WISEflow
Europe/Oslo(CEST)
28 Jun 2023

Handelshøyskolen BI

GRA 19703 Master Thesis

Thesis Master of Science 100% - W

Predefinert informasjon

Startdato:	09-01-2023 09:00 CET	**Termin:**	202310
Sluttdato:	03-07-2023 12:00 CEST	**Vurderingsform:**	Norsk 6-trinns skala (A-F)
Eksamensform:	T		
Flowkode:	202310||11184||IN00||W||T		
Intern sensor:	(Anonymisert)		

Deltaker

Navn:	Camilla Vannebo og Kristine Wilhelmsen

Informasjon fra deltaker

Tittel *:	Unraveling the Link between Environmental Metrics and Stock Returns: An In-Depth Analysis of Climate Risk Premiums and Climate News Hedge Portfolios
Navn på veileder *:	Giovanni Pagliardi

Inneholder besvarelsen konfidensielt materiale?:	Nei	**Kan besvarelsen offentliggjøres?:**	Ja

Gruppe

Gruppenavn:	(Anonymisert)
Gruppenummer:	227
Andre medlemmer i gruppen:	

Unraveling the Link between Environmental Metrics and Stock Returns: An In-Depth Analysis of Climate Risk Premiums and Climate News Hedge Portfolios

Master Thesis

by

Camilla Vannebo and Kristine Wilhelmsen

MSc in Business with Major in Finance

Supervisor

Giovanni Pagliardi

Oslo, June 28, 2023

Acknowledgements

We would like to sincerely thank our supervisor, Giovanni Pagliardi, for his guidance and support in writing this thesis. The support and encouragement he provided were crucial throughout the process of completing this semester and the thesis. We would also like to thank colleagues, friends and family for continuous support, insightful discussions, and unconditional motivation.

Abstract

This thesis examines a broad range of E(SG) scores and real environmental metrics' cross-sectional and time-series effect on excess stock return, in the European market from December 2010 to December 2022. The objective is to analyze how E(SG) scores and environmental metrics differ in their ability to capture climate risk exposure. Through Fama MacBeth regressions and characteristics-based portfolio sorting, the study reveals a compelling insight; adopting an investment strategy that goes long environmental underperforming firms (i.e. "brown") and short overperforming firms ("green") results in statistically significant risk premiums ranging from -7.83% to 5.22% annually. Furthermore, our analysis reveals a weak relationship between environmental metrics and E(SG) scores, extending existing research on ESG disagreement. Finally, using a mimicking portfolio approach we show that our sample of environmental variables proves insufficient in hedging innovations in climate change news.

Contents

Chapter 1

Introduction

As climate change materializes and the occurrence of extreme weather events increases, policymakers are pressured to introduce regulations that aim to limit companies' contribution to global warming. Consequently, over the past decade, we have witnessed a surge in the adoption of ESG and integrated reporting that is supposed to reflect companies' climate efforts, footprint, and climate risk. However, these measures are largely susceptible to bias and subjective judgments by rating agencies (Berg et al., 2022). Researchers have found contradicting results when exploring the relationship between ESG scores and stock return. While some research shows that high-emission portfolios provide positive stock excess return (Bolton & Kacperczyk, 2021), other research shows that high ESG-score portfolios outperform the benchmark (Pollard & Sherwood, 2018). In light of this, our finding of highly diverging risk premiums, ranging from -7.83% to 5.22%, is intriguing. Additionally, there is limited research on real environmental metrics' relationship with stock return and ESG scores. As climate change awareness and the preference for ESG stocks increases rapidly and unpredictably, research within the field of ESG is highly dependent on the time period studied (Pástor et al., 2021). It is evident that more research and a better understanding of the relationship between climate risk and stock return are needed. Consequently, we take a closer look at ESG scores and quantifiable environmental metrics in this study;

what these variables are based upon, and their effect on stock return, in order to explore whether investors truly consider climate risk in investment decisions.

Furthermore, we study the performance of portfolios based on these metrics during news about climate change events. By constructing and investing in portfolios that correlate with news about climate change, investors are essentially hedged against the realization of climate change events (Engle et al., 2020). Most research in this field has been conducted in American markets, while we will analyze companies on the STOXX Europe 600 index. According to Société Générale, the European Union now has the most advanced and extensive regulatory framework in the world, much due to the European Action Plan for Sustainable Finance ("EU Action Plan on Sustainable Finance", 2023). This way, the European market may be less susceptible to biases arising from insufficient sustainability reporting.

Investors seeking to manage climate risk rely on ESG data as their best proxy in investment decisions. According to Eccles et al. (2011), ESG factors have been discussed in the academic literature for more than 35 years. Various third-party agencies like Bloomberg, CDP, Sustainalytics, MSCI, Refinitiv, and S&P Global have created a methodology to score companies based on their ESG performance. However, though these ESG scores aim to reflect a realistic picture of a company's performance in the different categories, there is low correlation and high divergence between ratings (Berg et al., 2022). Even when we adjust for explicit differences in the definition of corporate social responsibility (CSR) held by different vendors, ESG ratings disagree substantially, implying the ratings have low validity (Chatterji et al., 2016).

Given the ambiguity in ESG data already posed by other researchers, exploring the extent and level of this ambiguity in our own sample is a natural first step of our research. In addition to studying several E(SG) scores, which encompass both E and ESG scores, we introduce ten additional environmental metrics such as carbon emissions, total waste, renewable energy usage, etc., hereby referred to as ”environmental metrics”. Combined with E(SG) scores, these metrics serve as firm-specific proxies for climate risk exposure. The environmental metrics are deliberately chosen to be so-called ”non-greenwashable”, as they are directly linked to environmental performance and are not estimated by third-party vendors. We analyze the relationship between the E(SG) scores and environmental metrics, providing new insight into what the ESG scores are actually based upon and whether environmental metrics are a better proxy for climate risk. One would think that these environmental metrics provide a more consistent relationship with stock return, but that is not always the case. Metrics such as carbon emissions have been proven by Bolton and Kacperczyk (2021) to provide a positive risk premium, while Gibson Brandon et al. (2021) on their end found that the companies where ESG ratings disagree the most provide a higher return. We therefore study our broad spectre of metrics and assess their effect on excess return, and whether there exists a premium on any of these metrics when using them to form green portfolios. To ensure the validity of our results we construct risk premiums in two ways; characteristics-based portfolio sorting, and Fama MacBeth cross-sectional regression.

There are some natural limitations to such risk premiums given the ambiguity and availability of ESG data and the rapidly increasing climate change awareness. An investor that wishes to hedge using ESG scores as a proxy for climate risk exposure might miss the mark completely given the fact that measurement is the biggest source of divergence (Berg et al., 2022). To give proactive investors insight into how they can hedge themselves against the realization of climate change, we show how one can use the same metrics to hedge climate change news, using a mimicking portfolio approach as shown by Engle et al. (2020). Our findings on what each E(SG) score actually measures bring relevant insight when creating climate risk premiums and performing the hedging.

In light of limited research on the link between climate risk exposure and stock return, our master thesis seeks to address a fundamental research question:

How do ESG scores and measurable environmental metrics differ in their ability to capture climate risk exposure, and what metrics are better suited to hedge climate risk?

Our hypothesis is that companies' sustainability reporting does not necessarily reflect their real climate change resilience or mitigating efforts, particularly not their E(SG) scores. With that in mind, we will test whether the link between E(SG) scores and measurable environmental metrics holds true. Furthermore, we will study whether there are portfolio(s) that provide a risk premium, and if so, in what way environmental performance affects stock excess return. The final step of our research will be to test all of our E(SG) scores and environmental metrics against the climate news index developed by Engle et al. (2020), to see whether any of the identified portfolios can serve as a hedge against climate risk exposure. This will bring insight into which scores may be more prone to "green-washing" and whether they truly measure environmental performance.

Chapter 2

Literature Review

2.1 ESG ambiguity

With ESG scores being a relatively new phenomenon, and its constituents being more difficult to report than standard financial numbers, assessing and determining the correct score can be a challenge for rating agencies. This is proven in the literature by the divergence in ESG scores, where Chatterji et al. (2016) has shown that ESG ratings from different providers disagree substantially. Furthermore, Berg et al. (2022) found varying definitions between rating agencies and fundamental disagreement about the underlying data. According to their research, the divergence among the most established ESG rating agencies mainly comes from differences in measurement (56% of the divergence), in addition to scope (38%), and weight (6%). Another interesting finding is that the main reason for measurement divergence is the rater effect, where the rating agency's view of the firm influences its measurement of the different categories, indicating that ESG ratings are not based on objective observations that can be ascertained. Therefore, the authors further stress the importance of paying more attention to how rating agencies generate ESG data. Grundström and Miedel (2021) specifically studied the relationship between sustainability scores and emissions and found that relationships varied between rating agencies, indicating that a high E score is not necessarily associated with lower emissions. We extend the research of Berg et al. (2022) by studying the link between a broad span of environmental metrics and established E(SG) scores. This way, we explore the divergence of E(SG) scores further and determine whether rating agencies actually measure what they say they measure.

Further substantiating the presence of ambiguity in ESG data, Gibson Brandon et al. (2021) explores the severity of ESG rating disagreement and discovers that stock returns are positively related to ESG rating disagreement, suggesting a risk premium for firms with higher ESG rating disagreement. The relationship is primarily driven by disagreement in the environmental dimension, further exemplifying the difficulties of using E(SG) scores to form an investment strategy.

2.2 Pricing of climate risk

Several studies have shown how investors deal with climate risk, and the research presents contradictory results. One of the most esteemed research papers on this topic is that of Bolton and Kacperczyk (2021), showing that there exists a risk premium on US stocks with higher total carbon emissions (Scope 1, 2, and 3), as well as higher changes in emissions. They use a standard cross-sectional regression model with pooled OLS, regressing Scope 1, 2, and 3 emissions, the year-to-year change in emissions, and the emissions intensity, to the corresponding stock returns. They include well-known, firm-specific control variables such as LOGSIZE, B/M, ROE, LEVERAGE, MOM, etc. (further explained in Chapter 3.2). Their research also reveals that institutional investors implement exclusionary screening based on direct emission intensity in certain industries. However, the carbon premium observed in stock returns is not solely driven by sin stock divestment, as divestment mainly occurs in specific industries such as oil and gas, utilities, and automobiles. Interestingly, there is no carbon premium associated with emission intensity outside these industries. The findings indicate that investors are already demanding compensation for their exposure to carbon emission risk, suggesting that they are pricing in the potential impact of carbon risk on firms. We draw on the methodology and research question this paper poses for the computation of our risk premiums, choosing the same eleven control variables used in this study in the cross-sectional regression. Rather than solely analyzing the relationship between emissions and stock return, we add more observable and objective environmental metrics, in addition to using a European sample. We are curious to see whether we are able to observe a risk premium for carbon emissions in our sample, and how it potentially differs from the findings of Bolton and Kacperczyk (2021).

Another study had somewhat opposite findings, identifying a significant premium on portfolios based on high ESG score-companies, i.e. "green" firms. Pollard and Sherwood (2018) measured the effect on the risk-adjusted return of including ESG data in a global equity portfolio (over 70% of the portfolio consists of American and Canadian stocks). They constructed two portfolios, a benchmark, and an ESG-integrated portfolio, to be observed and measured from 2007 to 2017. The portfolios were rebalanced quarterly, which for the ESG portfolio meant that the lowest performing stocks were replaced with stocks that had the relatively highest improvement in ESG score that quarter. Their results show that the integration of ESG factors into the ESG portfolio generated a consistent alpha over the period analyzed and a higher Sharpe Ratio than the benchmark portfolio. They conclude that ESG as a risk premium should be included alongside other well-known risk factors. Under the assumption that excess return of high ESG score companies should be somewhat correlated with excess return of low carbon emission companies, this finding contradicts that of Bolton and Kacperczyk (2021), who show that high carbon emission companies earn a relatively higher return.

To further substantiate the ambiguity in ESG data, Pástor et al. (2021) uses an equilibrium model to show that "green" assets have low expected returns, both because investors prefer holding them and because they act as a hedge against climate risk. However, they find that green assets outperform when positive shocks occur in the ESG factor, which captures shifts in customer and investor preferences for green products and holdings. The pricing discrepancy shown between green and brown stocks' alphas arises from investors' preferences for green holdings and the ability of green stocks to hedge climate risk. The authors note that their model describes the present and future world and that its applicability to different time periods should be explored further. We share the view that findings within the ESG sphere could be sample specific given the recent adoption of ESG reporting and lacking reporting standards in many areas. It is difficult to distinguish ex-ante versus ex-post effects of ESG concerns by looking at realized returns over periods during which ESG tastes shift (Pástor et al., 2021), which further underpins the need for more research in this field. The findings of Pástor et al. (2021) portray a similar risk story to that of Bolton and Kacperczyk (2021), and introduce the idea of hedging climate risk using market information about shocks in the ESG factor.

2.3 Hedging climate risk

Computing risk premiums and learning about how investors relate to climate risk gives important insight into how one can manage this risk. Well-known papers on the topic of hedging climate risk include Andersson et al. (2016), which show how to construct a portfolio of companies with a substantially lower carbon footprint than a US benchmark, with the idea that the portfolio will outperform the benchmark once carbon is priced in the market. This strategy might be suitable for long-time passive investors, but it fails to show the ex-post performance of the portfolio and take a dynamic approach. Additionally, there is a risk that this hedging strategy loses its efficiency as companies increasingly incorporate climate mitigation policies.

