, L.J., Villarreal-Samaniego, D., Santillán-Salgado, R.J., 2025. The Effects of ESG Scores and ESG Momentum on Stock Returns and Volatility: Evidence from U.S. Markets. J. Risk Financ. Manag. 18, 367. https://doi.org/10.3390/jrfm18070367

European Commission, 2021. Overview of sustainable finance [WWW Document]. URL https://finance.ec.europa.eu/sustainable-finance/overview-sustainable-finance_en (accessed 7.19.25).

Fama, E.F., French, K.R., 2012. Size, value, and momentum in international stock returns. J. Financ. Econ. 105, 457–472. https://doi.org/10.1016/j.jfineco.2012.05.011

Fama, E.F., French, K.R., 1998. Value versus Growth: The International Evidence. J. Finance 53, 1975–1999. https://doi.org/10.1111/0022-1082.00080

Fama, E.F., French, K.R., 1993. Common risk factors in the returns on stocks and bonds. J. Financ. Econ. 33, 3–56. https://doi.org/10.1016/0304-405X(93)90023-5

Freeman, R.E., 2010. Strategic Management: A Stakeholder Approach. Cambridge University Press.

Freeman, R.E., McVea, J., 2001. A Stakeholder Approach to Strategic Management. https://doi.org/10.2139/ssrn.263511

Friedman, M., 1970. The Social Responsibility of Business Is to Increase Its Profits. N. Y. Times Mag.

Gabrielsen, A., Marzo, M., Zagaglia, P., 2011. Measuring Market Liquidity: An Introductory Survey. SSRN Electron. J. https://doi.org/10.2139/ssrn.1976149

Griffin, D., Guedhami, O., Li, K., Lu, G., 2021. National Culture and the Value Implications of Corporate Environmental and Social Performance. J. Corp. Finance 71, 102123. https://doi.org/10.1016/j.jcorpfin.2021.102123

Halbritter, G., Dorfleitner, G., 2015. The wages of social responsibility — where are they? A critical review of ESG investing. Rev. Financ. Econ. 26, 25–35. https://doi.org/10.1016/j.rfe.2015.03.004

Hamilton, J.D., 1994. Time Series Analysis. Princeton University Press. https://doi.org/10.2307/j.ctv14jx6sm

Hartmann, J., Pelster, M., Sievers, S., 2021. Can the market identify prosperous activist engagements? Evidence from announcement and long-term buy-and-hold returns. https://doi.org/10.2139/ssrn.3800001

Hayashi, F., Jagannathan, R., 1990. Ex-day behavior of japanese stock prices: New insights from new methodology. J. Jpn. Int. Econ. 4, 401–427. https://doi.org/10.1016/0889-1583(90)90019-3

He, F., Feng, Y., Hao, J., 2023. Corporate ESG rating and stock market liquidity: Evidence from China. Econ. Model. 129, 106511. https://doi.org/10.1016/j.econmod.2023.106511

Heath, D., Macciocchi, D., Michaely, R., Ringgenberg, M.C., 2022. Do Index Funds Monitor? Rev. Financ. Stud. 35, 91–131. https://doi.org/10.1093/rfs/hhab023

Hermalin, B.E., Weisbach, M.S., 2005. Boards of Directors as an Endogenously Determined Institution: A Survey of the Economic Literature.

Hoepner, A.G.F., Oikonomou, I., Sautner, Z., Starks, L.T., Zhou, X.Y., 2024. ESG shareholder engagement and downside risk. Rev. Finance 28, 483–510. https://doi.org/10.1093/rof/rfad034

Hoffmann, V.H., Busch, T., 2008. Corporate Carbon Performance Indicators. J. Ind. Ecol. 12, 505–520. https://doi.org/10.1111/j.1530-9290.2008.00066.x

Holtz-Eakin, D., Newey, W., Rosen, H.S., 1988. Estimating Vector Autoregressions with Panel Data. Econometrica 56, 1371–1395. https://doi.org/10.2307/1913103

Hong, H., Kacperczyk, M., 2009. The price of sin: The effects of social norms on markets. J. Financ. Econ. 93, 15–36. https://doi.org/10.1016/j.jfineco.2008.09.001

Horn, M., Oehler, A., 2024. Constructing stock portfolios by sorting on ESG ratings: Does the rating provider matter? Int. Rev. Financ. Anal. 96, 103568. https://doi.org/10.1016/j.irfa.2024.103568

Janakiraman, S.N., Lambert, R.A., Larcker, D.F., 1992. An Empirical Investigation of the Relative Performance Evaluation Hypothesis. J. Account. Res. 30, 53–69. https://doi.org/10.2307/2491091

Jensen, M.C., Black, F., Scholes, M.S., 2006. The Capital Asset Pricing Model: Some Empirical Tests.