Engle et al. (2020) present a more dynamic approach, and their heavily cited and acknowledged paper on climate news hedging serves as the foundation for our final analysis. The paper demonstrates how to construct climate change hedge portfolios in the US market using a mimicking portfolio approach that performs well during news about climate change. They extract climate news series from textual analysis of news sources that capture the intensity of climate change discourse in the media (further explained in Chapter 3.2), serving as proxies for climate risk exposure. Engle et al. (2020) then construct hedge portfolios by projecting innovations in the climate news indices onto portfolios sorted based E scores from Sustainalytics and MSCI. By sorting the stocks in portfolios based on the characteristics that proxy for a firm's exposure to climate risk, the weights of the hedge portfolios are then parameterized in a mimicking portfolio approach. The effectiveness of the methodology is evaluated by comparing the performance of the hedge portfolios with alternative green ETFs. Finally, the hedge portfolios are regressed on the climate news indices along with the risk factors Size, Book Value, and Market. The resulting R-squared of the regressions shows that the hedge portfolios based on Sustainalytics' E-scores have the best fit, hedging 15-19% of the in-sample variation. This research methodology contributes to the literature on climate change and asset markets by providing a systematic framework for constructing climate risk hedge portfolios using publicly traded assets and news media data. The authors encourage further exploration and research in various aspects of their study. One such aspect is the differentiation between physical and regulatory-oriented climate risks as portrayed in the news media. In our own work, we address this distinction by including the carbon intensity of firms, a commonly used measure of regulatory climate risk. Given the relatively advanced regulatory environment in Europe compared to other financial markets, we believe it would be insightful to observe how these metrics perform.

Additionally, we extend the research of Engle et al. (2020) by performing climate risk hedges using a wider range of E(SG) scores and environmental metrics and analyzing them in relation to their risk premiums.

2.4 Gaps in the literature

Through our extensive literature search, we are left with two main observations. First, there exists a lot of uncertainty regarding the integrity and validity of the scores, as research in this field shows that there is high divergence and disagreement between rating agencies. With the overhanging risk of "greenwashing" when using E(SG) scores as a proxy for climate risk exposure, we wish to extend current research by introducing measurable environmental metrics and testing how they correlate with established ESG ratings. We also want to see whether the twelve-month change in scores and metrics can explain stock return and the relationship between scores and metrics.

Second, there have been numerous attempts at creating risk premiums based on ESG scores (Gibson Brandon et al., 2021) and emissions (Bolton & Kacperczyk, 2021), though the use of environmental data has been limited. As Pástor et al. (2021) point out in their study, it is challenging to conclude on ESG data's effect on stock return when the taste for ESG stocks is changing and data is time- and sample-specific. The most known papers in this field are studying US stocks, making it an interesting angle to test these relationships in a different developed market. We have yet to see research on climate news hedging in the European market and with such extensive use of environmental metrics and ESG scores, which is why we include the climate news hedging approach by Engle et al. (2020) as a last step of our analysis.

Chapter 3

Data

3.1 Data Sample and Screening

The analyzed period in this study is from December 2010 to December 2022 and covers 145 months of data. The period is limited to these years for two reasons. First, and as mentioned previously, it is only relevant to measure recent years as historical data neither includes realizations of extreme climate change effects nor includes the types of policies that could emerge going forward. Second, reporting on ESG scores and environmental metrics is a new phenomenon, and there is very limited data to be found before 2010. The necessary data for financial numbers, carbon emissions, ESG scores, etc. is available through the Bloomberg and Eikon terminal. One issue is the reliability of data on greenhouse gas (GHG) emissions and ESG scores. We have therefore obtained E(SG) scores from several agencies, such as Refinitiv, Sustainalytics, Bloomberg, CDP, and S&P Global. Furthermore, by scoping our research around European companies covered by the Non-Financial Reporting Directive (NFRD), requiring them to provide detailed reports on the company's environmental impact, we try to minimize the risk of reporting error and green-washing.

3.1.1 Selection of Assets

We believe climate risk to be universal, and that every company is in some way exposed to it. Geographical limitations are therefore not very important. However, The European Union (EU) is among the leading major economies when it comes to tackling GHG emissions, with the world's biggest and first major carbon market, EU Emissions Trading System, according to the European Commission. The availability of emissions- and ESG data is important to avoid sources of bias, for example where the data sample is over-represented by companies with high ESG performance. To minimize this risk, we are using STOXX Europe 600 listed companies. The sample has the advantage of covering European countries which have diverse legal systems. Furthermore, as the index consists of the largest companies in Europe, the companies' returns are not dominated by market microstructure issues (Engle et al., 2020). Finally, by scoping our study to the European stock market, we are differentiating ourselves from the majority of existing research in this field which has largely been focusing on the US market.

3.1.2 Selection of Timeframe

Taking the nature of climate change into account, and how (parts of) the world have only recently become aware of its associated risks, we believe it is accurate to use a relatively short time frame for the analysis. The first IPCC3 assessment report was published in 1990, underlining the importance of climate change as a challenge with global consequences. However, it was only in recent years that the broader public accepted the science and became fully aware of the immense negative impacts of a warming planet, the Paris Agreement in 2015 being a testament to this. Historical data is of little use for analyzing climate risk as it neither includes realizations of extreme climate change effects nor the types of policies that could emerge going forward (A. van Dijk, 2020). This is further emphasized by Bolton and Kacperczyk (2021) when they perform the 2005 cross-sectional distribution of total emissions to the 1990s, and find that there is no significant carbon premium. When they do the same for the 2017 cross-sectional regression, they find a significant carbon risk premium. Based on this, we believe the time period from 2010-2022 will be sufficient to study the effects of the materializing of climate risk.

3.2 Variable Description

In this section, we explain what the different ESG scores, environmental metrics, control variables, and financial metrics are based on according to the providers of these metrics. We also briefly comment on the rationale behind certain metrics where we think it is necessary.

3.2.1 E(SG) Scores and Environmental Metrics

Metric	Description
BB_ESG	Bloomberg ESG score, 0-10 where 10 is the best
CDP_CC_P	Carbon Disclosure Project (CDP) score, 1-5 where 5 is the best
CDP_Reg	CDP's regulatory risk score, dummy variable giving 1 if a company states it is prone to regulatory climate risk
Ref_E	Refinitiv E score, -100 to 100 where 100 is the best
SA_ESG	Sustainalytics E score, 0-100 where 0 is the best
SP_E	S&P's Environmental pillar score, 0-100 where 100 is the best
CO2intensity	Scope 1 and 2 carbon emissions divided by revenues
CDP_S1	Scope 1 emissions in MtCO2e, reported by CDP
CDP_S2	Scope 2 emissions in MtCO2e, reported by CDP
Green_rev	Percentage of the company's revenue that comes from green sources
Inv_Op_Sust	EUR mn spent on operational environmental and social compliance
Inv_Sust_Prod	EUR mn of investment in sustainable products
Raw_Material_Used	Total amount of raw materials consumed, in thousand Mt
Tot_Energy_Cons	Thousand MWh
Tot_Waste	Total amount of waste the company discards, in thousand Mt
RES_Use	Energy consumed generated by a renewable energy source, in thousand MWh.

Table 3.1: **Environmental metrics in sample:** The table gives a short description of each E(SG) score and environmental metric used throughout the report. With regard to the E(SG) scores, we define "best" as the most sustainable.

3.2.2 Financial Metrics

The financial metrics are obtained from the Bloomberg terminal, all reported in Euros. They include the companies on the STOXX 600 Index's revenues, capital expenditures, net income, PPE (plant, property, and equipment), return on equity (ROE), market beta, earnings per share, share price, book value, and market capitalization. Additionally, we have through the Bloomberg terminal obtained the 3-month annualized EURIBOR interest rate to calculate the excess return of our sample (see chapter 4.1 for calculation).

3.2.3 Control Variables

We use the same control variables as those used by Bolton and Kacperczyk (2021). These reflect the most common control variables to include in cross-sectional regression and are based on previous studies of Fama and MacBeth (1973) and Fama and French (1992). Hereunder follows an explanation of the different control variables and their construction.

Variable	Description
$\text{LOGSIZE}_{i,t}$	Natural logarithm of firm i's market capitalization
$\text{B/M}_{i,t}$	Firm i's book value divided by market capitalization
LEVERAGE	Book leverage of the company (D/E)
$\text{ROE}_{i,t}$	Firm i's earnings performance
$\text{MOM}_{i,t}$	Average of the most recent 12 months' returns on stock i
INVEST/A	Capital expenditures divided by book value of assets
LOGPPE	Natural logarithm of the firm's property, plant, and equipment
$\text{BETA}_{i,t}$	Market beta of firm i in year t
$\text{VOLAT}_{i,t}$	Standard deviation of returns based on the past 12 months
$\text{SALESGR}_{i,t}$	Dollar change in annual firm revenues normalized by last month's market capitalization
$\text{EPSGR}_{i,t}$	Dollar change in annual earnings per share normalized by the firm's equity price

Table 3.2: **Control variables:** This table gives an overview and short description of the control variables we have included in the cross-sectional.

3.2.4 News Index

For the final part of our analysis, we use the Climate News Index created by Engle et al. (2020), collected from the website of Johannes Stroebel (NYU, n.d.). The data is a monthly time series from 2009-2017 based on textual analysis of the daily Wall Street Journal issue. See detailed explanation about the construction of the index in Appendix A.3.2).

3.3 Descriptive Statistics

We calculate the standard deviation to test whether it is necessary to exclude any of the variables. As we can see in Table 3.3, there are no variables with a standard deviation equal to zero, meaning that we keep all variables in the sample. The number of observations (N) per variable is sufficient, and the number of observations increases during the analyzed time period (Appendix A.2). However, we do observe a lower number of observations for Green_rev and Inv_Sust_Prod, which might have implications later on in our analysis.

Figure 3.1 demonstrates the correlation between the different measures. Not surprisingly, the highest correlations are between E(SG) scores like S&P Global E score and Bloomberg ESG score as the highest one (0.51), and between consumption metrics like Total energy consumption and Reported MtCO2 equivalents (0.61). However, most correlations are lower than expected, indicating that E(SG) scores are not to a large extent capturing the variations of the environmental metrics in our sample. In light of existing literature on the rater effect (Chapter 2.1), these findings emphasize investors' need to find other ways to identify the most sustainable stocks, as ESG scores are so dependent on the vendor that made them.

Variable	Std. Dev.	Mean	Median	N
BB_ESG	2.072	3.65	3.58	47016
SP_E	37.637	67.16	71.00	41652
SA_ESG	7.754	19.76	19.08	14148
Ref_E	21.163	9.85	−1.22	30852
CDP_S1	17.900	4.34	0.07	56177
CDP_S2	1.494	0.73	0.12	52158
CDP_Reg_Risk	0.470	1.00	1.00	28734
CDP_CC_Performance	1.868	3.39	4.00	52940
Green_rev	0.032	0.44	0.36	300
Inv_Op_Sust	223.650	163.89	20.00	14283
Inv_Sust_Prod	188.110	948.11	112.50	696
RES_Use	3,791.857	1,681.45	117.87	34104
Tot_Energy_Cons	43,503.080	13,933.45	595.10	58690
Raw_Material_Used	30,570.740	21,728.75	550.00	11665
Tot_Waste	104,126.600	13,414.73	34.82	47742
CO2intensity	0.006	0.00037	0.00003	57509

Table 3.3: **Summary statistics:** This table provides an overview of the summary statistics of the E(SG) scores and environmental metrics used in our analysis, including their standard deviation, mean, median and number of observations

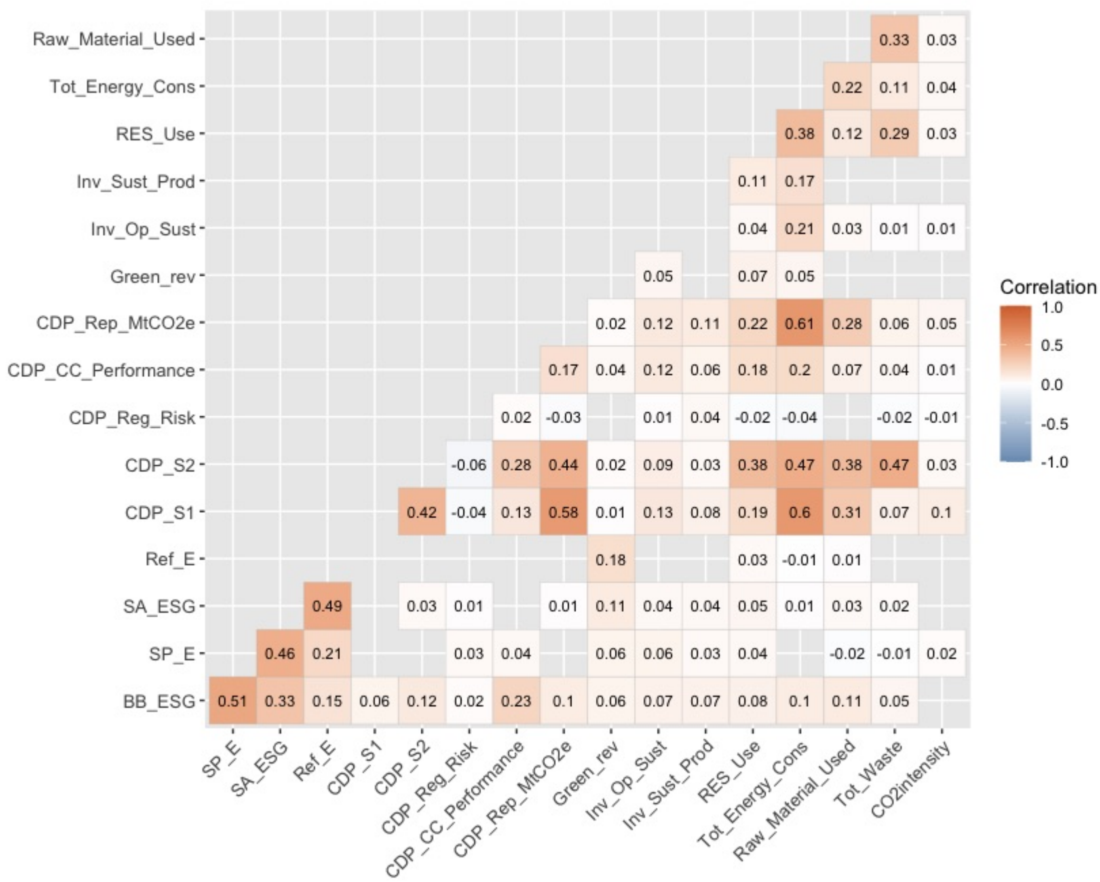


Figure 3.1: **Correlation plot of all environmental metrics in sample:** This figure displays the correlation between each E(SG) score and environmental metric. No correlation value implies a non-significant relationship.

Chapter 4

Methodology

The following section focuses on the structure and models that form the foundation of our research; exploring the ambiguity of E(SG) scores and environmental metrics, uncovering significant risk premiums among them, and performing the climate news hedge. We go through, step by step, how we build our analysis and the various regressions performed at each stage.

4.1 Linear regression

Before performing any regressions, each variable containing environmental metrics or E(SG) scores is standardized to ensure that all variables have a mean of zero and a standard deviation of one across the firms in the sample. We do not standardize the response variable (i.e. stock excess return), as we are interested in predicting its absolute value, not its standardized value.

The excess return used throughout all the regressions is computed by taking the monthly return of each stock and subtracting the risk-free rate (RF), previously defined as the EURIBOR 3-month annualized interest rate. This way we have for each company (i) their excess return (R):

$$R_{i,t} = \frac{Price_{i,t} - Price_{i,t-1}}{Price_{i,t-1}} - RF \tag{4.1}$$

We run all linear regressions in chapter 4.1 twice. First, we use the standardized environmental metrics and E(SG) scores. Second, we perform the same regressions using the *delta* of each metric per period, i.e. the 12-month change in value. By doing this we wish to understand whether excess return is more affected by significant changes in environmental performance or relative performance compared to peers. The delta of a given metric or score k is defined as the following:

$$\Delta k = k_t - k_{t-12}$$

Where Δk represents the change in the metric k, k_t denotes the value of the metric at time t, and k_{t-12} represents the value of the metric 12 time periods ago.

4.1.1 Does our environmental metrics forecast return?

Simple linear regression without fixed effects

We start by estimating a single intercept and slope for each environmental metric across all companies, under the assumption that there is no unobserved heterogeneity across the companies. We do this to see the basic effect of the independent variable on the dependent variable, before including control variables and comparing the results. We perform one simple linear regression for each environmental metric and E(SG) score in time (t) on excess return as the dependent variable (Y) in time (t+1) (see equation 4.2). We run this time-series regression across all 600 companies (i) in the sample. By studying the coefficients of the independent variables we can interpret each variable's effect on excess return, and explore whether this relationship is consistent in regression 4.2, 4.3 and 4.4.

$$Y_{i,t+1} = \beta_0 + \beta_1 X_{i,t} + \epsilon_t \tag{4.2}$$

Simple linear regression with country fixed effects

We then perform the same simple linear regression, including country-fixed effects, represented by dummy variables. Through this inclusion, we account for the potential variations in excess returns that are unique to each country, beyond the influence of the E(SG) scores and environmental metrics. It allows us to control for country-specific factors that may affect excess returns, such as differences in regulatory frameworks, market conditions, economic factors, or cultural factors. The following regression is used to control for country-fixed effects.

$$Y_{i,t+1} = \beta_0 + \beta_1 X_{i,t} + \gamma_2 D2_i + \gamma_3 D3_i + \cdots + \gamma_n Dn_i + \epsilon_t \tag{4.3}$$

Multiple Linear Regression

In this final step, we run a multiple linear regression where all environmental variables (k) are included. The purpose of this final regression is to check whether the coefficients are persistent throughout, in addition to understanding their collective influence on the dependent variable.

$$Y_{i,t+1} = \beta_0 + \beta_1 X_{1i,t} + \beta_2 X_{2i,t} + \cdots + \beta_k X_{ki,t} + \epsilon_t \tag{4.4}$$

4.1.2 How much of the ESG scores are explained by other environmental metrics?

To better understand what E(SG) scores are actually based on, we regress all the environmental metrics on each E(SG) score. We perform a multiple linear regression using equation 4.4, with slight alterations. In this analysis, each E(SG) score is the dependent variable (Y) in time (t+1), i.e. we run one regression per score, and the environmental metrics are independent variables in time (t). In this regression, we will carefully study the R-squared, which will give insight into what portion of the scores is explained by the environmental metrics.

4.2 Cross sectional regression

We perform a cross-sectional regression to further substantiate the findings of the previous linear regressions. The variables are associated with the same time period, meaning we take the average excess return over the whole time period, of all our companies, as our dependent variable (Y), giving us a 1xn vector of 600 average excess returns. We use the average value of the environmental metrics and E(SG) scores over the whole time period, per company, as our independent variable, giving us a 1xn vector of 600 average measures. Additionally, we include the control variables Size and Book-to-Market, which are explained in Appendix A.3.1. Regression 4.5 is run 14 times, one time per environmental variable, giving us a total of 14 coefficients to interpret in relation to the previous regressions.

$$\bar{Y}_i = \beta_0 + \beta_1 \bar{X}_i + \epsilon \tag{4.5}$$

4.3 Computing risk premiums

After exploring E(SG) scores and environmental metrics and their effect on excess return, we want to investigate whether these metrics provide a risk premium. By analyzing potential risk premiums of environmental metrics and scores, we should gain a deeper understanding of how these metrics are truly perceived and used by investors. We compute risk premiums in two ways; using the characteristics-based portfolio sorting method and Fama MacBeth's two-step cross-sectional regressions.

4.3.1 Characteristics-based portfolio sorting method

This method is widely used in modern empirical finance and has been deployed to test theories in asset pricing, construct a wide range of pricing anomalies, and identify investment strategies that are profitable (Soebhag & Vliet, 2022). There are various ways and conditions for how to do the sorting; we will use an 80/20 sorting. We construct portfolios by first sorting the companies based on one of the environmental metrics or E(SG) scores, from the lowest environmental performance (e.g. high emissions or low E(SG) score), hereby referred to as ”brown”, to the best performance (”green”). Each portfolio is rebalanced monthly. In the next step, we subtract the average excess return of the highest 20th percentile of the ranked companies from the average excess return of the lowest 20th percentile. Essentially, we construct the portfolios to go long the worst-performing companies within each metric or score, and go short the best performers. The definition of best or worst performance depends on the type of metric or score we are considering, always categorizing best performance as the most sustainable (e.g. low emissions and high E(SG) score). Lastly, we take the average of all the monthly computed risk premiums and are left with one risk premium for each environmental metric or E(SG) score in sample. With this portfolio construction method, a positive risk premium relates to a higher excess return for ”brown” companies, meaning that investors deem these stocks as riskier and are already expecting a higher return (Bolton & Kacperczyk, 2021) from such companies.

4.3.2 Fama MacBeth

The Fama-MacBeth method is a two-step regression of expected returns as the dependent variable on risk factors for individual stocks, developed by Fama and MacBeth (1973). The model uses the Fama-French Three-Factor Model as a starting point, but extends the model by estimating the factor loadings for each stock in a given time period using cross-sectional regression analysis. This allows for more precise estimates of expected excess return (R) of individual stocks, by estimating the risk premiums of any factors that impact asset prices.

In the first step, we conduct a time-series regression on the factors to estimate the beta coefficient for each factor. The regression assumes constant coefficients and constant expected returns. We estimate the coefficients using ordinary least squares (OLS), and perform the regression for all assets, $i = 1, \ldots, N$, where N represents the number of portfolios used as test assets:

$$R_{i,t} = \alpha_i + \beta_{i,1} F_{1,t} + \beta_{i,2} F_{2,t} + \ldots + \beta_{i,K} F_{K,t} + \varepsilon_{i,t}, \quad t = 1, \ldots, T \tag{4.6}$$

In these regressions, α_i represents the intercept, $\beta_{i,1}, \beta_{i,2}, \ldots, \beta_{i,K}$ are the factor loadings for each of the K factors, $\varepsilon_{i,t}$ is the error term, and T is the number of time step observations. The estimated factor loadings, denoted as $\hat{\beta}_{i,1}, \hat{\beta}_{i,2}, \ldots, \hat{\beta}_{i,K}$, are only approximations of the true factor loadings. We run regression 4.6 once for each environmental metric or E(SG) score, specified as $F_{1,t}$, where every regression includes the control variables described in Table 3.2, specified as $F_{2,t}$, ..., $F_{K,t}$.

The second step of the methodology involves regressing returns on the estimated factor loadings $\hat{\beta}_{i,1}, \hat{\beta}_{i,2}, \ldots, \hat{\beta}_{i,K}$ for each cross-sectional observation. This step yields the estimated risk premium for each of the K factors. The cross-sectional regression equation is as follows:

$$R_{i,t} = \lambda_{0,t} + \lambda_{1,t} \hat{\beta}_{i,1} + \lambda_{2,t} \hat{\beta}_{i,2} + \ldots + \lambda_{K,t} \hat{\beta}_{i,K} + \varepsilon_{i,t}, \quad i = 1, \ldots, N \tag{4.7}$$

Here, $\lambda_{0,t}$ represents the intercept, $\lambda_{1,t}, \lambda_{2,t}, \ldots, \lambda_{K,t}$ are the estimates of the risk premium for the K factors in period t, and $\varepsilon_{i,t}$ is the error term. Through the OLS regressions for each cross-section, T estimates of the risk premium are obtained for each factor. Since they are estimates, they are denoted as $\hat{\lambda}_{1,t}, \hat{\lambda}_{2,t}, \ldots, \hat{\lambda}_{K,t}$. We then calculate the average risk premium using the following formula:

$$\bar{\hat{\lambda}}_k = \frac{1}{T} \sum_{t=1}^{T} \hat{\lambda}_{k,t}, \quad k = 1, \ldots, K \tag{4.8}$$

In this equation, $\bar{\hat{\lambda}}_k$ represents the average risk premium for factor k, and $\hat{\lambda}_{k,t}$ represents the estimate of the risk premium for factor k at time t. T represents the total number of time periods.

4.3.3 Hedging Climate Change News

For the final step of our analysis, we use the same regressions, control variables, and news indices as Engle et al. (2020) to hedge climate change news, but extend the research by including a wider range of E(SG) scores and environmental metrics. In the first step of the mimicking portfolio approach, we create a matrix of firm-level characteristics Z_t, for each E(SG) score and environmental metric, appropriately cross-sectionally normalized to construct the portfolio returns as:

$$\tilde{r}_t = Z'_{t-1} r_t, \tag{4.9}$$

where r_t are excess returns of individual stocks, per month, and portfolio weights are equal to the normalized characteristics. The normalized Z_t for each metric is constructed so that the portfolio is overweight "green" stocks and underweight "brown" stocks, e.g. stocks with high emissions or low E(SG) score.

The second is to measure the portfolios' climate risk exposure. We wish to explore how much of the variation in climate change news, CC_t, is explained by each characteristic-based portfolio, $Z^{\text{ENV'}}$. We use the portfolio returns computed with equation 4.9, for one portfolio at a time, and include the same well-known risk factors as Engle et al. (2020); SIZE, HML, and MKT. See Appendix A.4 for further details on how these risk factors were constructed. w_{SIZE}, w_{HML} and w_{MKT} act as scalars that capture the weights of the corresponding portfolios in the mimicking hedge portfolio for CC_t.

$$CC_t = \xi + w_{ENV} {Z^{\text{ENV}'}}_{1-t} r_t + w_{SIZE} {Z^{\text{SIZE}'}}_{1-t} r_t + w_{HML} {Z^{\text{HML}'}}_{1-t} r_t + w_{MKT} {Z^{\text{MKT}'}}_{1-t} r_t + e_t \tag{4.10}$$

After performing this regression, we test whether the betas (w's) are statistically significant. For all metrics that have statistically significant betas, we explore the in-sample fit of our hedge returns to the climate news index over the full sample period. This brings insight into what type of environmental firm characteristics produce a reliable hedge in periods of high climate news coverage. Furthermore, we will view these findings in the light of our previous analysis on risk premiums, to understand which environmental metrics best capture climate risk exposure.

4.4 Validity

To ensure the robustness and reliability of our findings, we conduct a range of statistical tests and consider potential biases that could influence our results. The significance of our results is assessed using the t-statistic across all regression analyses. Additionally, we address the issue of multicollinearity by calculating the variance inflation factors (VIF) for the E(SG) scores and environmental metrics. Table 4.1 demonstrates that multicollinearity within our sample is generally low, with all variables exhibiting estimated VIF values below 2. We only exclude the environmental metric related to total reported emissions ("CDP MtCO2e") due to its naturally high multicollinearity with Scope 1 and Scope 2 emissions. See Appendix A.2 for a detailed explanation of the validity tests conducted, including the computation of the t-statistic, ADF test, omitted variable bias, selection bias, and reverse causality.

Statistic	N	Mean	St. Dev.	Min	Max
VIF	14	1.238	0.209	1.004	1.620

Table 4.1: **Variance Inflation Factor:** This table shows the results from computing the Variance Inflation Factor (VIF) of our ESG scores and environmental metrics.

Chapter 5

Empirical Findings and Analysis

This section presents the results and findings of the methodology introduced in section 4. We first examine whether excess stock return can be forecasted using environmental metrics. Then, we investigate the link between the E(SG) scores and our sample of measurable environmental metrics, to show how much of the overall E-score is driven by environmental factors. Furthermore, we present the estimated environmental risk premiums, calculated in two ways. Lastly, we present the in-sample fit of our climate news hedge portfolios, constructed with the mimicking portfolio approach.