Jones, R., Murrell, A.J., 2001. Signaling Positive Corporate Social Performance: An Event Study of Family-Friendly Firms. Bus. Soc. 40, 59–78. https://doi.org/10.1177/000765030104000105

Karpoff, J.M., 2001. The Impact of Shareholder Activism on Target Companies: A Survey of Empirical Findings. https://doi.org/10.2139/ssrn.885365

Khan, M., Serafeim, G., Yoon, A., 2016. Corporate Sustainability: First Evidence on Materiality. https://doi.org/10.2139/ssrn.2575912

Krueger, P., Sautner, Z., Tang, D.Y., Zhong, R., 2024. The Effects of Mandatory ESG Disclosure Around the World. J. Account. Res. 62, 1795–1847. https://doi.org/10.1111/1475-679X.12548

Kumar Mandal, S., Madheswaran, S., 2010. Causality between energy consumption and output growth in the Indian cement industry: An application of the panel vector error correction model (VECM). Energy Policy, Energy Efficiency Policies and Strategies with regular papers. 38, 6560–6565. https://doi.org/10.1016/j.enpol.2010.07.042

Kumar, S., Sharma, D., Rao, S., Lim, W.M., Mangla, S.K., 2025. Past, present, and future of sustainable finance: insights from big data analytics through machine learning of scholarly research. Ann. Oper. Res. 345, 1061–1104. https://doi.org/10.1007/s10479-021-04410-8

Lehavy, R., Sloan, R.G., 2005. Investor Recognition and Stock Returns. https://doi.org/10.2139/ssrn.817066

Lewandowski, S., 2017. Corporate Carbon and Financial Performance: The Role of Emission Reductions. Bus. Strategy Environ. 26, 1196–1211. https://doi.org/10.1002/bse.1978

Lin, W.L., Law, S.H., Ho, J.A., Sambasivan, M., 2019. The causality direction of the corporate social responsibility – Corporate financial performance Nexus: Application of Panel Vector Autoregression approach. North Am. J. Econ. Finance 48, 401–418.

Lins, K.V., Servaes, H., Tamayo, A., 2016. Social Capital, Trust, and Firm Performance: The Value of Corporate Social Responsibility during the Financial Crisis. https://doi.org/10.2139/ssrn.2555863

London Stock Exchange Group, 2025. Company Data - Environmental, Social and Corporate Governance - ESG [WWW Document]. LSEG Data Anal. URL https://www.lseg.com/en/data-analytics/financial-data/company-data/esg-data (accessed 3.31.25).

London Stock Exchange Group, 2024. Environmental, Social and Governance scores from LSEG.

Luo, D., 2022. ESG, Liquidity, and Stock Returns. J. Int. Financ. Mark. Inst. Money 78. https://doi.org/10.1016/j.intfin.2022.101526

Lütkepohl, H., 2005. New Introduction to Multiple Time Series Analysis. Springer Books.

Meng-tao, C., Da-peng, Y., Wei-qi, Z., Qi-jun, W., 2023. How does ESG disclosure improve stock liquidity for enterprises — Empirical evidence from China. Environ. Impact Assess. Rev. 98, 106926. https://doi.org/10.1016/j.eiar.2022.106926

Merton, R.C., 1987. A Simple Model of Capital Market Equilibrium with Incomplete Information [WWW Document]. URL https://onlinelibrary.wiley.com/doi/10.1111/j.1540-6261.1987.tb04565.x (accessed 8.22.25).

Molino, A., Prosperi, L., Zicchino, L., 2023. Unveiling the Liquidity Greenium: Exploring Patterns in the Liquidity of Green versus Conventional Bonds. https://doi.org/10.2139/ssrn.4503175

Nagy, Z., Kassam, A., Lee, L.-E., 2016. Can ESG Add Alpha? An Analysis of ESG Tilt and Momentum Strategies. J. Invest. 25, 113–124. https://doi.org/10.3905/joi.2016.25.2.113

Napier, E., Knight, G., Luo, Y., Delios, A., 2023. Corporate social performance in international business. J. Int. Bus. Stud. 54, 61–77. https://doi.org/10.1057/s41267-022-00584-w

Nickell, S., 1981. Biases in Dynamic Models with Fixed Effects. Econometrica 49, 1417–1426. https://doi.org/10.2307/1911408

Novy-Marx, R., 2013. The other side of value: The gross profitability premium. J. Financ. Econ. 108, 1–28. https://doi.org/10.1016/j.jfineco.2013.01.003