5.1 Does environmental metrics forecast return?

5.1.1 Linear regression findings

When analyzing our regression findings in Table 5.1, it is evident that some environmental metrics do explain parts of the variation in excess stock return. In our sample, we have four metrics with a consistent and statistically significant effect on excess stock return: Bloomberg ESG score, Sustainalytics ESG score, Refinitiv E-score, and Green Revenue. These metrics are all significant at the 5% significance level, whereas Sustainalytics and Refinitiv have the highest

Variable	SLR	SLR with FE	MLR
(Intercept)	0.0023***	0.0023***	0.0023***
	(0.0004)	(0.0004)	(0.0004)
BB_ESG	-0.0023***	-0.0024***	-0.0023***
	(0.0005)	(0.0007)	(0.0005)
SP_E	-0.0004	-0.0017**	-0.0004
	(0.0005)	(0.0006)	(0.0005)
SA_ESG	-0.0286***	-0.0292***	-0.0286***
	(0.0005)	(0.0005)	(0.0005)
Ref_E	0.0205***	0.0230***	0.0205***
	(0.0010)	(0.0011)	(0.0010)
CDP_S1	-0.0019	-0.0032*	-0.0019
	(0.0012)	(0.0017)	(0.0012)
CDP_S2	0.00004	0.0021	0.00004
	(0.0006)	(0.0017)	(0.0006)
CDP_Reg_Risk	0.0006	0.0075***	0.0006
	(0.0004)	(0.0017)	(0.0004)
CDP_CC_Performance	-0.0009	0.0146***	-0.0009
	(0.0005)	(0.0014)	(0.0005)
CDP_Rep_MtCO2e.y	0.0007	0.0012	0.0007
	(0.0006)	(0.0007)	(0.0006)
Green_rev	0.0034***	0.0037***	0.0034***
	(0.0004)	(0.0004)	(0.0004)
Inv_Op_Sust	-0.0002	-0.0006	-0.0002
	(0.0004)	(0.0007)	(0.0004)
Inv_Sust_Prod	0.0007	0.0014*	0.0007
	(0.0004)	(0.0006)	(0.0004)
Residual SE	0.1164	0.1164	0.1164
R^2	0.02024	0.02024	0.01925
Adjusted R^2	0.02005	0.02005	0.01906
F-statistic	105.7	105.7	100.4
DF (numerator)	17	17	17
DF (denominator)	86981	86981	86982

Table 5.1: **Regression results of environmental metrics on excess return:** This table shows the regression results of the simple linear, simple linear with fixed effects and multiple linear regressions. We can see the coefficients of each of the ESG scores and environmental metrics, with their corresponding T-statistic, with stars indicating significance.

coefficients at approximately -0.029 and 0.021, respectively. Because our environmental metrics are standardized, one standard deviation increase in Sustainalytics E score (indicating an increased level of unmanaged risk, i.e. reduced environmental performance) is associated with a 2.9% decrease in stock excess return. The same interpretation holds true for the Refinitiv E score, as one standard deviation increase in Refinitiv E score (indicating increased environmental performance) is associated with a 2.1% increase in stock excess return.

It is worth noting that among the consistently significant variables, Green Revenue is the only variable not provided by third-party E(SG) rating agencies. Green Revenue represents the percentage of a company's revenues derived from "green" operations. This finding aligns with the research conducted by Bolton and Kacperczyk (2021), which highlights that institutional investors tend to conduct exclusionary climate risk screening in only a few salient industries. As a result, companies that generate a portion of their revenues from renewable resources may experience higher excess returns due to divestment from fossil fuel-intensive industries such as oil and gas.

The presence of ambiguity

Despite some indications that environmental metrics could predict excess return, the regression results are ambiguous. The ESG score from Bloomberg and the E score from S&P Global have opposite coefficient signs than the other third-party E(SG) scores, keeping in mind that Sustainalytics has a reverse scoring system where a high score indicates low environmental performance. Interestingly, where the more "sustainable" companies according to Bloomberg and S&P exhibit a negative relationship with excess stock return, the opposite is true for Refinitiv and Sustainalytics. These findings suggest significantly different outcomes for an impact investor, depending on which scoring system they employ to construct their portfolios. Furthermore, our regression results reveal contrasting coefficient signs for Scope 1 and Scope 2 emissions. This indicates a substantial discrepancy in evaluating the performance of high-emitting companies, depending on which emission metric is utilized to predict excess return. Furthermore, this discrepancy underscores the complexities and nuances involved in analyzing the relationship between environmental metrics and excess stock return.

Need for further investigations

The ambiguity could be explained by several things, and it is evident that we need to investigate further before drawing any conclusions. In light of the findings of Berg et al. (2022), showing that there is a divergence in ESG ratings, the above results call for greater attention to how the data underlying ESG ratings are generated. In addition, there is a risk of selection bias as better-performing companies report more extensively on sustainability than lower-performing companies. To assess whether the shown relationship between a few environmental metrics and stock excess return holds across our sample, we introduce cross-sectional analysis and include firm-specific effects to predict returns.

5.1.2 Cross-sectional regression findings

Running a cross-sectional regression further highlights the ambiguity surrounding the relationship between our environmental variables and excess return. Table 5.2 presents the findings, showing that the only statistically significant variable is the environmental score from Refinitiv, while other variables such as the Bloomberg ESG score, S&P E score, and Green Revenue are no longer found to be significant.

This inconsistency in our regression results may be attributed to several factors. First, the introduction of control variables, namely SIZE and BM, may have influenced the significance of the environmental variables. The control variables account for company-specific characteristics and market valuation, potentially attenuating the impact of environmental metrics and E(SG) scores on excess return.

Second, our previous time-series regressions might not have adequately captured the heterogeneity across companies. For instance, while the time-series regressions showed a statistically significant effect of Green Revenue on excess return, this relationship may be restricted to specific companies rather than being universally applicable across all companies in our sample. This variation in the relationship across different companies could explain the absence of statistical significance in the cross-sectional regression.

Variable	Coefficient	T-stat_ENV	T-stat_SIZE	T-stat_BM	R^2
BB_ESG	-0.00059	-0.8848	-3.7827(***)	-1.6890	0.034
SP_E	0.00035	0.4320	-2.9191(**)	-2.9239(**)	0.029
SA_ESG	0.00025	0.4418	-4.4541(***)	-2.6205(**)	0.044
Ref_E	-0.00415	-2.5534(*)	-4.5186(***)	-2.0355(*)	0.057
CDP_S1	-0.00016	-0.3713	-5.4175(***)	-2.7039(**)	0.071
CDP_S2	0.00021	0.7378	-6.6311(***)	-2.5096(*)	0.120
CDP_CC_Perf.	0.00043	1.4272	-5.8098(***)	-2.1949(*)	0.099
Green_rev	-0.00072	-0.9044	-1.9604	-1.3015	0.213
Inv_Op_Sust	0.00043	0.6857	-2.9324(**)	-0.5522	0.079
Inv_Sust_Prod	-0.00111	-0.5081	-0.9516	-1.2341	0.339
RES_Use	0.00009	0.1459	-3.0631(**)	-1.8783	0.030
Tot_Energy_Cons	-0.00016	-0.2206	-3.0729(**)	-1.7056	0.023
Raw_Material_Used	0.00087	0.6676	-0.6496	0.1708	0.011
Tot_Waste	0.00027	0.4708	-3.3590(***)	-2.3642(*)	0.039
CO2intensity	-0.00011	-0.3264	-5.4781(***)	-2.8325(**)	0.072

Table 5.2: **Cross-sectional regression results:** The table shows the regression findings from the cross-sectional regression, with the coefficients for each ESG score and environmental metric in the first column. The preceding columns contain the T-statistic of the score or metric, the size variable and the BM variable, and lastly the R-squared.

Overall, the results from the cross-sectional regression amplify previous conclusions that the findings are not coherent, and that more robust regressions are needed to outline a true relationship between stock excess return and environmental performance.

5.2 How much of the ESG scores are explained by other environmental metrics?

To further understand the dynamic between third-party E(SG) scores (both ESG scores or E scores) and the "non-greenwashable" environmental metrics, we will investigate what explains the E(SG) scores. After running one multiple linear regression of all the environmental metrics on the five E(SG) scores, we are surprised to see that the environmental metrics in our sample explain such a small proportion of the E(SG) scores. Every E(SG) score has an R-squared below 4%, except for the CDP Climate Change Performance score that has an R-squared at 11.3% (see Table 5.3. for the complete results). The fact that the ESG scores are not explained to a high degree by CO2 emissions, total waste, investments in sustainable products or other environmental metrics could be explained by a low score weighting to the "E" (environmental pillar) in the ESG score. However, for our three environmental scores, Refinitiv, S&P Global and CDP Climate Change Performance, the findings are surprising.

Disagreement between agencies' claimed and actual measures

Refinitiv is the third-party score with the lowest R-squared, with only two statistically significant variables in our sample, namely "Green Revenues" and "Regulatory Risk Exposure", both at the 0.1% level. According to Refinitiv, their Environmental Pillar score is based on three categories; resource use, emissions, and innovation. It is therefore remarkable that this score does not display a significant relationship with neither emissions (scope 1 or 2) nor resource use metrics such as total energy consumption or total waste. This observation suggests a discrepancy between what the rating agencies say they measure and what they actually measure. Furthermore, it is interesting to observe that no E scores in our sample show a statistically significant relationship with CO2 intensity or emissions.

Metric	Sustainalytics ESG	Bloomberg ESG	Refinitiv E	S&P E	CDP CC
CDP_S1	0.02	-0.04	0.00	0.00	-0.04
	(3.36)**	(-9.63)***	(-0.21)	(-0.03)	(-8.88)**
CDP_S2	-0.02	0.07	-0.01	0.00	0.31
	(-4.34)***	(15.26)***	(-1.87)	(-0.30)	(69.14)***
CDP_Reg_Risk	-0.01	0.02	0.02	0.02	0.04
	(-1.52)	(7.20)***	(4.63)***	(7.04)	(12.59)**
CDP_Rep_MtCO2e	-0.01	0.03	0.00	0.00	0.01
	(-1.38)	(7.59)***	(-0.30)	(-0.25)	(2.85)
Green_rev	-0.10	0.06	0.04	0.05	0.02
	(-30.54)***	(16.47)***	(11.89)***	(14.98)	(7.37)
Inv_Op_Sust	-0.03	0.06	0.00	0.07	0.08
	(-9.92)***	(16.25)***	(-1.24)	(19.15)	(25.14)***
Inv_Sust_Prod	-0.04	0.06	0.00	0.03	0.03
	(-12.68)***	(16.35)***	(-0.58)	(8.85)	(9.47)
RES_Use	-0.04	0.03	0.00	0.05	0.09
	(-10.24)***	(7.99)***	(1.21)	(12.32)***	(23.59)***
Tot_Energy_Cons	0.03	0.02	0.00	-0.03	0.03
	(6.17)***	(4.65)***	(-0.30)	(-6.03)***	(7.26)***
Raw_Material_Used	-0.03	0.08	0.00	-0.01	-0.02
	(-6.70)***	(22.03)***	(1.09)	(-2.76)	(-4.88)***
Tot_Waste	0.01	-0.02	0.00	-0.02	-0.12
	(2.43)*	(-4.11)***	(0.14)	(-4.76)***	(-32.29)***
CO2intensity	0.01	-0.01	0.00	0.02	0.00
	(1.51)	(-2.80)**	(-0.34)	(6.59)	(-0.94)
R^2	0.01675	0.03387	0.00199	0.0113	0.1129
Adjusted R^2	0.01661	0.03374	0.00186	0.01116	0.1128
F-statistic	123.5	254.1	14.52	82.86	922.3
DF	12 and 86987	12 and 86987	12 and 86987	12 and 86987	12 and 86987

Table 5.3: **Regression findings of environmental metrics on E(SG) scores:** This table shows the regression results from the multiple linear regression studying the effects of environmental metrics on E(SG) scores, with corresponding T-statistics. The R-squared of these regressions on the bottom of the table tells us how much of the E(SG) score is explained by the environmental metrics.

In contrast to Refinitiv, the remaining four E(SG) scores demonstrate statistically significant relationships with most of our measurable environmental metrics, particularly the Bloomberg ESG score, which exhibits significant relationship with all metrics. This finding is more closely aligned with both our expectations and with what ESG raters communicate.

Not only is there a disagreement between the various ESG rating agencies, but it also seems like the ratings carry important unintended exposures. According to research by LaBella et al. (2019), two of the most notable unintended exposures are in company size and geography. Size exposure occurs as rating agencies rely on survey and policy disclosure data, leading to a consistent skew favouring large and multi-national companies, while geographical exposure happens when some jurisdictions have a higher quality of formal reporting requirements. For an investor, the low R-squared values indicate that relying solely on E(SG) scores may not adequately capture a company's exposure to climate risk. Investors might consider incorporating additional metrics, including specific climate-related metrics such as emissions, energy consumption, and waste management, to gain a more accurate assessment of a company's climate risk profile. However, if most climate risk-averse investors select stocks using E(SG) scores as their proxy for climate risk, these scores might be more related to excess stock return than other "non-greenwashable" metrics like emissions. Further investigation into this relationship is therefore needed.

5.2.1 Delta

When assessing the twelve-month change (delta) of the environmental variables, the story is slightly different. Among the E(SG) scores, Refinitiv E-score exhibits the highest R-squared value at 5.8%, while the remaining E(SG) scores in our sample are all below 3%. Although the low R-squared measure makes it difficult to conclude, this finding suggests that Refinitiv E-score might be more effective

Model	$\mathbf{R^2}$	$\mathbf{R^2}$ **Delta**
$\text{Sustainalytics}_{\text{ESG}}$	0.01675	0.01914
$\text{Bloomberg}_{\text{ESG}}$	0.03387	0.00499
$\text{Refinitiv}_{\text{E}}$	0.00199	0.05869
$\text{S\&P}_{\text{E}}$	0.01130	0.00302
$\text{CDP}_{\text{Climate Change Performance}}$	0.11287	0.02632

Table 5.4: **R-squared measures:** This table shows the R-squared measures from the various multiple linear time-series regressions, where environmental metrics are regressed on the various E(SG) scores. The table compares the R-squared measures of regressing the absolute values of the environmental metrics and regressing the 12 months change in absolute values (delta).

in capturing fluctuations in environmental performance rather than solely focusing on absolute performance levels. This implication is important for investors seeking to hedge climate change risk. It suggests that companies with higher emissions, for instance, will achieve a high environmental score from Refinitiv if they demonstrate a positive trend of reducing their emissions. Overall, we observe that some E(SG) rating agencies rely more heavily on changes (delta) in environmental performance (e.g. emission reduction), while others give more weight to absolute performance (e.g. total emissions). Investors need to be mindful of this disparity between different rating agencies when using E(SG) scores to construct portfolios and manage climate risk.

5.3 Calculation of risk premiums

The main analysis is to investigate whether there is a risk premium associated with environmental performance. Specifically, which E(SG) scores and environmental metrics represent a material risk for investors that is reflected in the cross-section of stock returns and portfolio holdings. We have calculated risk premiums using two methods; characteristic-based portfolio sorting and Fama MacBeth two-step regressions.

5.3.1 Characteristic-based portfolio sorting

Based on our characteristic-based portfolio sorting method, our empirical findings reveal the presence of a climate risk premium at the firm level for multiple environmental metrics. As shown in Table 5.5, constructing portfolios by shorting assets with favourable environmental characteristics (i.e. "green" companies) and taking long positions in assets with unfavourable environmental characteristics ("brown") yields statistically significant risk premiums for the following variables: Bloomberg ESG score, CDP Climate Change Performance score, total energy consumption, total waste, and both scope 1 and scope 2 emissions. The risk premiums associated with portfolios built on these six sorting characteristics are statistically significant at a 5% level, while total energy consumption and scope 2 emissions exhibit significance at a more stringent 1% level.

However, this climate risk premium is not persistent across all statistically significant environmental metrics, contributing to the continued ambiguity in our findings. There is a disagreement between the signs of the climate risk premiums, where some metrics have negative risk premiums and others have positive ones. The interpretation of our risk premiums will therefore depend on what E(SG) scores and environmental metrics are used to construct the portfolios.

Constructing portfolios using Bloomberg ESG scores and CDP Climate Change Performance scores results in positive risk premiums. Portfolios investing in companies with low scores, indicating a higher climate risk exposure, and short-selling companies with high scores, obtain a stock excess return of 5.22% and 4.63%, respectively. This finding is aligned with common risk premium theory, stating that investors demand compensation for bearing the additional climate-related risks associated with these companies. On the other hand, a contrasting pattern emerges when we construct portfolios based on "non-greenwashable" metrics such as total waste, total energy consumption, and scope 1 and 2 emissions. For companies characterized by high levels of energy consumption, waste generation, or emissions, we identify a significant negative climate risk premium of -3.15%, -2.38%, -2.52% and -3.20%, respectively. In other words, these high-emitting (i.e. "brown") companies underperform and exhibit lower stock excess returns over time. The presence of a negative risk premium suggests that investors penalize companies with substantial environmental footprints, possibly due to concerns about their long-term sustainability.

The presence of opposing risk premiums, based on the construction of portfolios using either third-party E(SG) scores or reported environmental metrics, highlights an intriguing ambiguity in our findings. It raises the question of whether there are two distinct types of investors, each employing different strategies to reduce portfolio climate risk. One group of investors seeks to mitigate climate risk by incorporating companies with high E(SG) scores, accepting a slight reduction in return. Conversely, another group of investors aims to reduce climate risk by investing in companies with lower actual emissions or consumption levels. These investors prioritize tangible environmental metrics when constructing their portfolios.

The coexistence of these opposing strategies is particularly interesting, given our previous findings that showed a weak to non-existing relationship between E(SG) scores and measurable environmental metrics. If investors who seek to reduce climate risk exposure adopt these diametrically opposite approaches, the observed ambiguity in our results becomes more understandable. This dual investor perspective implies that effectively reducing portfolio climate risk is a complex endeavor, and necessitates thorough research and understanding of the specific metrics employed to construct a portfolio. Investors must carefully evaluate whether they prioritize third-party E(SG) scores, which may reflect broader sustainability considerations but exhibit weak connections to environmental performance, or focus on direct measures to mitigate climate risk.

5.3.2 Fama MacBeth

To investigate the hypothesis of dual-investor perspectives and delve deeper into the implications of our findings, we employ the Fama-MacBeth approach, a widely recognized method to construct risk premiums. The results from the Fama-MacBeth two-step regression are presented in Table 5.4. These regressions are conducted separately for each environmental metric, with the inclusion of the same set of control variables as in Table 3.1.

Controlling for other known risk factors and firm characteristics (Bolton & Kacperczyk, 2021) will let us fully understand how financial markets price climate risk. This is important given that risk premiums calculated using the characteristics-based portfolio sorting method are at risk of having bias in the environmental metrics and E(SG) scores. Two potential biases are company size or geography bias, e.g. higher emissions for larger companies, or higher E(SG) scores for companies in jurisdictions with high quality of reporting requirements. By including 12 well-known control variables (Table 3.1.) we aim at increasing the robustness of our Fama MacBeth cross-sectional regression findings.

When analyzing the regression results, we observe that there are environmental risk premiums significantly different from zero. Specifically, the metrics that are statistically significant are the Bloomberg ESG score, S&P Global E score, Refinitiv E score, CDP Climate Change Performance score, investments in sustainable products, and renewable energy usage. Out of these six variables, the ones with positive risk premiums are still solely third-party E(SG) scores, namely the Bloomberg ESG score and Refinitiv E score. The two yield statistically significant risk premiums at the 0.1% and 1% levels, respectively. On the other hand, we continue to observe negative risk premiums as well. As depicted in Table 5.4, S&P E score, CDP Climate Change Performance Score, sustainable investments, and renewable energy usage all exhibit negative risk premiums with statistically significant deviations from zero at varying significance levels (0.1%, 5%, 0.1%, and

Metric	Sorting RP	T-stat	Fama MacBeth RP	T-stat
BB_ESG	5.22%	2.39(∗)	15.86	26.50(∗∗∗)
SP_E	-2.15%	-1.38	-73.69	-8.04(∗∗∗)
SA_ESG	-8.65%	-1.57	NA	NA
Ref_E	-5.95%	-1.83	105.26	2.69(∗∗)
CDP_CC_Perf.	4.63%	2.26(∗)	-4.42	-2.21(∗)
Green_rev	-8.20%	-0.49	NA	NA
Inv_Op_Sust	3.75%	1.36	-98.29	-16.89(∗∗∗)
Inv_Sust_Prod	1.98%	0.35	-109.05	-1.95
RES_Use	3.70%	1.70	-1.96	-5.22(∗∗∗)
Tot_Energy_Cons	-7.83%	-3.15(∗∗)	-0.79	-1.01
Raw_Material_Used	-6.73%	-1.92	0.79	0.37
Tot_Waste	-4.68%	-2.38(∗)	-0.20	-0.27
CO2intensity	-2.55%	-1.12	-0.00	-0.45
CDP_S1	-5.24%	-2.52(∗)	0.01	1.57
CDP_S2	-6.58%	-3.20(∗∗)	-0.01	-1.21

Table 5.5: **Risk premium calculations:** This table shows the risk premiums computed for the E(SG) scores and environmental metrics using both the characteristics-based sorting method (column 1) and the cross-sectional regression method of Fama MacBeth (column 3), both with corresponding T-statistics. Column 3 (Fama MacBeth RP) displays the average estimated risk premiums for each of the K factors, i.e. $\bar{\hat{\lambda}}_k$

0.1%, respectively).

The hypothesis that there is a different pattern for portfolios sorted on E(SG) scores and those sorted on measurable environmental metrics does not hold true for our Fama MacBeth cross-sectional regressions. We observe a noteworthy deviation in the risk premiums of the two E(SG) scores from Refinitiv and S&P Global. Interestingly, these two E(SG) scores display negative risk premiums, aligning them with the other ”non-greenwashable” metrics rather than the remaining E(SG) scores. In light of the findings of Pástor et al. (2021), this discrepancy might be explained by investor preference, as they found that investors with stronger ESG preferences hold portfolios with a green tilt, while those with weaker preferences take a brown tilt.

Our findings challenge existing research and contradict previous findings in US data samples, leaving the relationship between climate risk, environmental performance, and stock returns unclear. Despite identifying statistically significant risk premiums, the implications are not straightforward. When analyzing the risk premiums, we note two segments: those based on environmental metrics and those based on third-party E(SG) scores. Environmental metrics consistently result in negative risk premiums, while E(SG) scores provide both positive and negative risk premiums, revealing the ambiguity and unreliability of third-party scores. Knowing that ESG rating divergence is very pronounced (Berg et al., 2022) and that the scores are explained to a low degree by measurable environmental metrics, using E(SG) ratings in portfolio construction does not seem like a robust way of pricing climate risk.

5.4 Hedging Climate Change News

The final test to fully understand the relationship between climate risk and portfolio return is to implement the mimicking portfolio approach to dynamically hedge climate change risk in our sample. Our findings demonstrate that it is possible to create concise and well-diversified investment portfolios. However, none of these portfolios performs well in hedging in-sample innovations in climate news.

First, we explore the in-sample fit of each version of regression 4.10 over the full sample period. Table 5.6 shows the regression results when hedging innovations to the WSJ Climate Change News Index, ${CC_{\mathrm{t}}}^{\mathrm{WSJ}}$, on each portfolio sorted by our E(SG) scores and environmental metrics. There is a clear and statistically significant relationship between the portfolios sorted by characteristics like Bloomberg ESG scores, CDP Scope 2 emissions, raw material usage, and total waste, and ${CC_{\mathrm{t}}}^{\mathrm{WSJ}}$, at significance levels 10%, 5%, 5% and 5%, respectively. In other words, during periods characterized by a higher frequency of innovations in climate-related news, portfolios favouring "green" firms (i.e., long companies with low emissions/material usage or high E(SG) scores) generate relatively higher excess returns. These findings are especially noteworthy in light of our prior research, which established that constructing portfolios based on the same variables provides investors with a risk premium of +/- 4.5-7%. In addition to this, we now observe that investors appear to adjust their portfolio allocations based on climate change news coverage, likely because news coverage increases their climate risk aversion through increased awareness. Considering the expected escalation of global warming in the upcoming years and the subsequent surge in climate news reporting, it is intriguing to observe the direct influence of such developments on investor behaviour, prompting a shift towards greener assets and reducing exposure to climate-related risks.

When examining the R-squared measures of these regressions, it becomes evident that the portfolios based on the E(SG) scores and environmental metrics in our sample demonstrate limited effectiveness in hedging the in-sample variation in CC_{t}. Most R-squared measures are below 1%, with the highest value observed in the non-significant metric Inv_Sust_Prod, exhibiting an R-squared of 1.7%. This finding is unexpected, considering that Engle et al. (2020) reported that the

Proxy	(1)	(2)	(3)	(4)	(5)	(6)
BB_ESG	0.0003*					
	(0.0001)					
CDP		0.0001				
		(0.0001)				
CDP_S1			0.0001			
			(0.0001)			
CDP_S2				0.0003**		
				(0.0001)		
CO2i					0.0004	
					(0.0006)	
Inv_Op_Sust						-0.0001
						(0.0004)
ZMKT	0.0721	0.1127***	0.1149***	0.1157***	0.1177***	0.1467***
	(0.0557)	(0.0180)	(0.0179)	(0.0179)	(0.0178)	(0.0404)
ZBM	0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
ZSIZE	0.0000	0.0001*	0.0000*	0.0000*	0.0000*	0.0001*
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Additional Regression Information						
R-squared	0.0005997	0.001575	0.001681	0.001931	0.001701	0.002143
N	13583	27757	27044	27604	27675	6953

Proxy	(7)	(8)	(9)	(10)	(11)	(12)
Raw_mat	0.0005**					
	(0.0002)					
RES_Use		0.0000				
		(0.0002)				
Ref_E			0.0000			
			(0.0003)			
SP_E				0.0002		
				(0.0002)		
Tot_Energy					0.0001	
					(0.0001)	
Tot_Waste						0.0003**
						(0.0001)
ZMKT	0.0796*	0.0935***	-0.0051	-0.0077	0.1177***	0.1091***
	(0.0375)	(0.0252)	(0.0541)	(0.0567)	(0.0178)	(0.0194)
ZBM	0.0000	-0.0000	-0.0000	0.0000	0.0000	-0.0000
	(0.0001)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
ZSIZE	0.0000	0.0000	-0.0000	0.0000	0.0000*	0.0000*
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Additional Regression Information						
R-squared	0.002944	0.001262	0.002464	0.0002352	0.001527	0.001973
N	5132	11364	6358	5910	5132	21656

Table 5.6: **Full-sample regression of WSJ Climate Change News Index:** This table shows results from regression 4.10. The dependent variable captures innovations for the WSJ-Based Climate News measure. The unit of observation is one month, and the sample runs between December 2010 and June 2017. Standard errors are presented in parentheses. $*p < .1$; $**p < .05$; $* * *p < .01$. The environmental metrics in sample not present are excluded because of limited data to perform the hedge.

portfolios based on the Sustainalytics E-Scores can hedge 15%–19% of the in-sample variation in CC_t. One plausible explanation for this disparity could be attributed to the fact that Engle et al. (2020)'s study focused on a sample of US assets, while our analysis exclusively encompasses European stocks. Both studies rely on a news index constructed on daily news reports from the Wall Street Journal, an American news paper, which suggests the possibility that our shared assumption of climate change news being universally global might not hold true.