Pastor, L., Stambaugh, R.F., Taylor, L.A., 2022. Dissecting Green Returns. https://doi.org/10.2139/ssrn.3864502

Pedersen, L.H., Fitzgibbons, S., Pomorski, L., 2021. Responsible investing: The ESG-efficient frontier. J. Financ. Econ. 142, 572–597. https://doi.org/10.1016/j.jfineco.2020.11.001

Pedersen, L.H., Fitzgibbons, S., Pomorski, L., 2019. Responsible Investing: The ESG-Efficient Frontier. https://doi.org/10.2139/ssrn.3466417

Qureshi, M.A., Akbar, M., Poulová, P., Akbar, A., 2021. Do ESG Endeavors Assist Firms in Achieving Superior Financial Performance? A Case of 100 Best Corporate Citizens. ResearchGate. https://doi.org/10.1177/21582440211021598

Renneboog, L., Ter Horst, J., Zhang, C., 2008. The price of ethics and stakeholder governance: The performance of socially responsible mutual funds. J. Corp. Finance, Special Issue: Contractual Corporate Governance 14, 302–322. https://doi.org/10.1016/j.jcorpfin.2008.03.009

Roodman, D., 2009. How to do Xtabond2: An Introduction to Difference and System GMM in Stata. Stata J. 9, 86–136. https://doi.org/10.1177/1536867X0900900106

Ruf, B.M., Muralidhar, K., Brown, R.M., Janney, J.J., Paul, K., 2001. An Empirical Investigation of the Relationship Between Change in Corporate Social Performance and Financial Performance: A Stakeholder Theory Perspective. J. Bus. Ethics 32, 143–156. https://doi.org/10.1023/A:1010786912118

Shackleton, M., Yan, J., Yao, Y., 2022. What drives a firm's ES performance? Evidence from stock returns. J. Bank. Finance 136, 106304. https://doi.org/10.1016/j.jbankfin.2021.106304

Sharpe, W.F., 1964. Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk. J. Finance 19, 425–442. https://doi.org/10.1111/j.1540-6261.1964.tb02865.x

Sims, C.A., 1980. Macroeconomics and Reality. Econometrica 48, 1–48. https://doi.org/10.2307/1912017

The Global Sustainable Investment Alliance, 2022. Global Sustainable Investment Review 2022.

Trumpp, C., Guenther, T., 2017. Too Little or too much? Exploring U-shaped Relationships between Corporate Environmental Performance and Corporate Financial Performance. Bus. Strategy Environ. 26, 49–68. https://doi.org/10.1002/bse.1900

Useche, A.J., Martínez-Ferrero, J., Reyes, G.E., 2024. Taking ESG strategies for achieving profits: a dynamic panel data analysis. J. Econ. Finance Adm. Sci. 30, 61–78.

van Emous, R., Krušinskas, R., Westerman, W., 2021. Carbon Emissions Reduction and Corporate Financial Performance: The Influence of Country-Level Characteristics. Energies 14, 6029. https://doi.org/10.3390/en14196029

Wang, K., Li, T., San, Z., Gao, H., 2023. How does corporate ESG performance affect stock liquidity? Evidence from China. Pac.-Basin Finance J. 80, 102087. https://doi.org/10.1016/j.pacfin.2023.102087

Watson, R., 2011. Ethical and socially responsible investment funds. J. Financ. Regul. Compliance 19, 100–110. https://doi.org/10.1108/13581981111123825

Wedajo, A.D., Salah, A.A., Bhat, Mohd.A., Iqbal, R., Khan, S.T., 2024. Analyzing the dynamic relationship between ESG scores and firm value in Chinese listed companies: insights from generalized cross-lagged panel model. Discov. Sustain. 5, 336. https://doi.org/10.1007/s43621-024-00546-2

Wilson, R., 1997. Islamic finance and ethical investment. Int. J. Soc. Econ. 24, 1325–1342. https://doi.org/10.1108/03068299710193624

Yu, B., Wu, S., Lenard, M.J., 2022. Do Ethical Companies Have High Stock Prices or High Returns? J. Risk Financ. Manag. 15, 81. https://doi.org/10.3390/jrfm15020081

Yu, D., Meng, T., Zheng, M., Ma, R., 2024. ESG uncertainty, investor attention and stock price crash risk in China: evidence from PVAR model analysis. Humanit. Soc. Sci. Commun. 11, 1152. https://doi.org/10.1057/s41599-024-03621-1

Zhang, C., Hao, D., Gao, L., Xia, F., Zhang, L., 2024. Do ESG ratings improve capital market trading activities? Int. Rev. Econ. Finance 93, 195–210. https://doi.org/10.1016/j.iref.2024.03.027