In addition to the E(SG) scores and environmental metrics, we see from Table 5.6 that market value (MKT) appears to correlate with climate change exposure. Contrary to the findings of Engle et al. (2020), we observe that larger firms with higher market capitalization display a lower vulnerability to climate change news compared to smaller firms, as they perform better when there is an increase in climate change news coverage in the Wall Street Journal (WSJ). However, in light of our previous finding that investors base their investments on environmental characteristics, it is plausible that larger firms allocate more resources towards sustainability reporting and ensuring a positive perception of their environmental practices. As a result, it may not be solely the size of the firms that renders them less vulnerable to climate change news, but rather that investors prefer companies with high-quality sustainability reporting during times of heightened perceived climate risk. More research is, however, needed to fully understand this relationship.

Figure 5.1 presents the in-sample performance of portfolios constructed to hedge innovations in the WSJ Climate Change News Index. Each panel show portfolios constructed using absolute values of the statistically significant E(SG) scores and environmental metrics in the first step of the mimicking portfolio approach. The left panels present scatter-plots of the in-sample returns of the hedge portfolios together with the realizations of the innovation of climate news. The right panels plot the time series of the climate news series and the return series

	CC_t	H_{BB}	H_{CDP}	H_{S1}	H_{S2}	H_{CO2i}	H_{InvOp}	H_{RawM}	H_{Ref}	H_{RES}	H_{Energy}	H_{Waste}
CC	1.00											
H_{BB}	-0.34	1.00										
H_{CDP}	-0.04	-0.04	1.00									
H_{S1}	-0.05	-0.05	1.00	1.00								
H_{S2}	-0.04	-0.04	1.00	1.00	1.00							
H_{CO2i}	-0.04	-0.04	1.00	1.00	1.00	1.00						
H_{InvOp}	0.01	0.01	-0.02	-0.02	-0.02	0.97	1.00					
H_{RawM}	-0.03	-0.03	0.96	0.97	0.96	0.96	0.96	1.00				
H_{Ref}	0.16	0.16	0.04	0.04	0.04	0.09	0.06	1.00	0.05			
H_{RES}	-0.02	-0.02	0.98	0.98	0.98	0.98	0.98	0.94	0.05	1.00		
H_{Energy}	-0.02	-0.02	1.00	0.99	1.00	1.00	0.98	0.97	0.05	0.99	1.00	
H_{Waste}	-0.02	-0.02	1.00	1.00	1.00	1.00	0.98	0.97	0.06	0.99	1.00	1.00

Table 5.7: **Cross-correlations of portfolio hedge returns and WSJ Climate Change News Index:** This table shows cross-correlations of different portfolios and innovations in the WSJ Climate Change News Index. The table focuses on the performance of hedge portfolios from our in-sample approach.

of the hedge portfolios. There is a negative in-sample correlation for all the four hedge portfolios, indicating that an investor can hedge in-sample variation of climate risk exposure (CC_t) using these metrics. Overall, the in-sample correlation between realization of climate change news and the hedge portfolios are -0.34 when using Bloomberg ESG scores, -0.04 when using CDP Scope 2 emissions, -0.03 when using raw material usage, and -0.02 when using total waste. The hedging ability of the Bloomberg ESG score is in this case much higher than for the other environmental characteristics, suggesting that the Bloomberg ESG score is more suited to capture climate change news. The substantially lower hedging performance of the other three portfolios highlight the importance of choosing characteristics that properly capture cross-sectional variation in exposure to climate risk. This is further exemplified in column 1 in Table 5.7, where we see that the hedging return of the portfolio constructed using Bloomberg ESG characteristics has a substantially higher correlation with CC_t than the hedging returns when using the other environmental characteristics.

Our findings reveal that an investor can construct portfolios with a relatively high correlation between in-sample hedged portfolio returns and climate news series (-34% for Bloomberg ESG). However, the results also highlight two noteworthy limitations. First, the ambiguity between our E(SG) scores and environ-

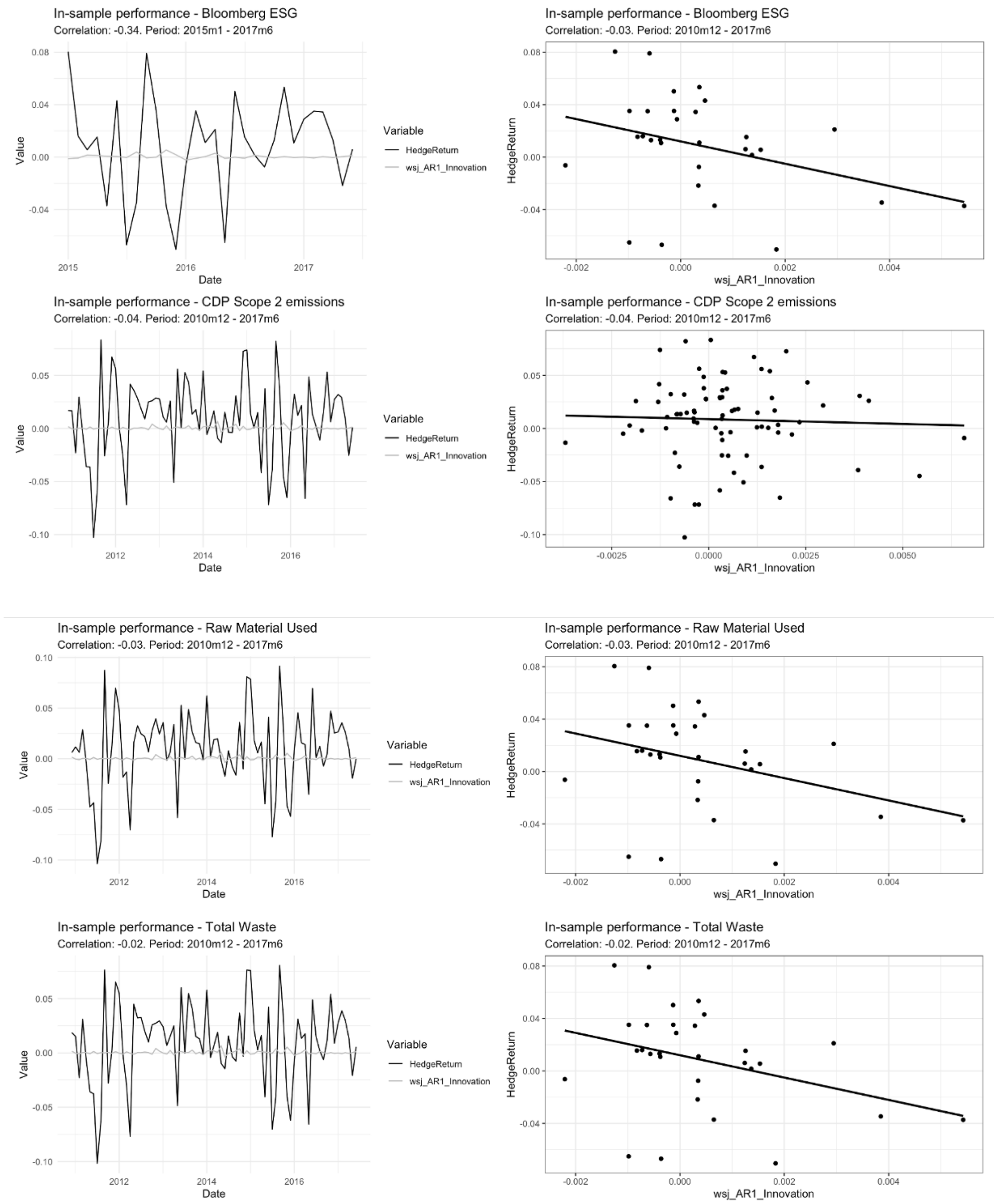


Figure 5.1: **In-sample fit: WSJ Climate Change News Index:** This figure explores the in-sample performance of the hedge portfolios constructed to hedge the WSJ-Based Climate News Measure, using the four statistically significant environmental measures in our sample. The top panel presents hedge portfolios built on the absolute values of the Bloomberg ESG score, then the CDP Scope 2 emissions, Raw Material Usage, and lastly Total Waste Generation.

mental metrics persists. Not all portfolios sorted by climate exposure (i.e. long "green" assets and short "brown" assets) exhibit the same relationship between hedge returns and innovations in climate change news. Second, the R-squared value of our European sample is considerably lower compared to that of Engle et al. (2020) in their US sample, indicating the potential influence of sample-specific characteristics on the obtained hedging results. Both the persisting ambiguity and the hypothesis of sample-specific results present intriguing avenues for future research.

5.5 Limitations and future research

The essence of our results is subject to discussion, and potential criticisms could be aimed at the fundamentals of our approach. Researchers have raised concerns about the empirical validity of risk factors, with Lo and MacKinlay (1990) categorizing the pursuit of risk factors as mere data mining. Other researchers like Ferson and Harvey (1991) do, however, argue that redundant factors can still possess explanatory power. Furthermore, Lakonishok et al. (1994) propose that factor premiums may stem from irrational investor behaviour rather than compensation for systematic risk. This irrationality becomes particularly relevant in the context of climate risk, as concerns over global warming and carbon emissions from human activity have only recently gained significant attention. Moreover, the intricate relationship between corporate environmental performance and financial performance remains uncertain, leaving investors seeking to hedge climate risk in a difficult position.

Another limitation of this study concerns the use of environmental data. Despite sustainability reporting having gained massive momentum, the data itself is prone to inaccuracies and discrepancies. Specifically, when it comes to ESG data, variations exist in the nature and scope of reporting practices across different companies, diminishing comparability. Consequently, the findings presented in this study may be constrained by the quality and accessibility of the employed data. It is important to acknowledge these limitations when interpreting the results and drawing conclusions.

We acknowledge that further investigation on this topic is needed, specifically in ESG divergence on stock returns. The research on this subject is still limited in the European market, and in the coming years we expect that more companies will report on climate-related metrics and obtain ESG scores. As highlighted by Engle et al. (2020), future research should also focus on distinguishing between physical and transitional risk, as these risks can have different, if not opposing implications for investors, e.g. an investor in real estate is more prone to physical climate risk than transitional. Furthermore, we propose to distinguish between sectors, as there may be large industry-specific patterns of divestment, such as in oil and gas or other fossil fuel reliant industries. This idea is supported by the research of Bolton and Kacperczyk, 2021, who found that large institutional investors divest from carbon-intensive industries in order to adhere to new negative exclusionary screening investment strategies.

Lastly, as we have established that some environmental metrics in our sample have an effect on excess stock return, we propose a further investigation into this relationship. More specifically, we suggest studying the cash flow effect on corporate sustainability, to better understand whether investments in mitigating policies or adopting abatement technologies can yield an increase in stock excess returns. Such analysis will provide valuable insights into the financial implications of concrete sustainability measures for companies and their potential influence on market performance.

Chapter 6

Conclusion

Our research addresses the longstanding challenge of convincing investors about the financial legitimacy of climate risk. Historically, many investors have held the belief that actively managing climate risk would likely result in diminished investment returns. However, in light of the pressing threat of climate change, our study offers a deeper understanding of market rationale on sustainable investment.

This thesis investigates whether a broad span of environmental metrics, both measurable and non-measurable, affect the cross-section of European stock excess returns. Our findings reveal a compelling insight; adopting an investment strategy focused on long environmental underperforming firms (i.e. "brown") and short overperforming firms ("green") can yield opposing risk premiums.

Previous research has shown that investing in carbon-efficient companies can be financially rewarding, even without the presence of government incentives (In et al., 2019). Our study contradicts this to some degree, showing that it depends on what measures are used to proxy climate risk exposure. Using third-party E(SG) scores like the Bloomberg ESG score or CDP Climate Change Performance score, an investor can earn a positive and significant risk premium of 5.22% and 4.63%, respectively. On the other hand, by creating portfolios based on "non-greenwashable" metrics like total energy consumption, total waste, scope 1 or scope 2 emissions, an investor would earn a reduced excess return of -3.15%,

-2.38%, -2.52% and -3.20%, respectively. When analyzing the diverging risk premiums, we observe two segments: Risk premiums created using measurable metrics and those using third-party E(SG) scores. For the first segment, we get solely negative risk premiums, suggesting that investors penalize companies with substantial environmental footprints, possibly due to long-term sustainability concerns. For the second segment, the interpretation is difficult. We obtain statistically significant risk premiums with diverging signs, highlighting the ambiguity and low reliability of third-party scores. This is further exemplified through our E(SG) score analysis, showing that the environmental metrics in our sample explain a small portion of the E(SG) scores (sub 4% R-squared). This disagreement weakens the reliability of scores when sorting companies on climate risk exposure.

Lastly, we showed that no environmental variable sufficiently work when hedging in-sample variation in climate news, as all the significant metrics exhibited a low R-squared between our hedged portfolio returns and the Climate News Index. However, we did observe that investors adjust portfolio allocations in times with high reporting on negative climate change news in favour of more "green" assets.

Out of all the variables in our sample, the Bloomberg ESG score is the most reliable metric, displaying statistically significant risk premiums in both the sorting method (5.22%) and the Fama MacBeth method (15.86). In addition, BB_ESG exhibits by far the highest correlation (34%) between its portfolio hedge returns and innovations in climate news, establishing it as the preferred variable when considering climate risk in a European portfolio.

To conclude, our findings challenge existing research and contradict previous findings in US data samples, leaving the relationship between climate risk, environmental performance, and stock returns unclear. Despite identifying statistically significant risk premiums, the implications are not straightforward, and further research on the intersection between financial performance and climate risk exposure would be beneficial.

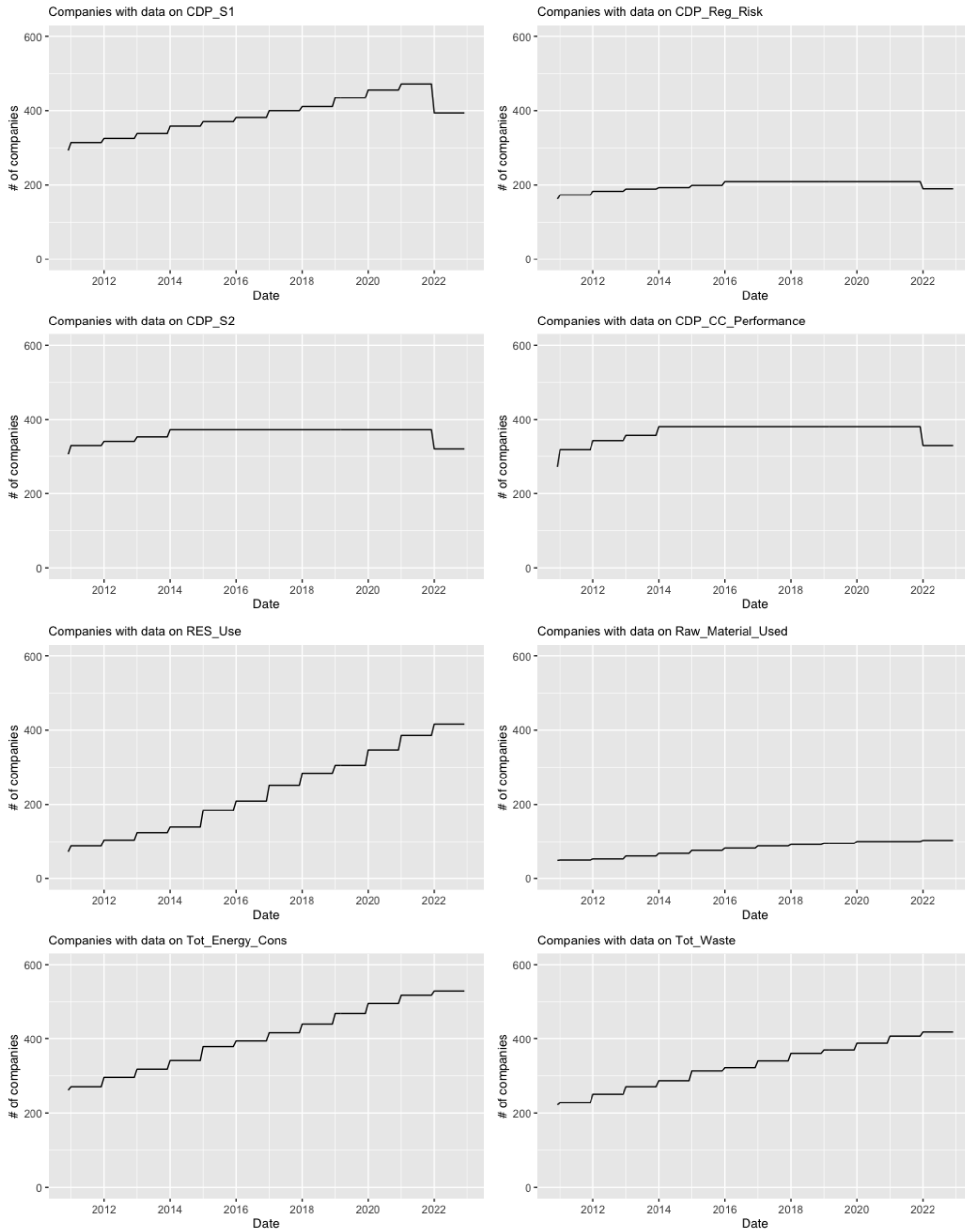
Appendix A

Appendix

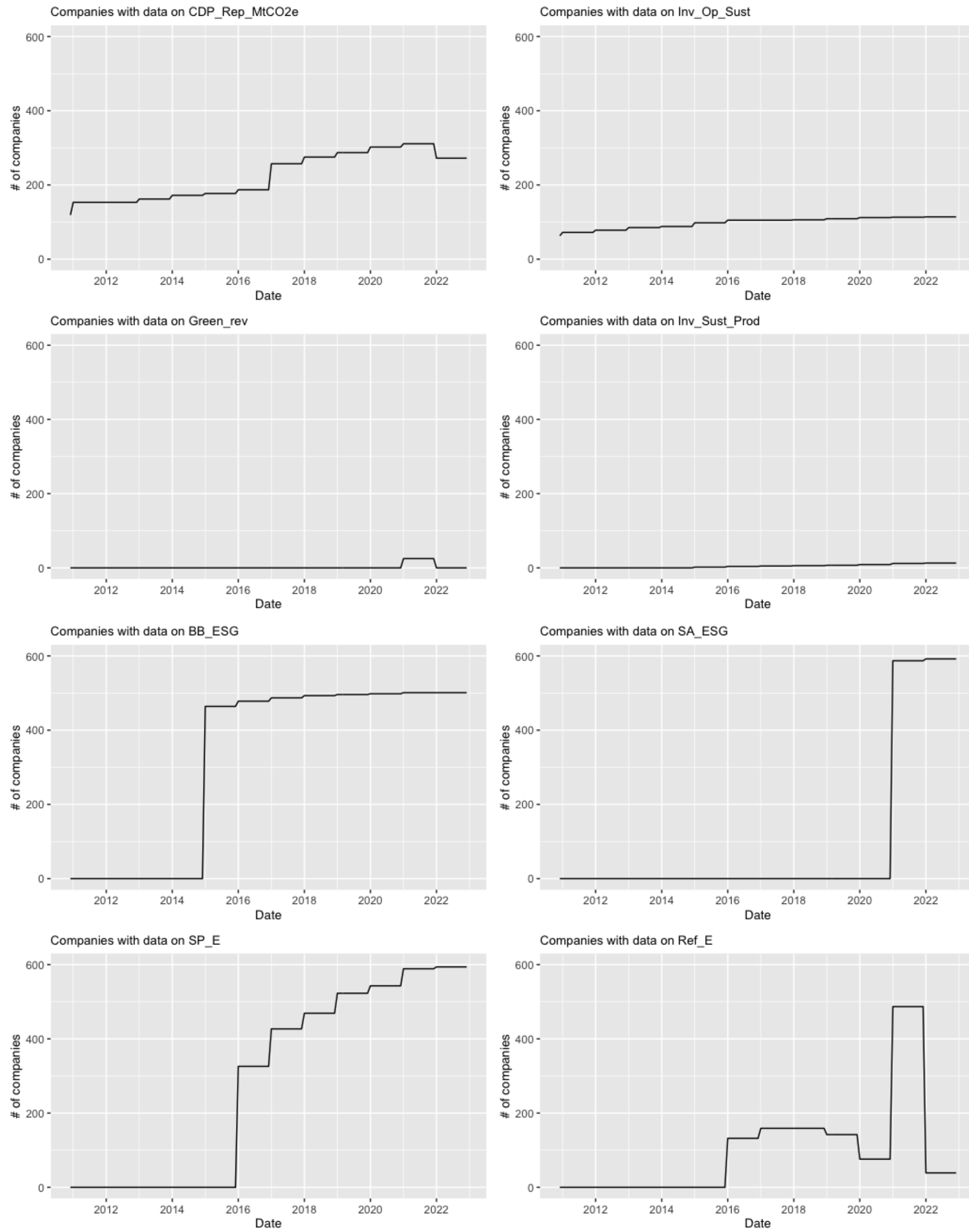
A.1 Summary Statistics Detailed Visualization

A.1.1 Number of data points per variable

The charts below present the number of firms, out of our total sample of 600 firms, that have data on each environmental metric and E(SG) score in our sample. As the data is a time series, the charts also present how this number evolves over the time period analyzed. For most variables, an upward trend is observed, indicating a growing number of companies reporting on climate-related metrics and receiving E(SG) scores.

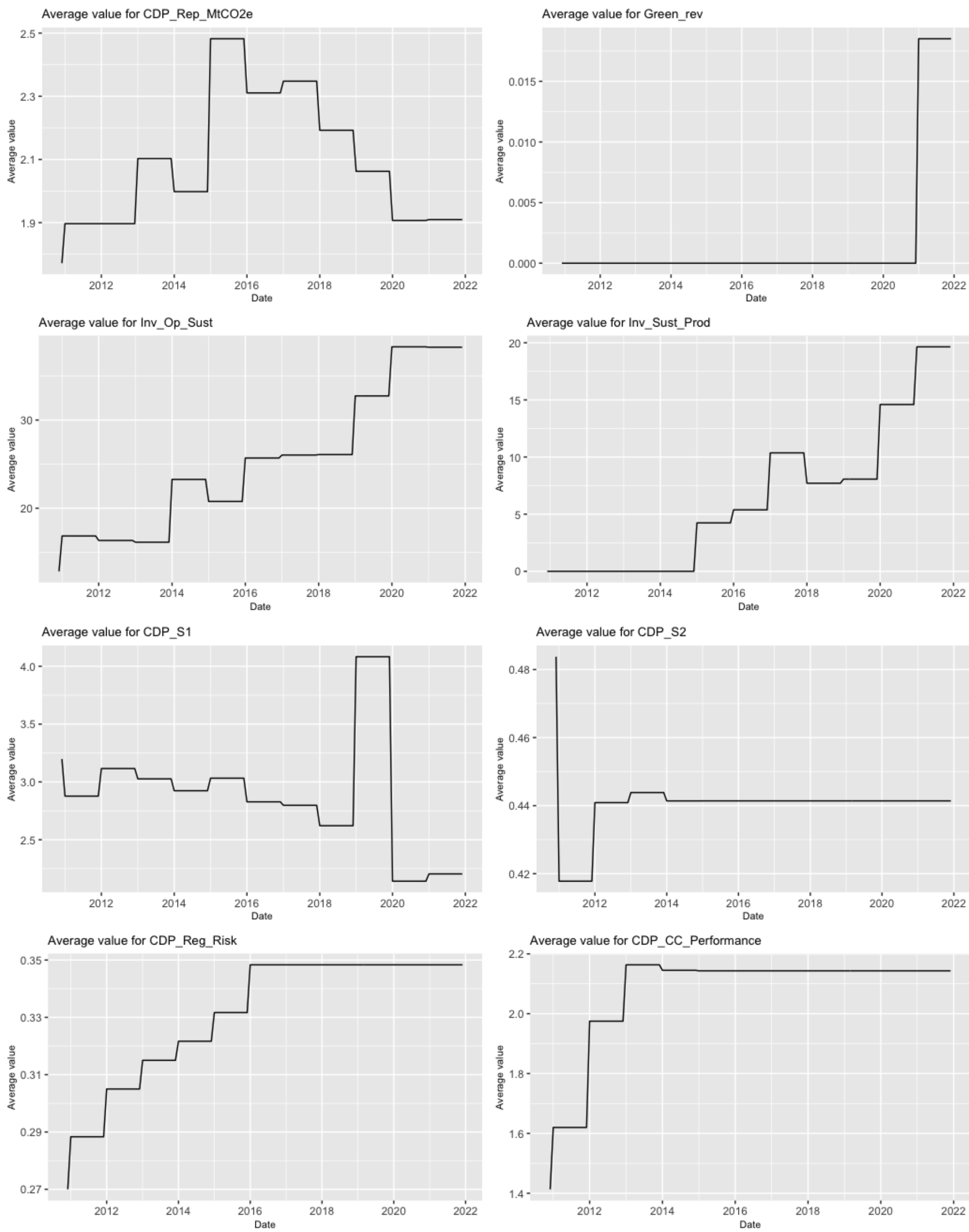


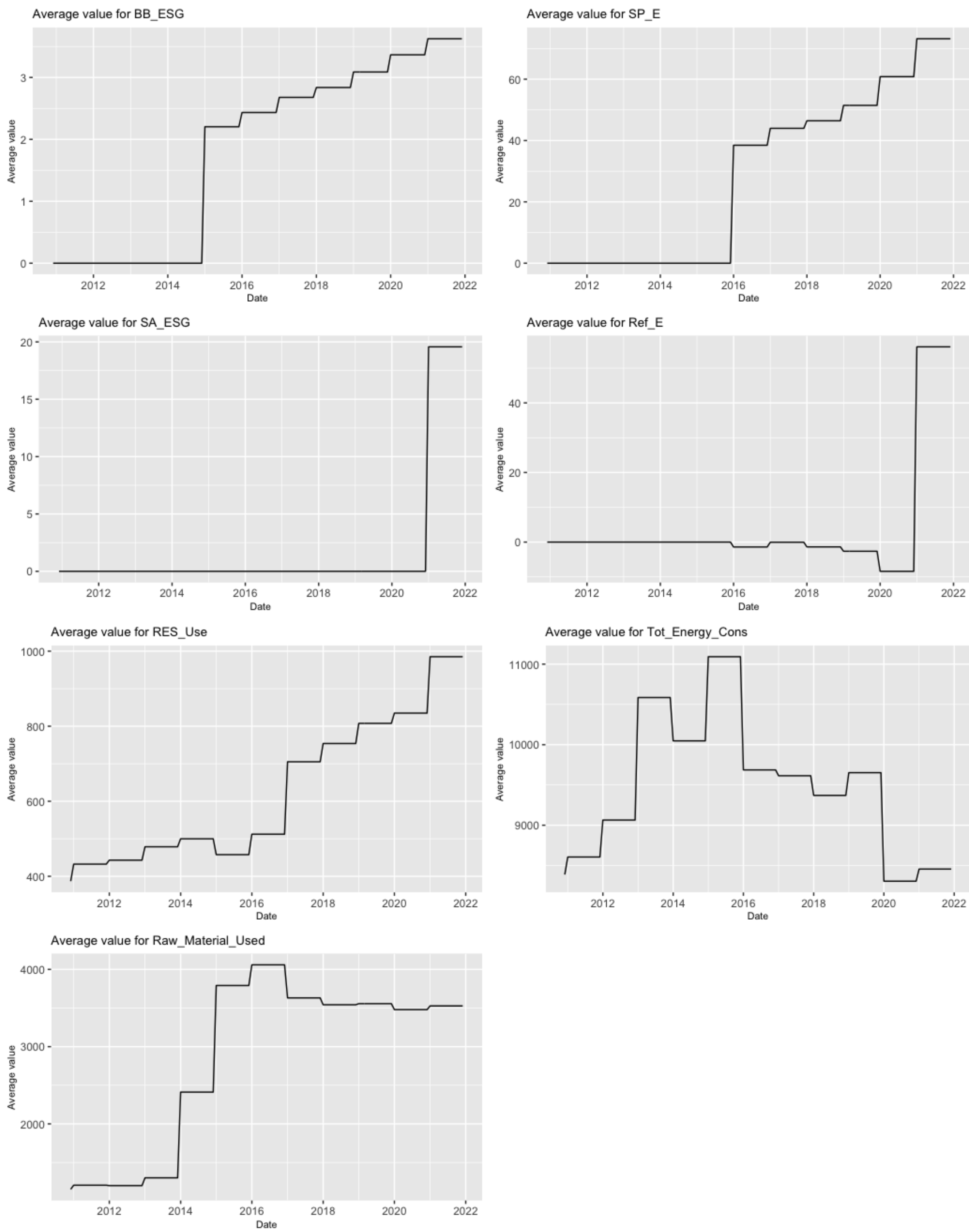
Companies with data on CDP_S1
Companies with data on CDP_Reg_Risk
Companies with data on CDP_S2
Companies with data on CDP_CC_Performance
Companies with data on RES_Use
Companies with data on Raw_Material_Used
Companies with data on Tot_Energy_Cons
Companies with data on Tot_Waste
# of companies
Date



A.1.2 Average reported value per variable

The charts below present the average value over time per variable in our sample. For the majority of the variables, the average value increases over the period analyzed, while for some variables (e.g. Tot_Energy_Cons) the average value has already reached its peak. The rising trend of most values can be linked to both higher consumption/emissions etc., or simply the fact that sustainability reporting is increasingly more common.


Average value for CDP_Rep_MtCO2e
Average value for Green_rev
Average value for Inv_Op_Sust
Average value for Inv_Sust_Prod
Average value for CDP_S1
Average value for CDP_S2
Average value for CDP_Reg_Risk
Average value for CDP_CC_Performance
Average value
Date



A.2 Validity tests

A.2.1 T-statistic

The t-statistic is employed to test the significance of the risk premiums and regression coefficients, to determine if they are statistically different from zero and thus provide meaningful compensation for the associated risks. The equation to calculate the t-statistic for testing risk premiums is as follows:

$$t = \frac{\hat{RP}}{\text{Standard Error}(\hat{RP})}$$

where t represents the t-statistic, $\hat{RP}$ denotes the estimated risk premium, and Standard Error($\hat{RP}$) is the standard error of the risk premium. To determine the significance of the estimated risk premium, we compare the calculated t-statistic to critical values from the t-distribution. For a significance level of 1% (a threshold often denoted as $\alpha = 0.01$), the corresponding critical value is approximately 2.576, and for a significance level of 5% (a threshold denoted as $\alpha = 0.05$), the critical value is approximately 1.960. The p-value associated with the t-statistic represents the probability of obtaining a test statistic as extreme as or more extreme than the observed value, assuming the null hypothesis is true. For a two-tailed test, the p-value is the probability of observing a t-statistic greater than the absolute value of the calculated t-statistic. If the p-value is less than the significance level (1% or 5%), the risk premium is considered statistically significant.

A.2.2 Stationarity - ADF Test

The Augmented Dickey-Fuller (ADF) test is performed to assess the stationarity of a time series dataset, i.e. that the mean and variance stay constant over time. The ADF test determines whether a unit root exists in the data, which indicates non-stationarity. After running the ADF test, we find that all variables are stationary. This suggests that no variable exhibit a trend or a systematic pattern of change over time, indicating that the relationships observed in our regressions will not be driven by long-term trends or spurious correlations arising from non-stationary variables. Instead, it implies that the relationships captured in your models are more likely to represent genuine and meaningful associations between the variables. We note that the ADF test only examines the stationarity of the individual variables and does not account for potential cointegration or other forms of interdependence between the variables. This is, however, tested using the Variance Inflation Factor (VIF) in Chapter 4.4. The test is conducted using an autoregressive model with lagged differences in the time series data. The equation for the ADF test is the following:

$$\Delta y_t = \alpha + \beta t + \gamma y_{t-1} + \sum_{i=1}^{p} \delta_i \Delta y_{t-i} + \varepsilon_t$$

Δy_t represents the differenced time series at time t, α is the intercept term, β represents the coefficient of a linear trend, and y_{t-1} is the lagged value of the time series. Additionally, γ represents the coefficient being tested for significance, and δ_i are the coefficients of the lagged differences of the time series. The significance of the coefficient γ is assessed to determine the presence of a unit root. If γ is statistically significant, it suggests the presence of a unit root and indicates non-stationarity in the data. On the other hand, if γ is not statistically significant, it suggests stationarity in the time series.

A.2.3 Omitted Variables

Omitted variable bias can lead to incorrect estimates of the relationships between the variables that are included in the model due to not including the correct variables. In this study, we run the risk of not having included enough control variables in some of our cross-sectional and time series regressions, which would lead to a misconception of the environmental metrics' effect on excess return. Additionally, we run the risk of not having included enough of the right E(SG) scores and environmental metrics to best explain excess return. As we have discussed previously, with ESG scores and environmental data being relatively new, this is a challenge we expected to face, and we believe it is difficult to overcome until the reporting standards are raised.

A.2.4 Selection Bias

Selection bias is especially likely to occur when dealing with ESG data. We are only able to observe and analyze the companies from the STOXX 600 that have the available data on ESG and various environmental measures for the time period we have chosen. Since ESG reporting was introduced in the late 2000s and became common practice only a few years ago, many companies do not have the data needed. This limits the sample to the share of companies of the STOXX 600 that have been reporting on ESG measures the longest. This may lead to a biased result if there is a specific reason these companies were the first ones to report on ESG and this reason affects the response variables of the sample.

A.2.5 Reverse Causality

Reverse causality is when the direction of causation between two variables is the opposite of what is assumed in a regression analysis. For example, a high stock return can lead to higher ESG ratings of a company because higher earnings allow the company to implement effective environmental initiatives, or the relationship could be the opposite, that higher ESG scores lead to higher earnings because consumers choose more environmentally friendly products. To address reverse causality, we are using lagged variables in our regressions, meaning that the dependent variable is one time period behind the independent variable(s). Additionally, by adding the analysis of the deltas (changes) of all the environmental metrics' effects on ESG scores and excess return, we further strengthen our understanding of the causality of our variables.

A.3 Hedging Climate Change News

A.3.1 Construction of Risk Factors

As explained by Engle et al. (2020), one of the conditions for the mimicking portfolio approach to isolate climate change risk is that the projection portfolios have to span all the risk factors driving returns. We therefore include in regression 4.10 three additional factors that might be correlated with climate risk and that are known to be important in explaining the cross-section of returns: size (using cross-sectionally standardized market value to create Z_t, so that half the firms, sorted by market value, have positive weight, and half have negative weight; note that this portfolio will be long large firms and short small firms), value (using cross-sectionally standardized values of book-to-market to create Z_t), and the market (setting Z_t to equal the share of total market value). The size and market factors are also used in the cross-sectional regression 4.5 in Chapter 4.2.

A.3.2 Construction of News Index

As previously mentioned, we obtain the news index used in the hedge from Engle et al. (2020) from Johannes Stroebel's website (NYU, n.d.). Here follows a more detailed explanation of how this monthly news index is constructed using textual analysis of the Wall Street Journal. The WSJ Climate Change News Index was constructed based on the desire to measure climate news relevant to investors and the availability of complete text access to The Wall Street Journal (WSJ) articles since the 1980s. To quantify the intensity of climate news coverage in the WSJ, the news content was compared to a collection of authoritative climate change texts, including white papers and glossaries. A Climate Change Vocabulary (CCV) was formed by aggregating the unique terms and their frequencies from these texts. The WSJ term counts were converted into tf-idf scores, highlighting representative terms. The CCV served as the reference for identifying climate change-related news. The daily WSJ editions were treated as documents, and tf-idf scores were calculated for each edition. The cosine similarity between the tf-idf scores of the CCV and each daily WSJ edition was used to construct the WSJ Climate Change News Index, representing the fraction of the WSJ dedicated to climate change topics. The index was scaled for interpretability, and its values indicated the magnitude of innovations in the index. Monthly hedge targets were derived by averaging the daily values and obtaining residuals from an autoregressive AR(1) model, resulting in the $CC^{WSJ}{}_{t}$ series capturing monthly innovations in the WSJ Climate Change News Index.

Appendix B

Tables

B.1 Fama MacBeth Risk Premiums

This section displays the complete Fama MacBeth two-step regression results for all the statistically significant environmental variables: BB_ESG, Tot_Waste, CDP_CC_Performance, CDP_S1, and CDP_S2.

Table B.1: BB ESG: Results from Fama MacBeth regression

Factor	Risk_Premium	T_Stat	Std_Error	R^2	R^2_Adj
BB_ESG	15.864	26.49627 (∗∗∗)	0.599	0.627	0.626
LOGSIZE	-7.406	-53.89519 (∗∗∗)	0.137	0.874	0.874
BM	-1.514	-12.30583 (∗∗∗)	0.123	0.267	0.266
LEVERAGE	5.723	2.733002 (∗∗)	2.094	0.018	0.015
ROE	10.833	0.663626	16.323	0.001	-0.001
MOM	-0.804	-1.173642	0.685	0.003	0.001
INVEST	0.001	0.3258314	0.003	0.0003	-0.002
BETA	16.255	1.629693	9.974	0.006	0.004
LOGPPE	-7.466	-15.60283 (∗∗∗)	0.478	0.373	0.372
VOLAT	-25.125	-17.59577 (∗∗∗)	1.428	0.429	0.428
SALESGR	4.637	1.390804	3.334	0.005	0.002
EPSGR	-0.417	-5.628899 (∗∗∗)	0.074	0.072	0.069

Table B.2: Tot Waste: Results from Fama MacBeth regression

Factor	Risk_Premium	T_Stat	Std_Error	R^2	R^2_Adj
Tot_Waste	−0.198	-0.269755	0.735	0.0002	−0.003
LOGSIZE	−10.399	-192.1027 (∗∗∗)	0.054	0.990	0.990
BM	−3.721	-42.19782 (∗∗∗)	0.088	0.837	0.836
LEVERAGE	−0.040	-0.01978551	2.044	0.00000	−0.003
ROE	13.784	1.943258	7.093	0.012	0.009
MOM	4.786	1.138314	4.204	0.004	0.001
INVEST	0.003	1.830405	0.002	0.011	0.007
BETA	75.696	1.071619	70.637	0.003	0.000
LOGPPE	−7.942	-39.41569 (∗∗∗)	0.202	0.838	0.837
VOLAT	−10.120	-0.6540667	15.472	0.001	−0.002
SALESGR	−166.557	-4.488871 (∗∗∗)	37.104	0.053	0.051
EPSGR	0.208	0.5494623	0.379	0.001	−0.002

Table B.3: CDP CC: Results from Fama MacBeth regression

Factor	Risk_Premium	T_Stat	Std_Error	R^2	R^2_Adj
CDP_CC	−4.424	-2.208219 (∗∗)	2.003	0.018	0.014
LOGSIZE	−8.100	-15.82151 (∗∗∗)	0.512	0.411	0.410
BM	−1.156	-7.725887 (∗∗∗)	0.150	0.143	0.141
LEVERAGE	−0.045	-0.004226051	10.654	0.00000	−0.003
ROE	−8.523	-0.5000929	17.043	0.001	−0.002
MOM	−0.136	-0.9147738	0.148	0.002	−0.001
INVEST	0.003	1.532317	0.002	0.007	0.004
BETA	−3.024	-1.089938	2.774	0.003	0.001
LOGPPE	−6.385	-13.66933 (∗∗∗)	0.467	0.344	0.342
VOLAT	−1.811	-4.850926 (∗∗∗)	0.373	0.062	0.059
SALESGR	−1.408	-1.065594	1.322	0.003	0.000
EPSGR	−0.006	-0.06438437	0.091	0.00001	−0.003

Table B.4: CDP S1: Results from Fama MacBeth regression

Factor	Risk_Premium	T_Stat	Std_Error	R^2	R^2_Adj
CDP_S1	0.010	1.565537	0.007	0.006	0.004
LOGSIZE	−8.934	-34.03566 (∗∗∗)	0.262	0.734	0.733
BM	−1.208	-14.18137 (∗∗∗)	0.085	0.330	0.329
LEVERAGE	74.730	4.703436 (∗∗∗)	15.888	0.052	0.049
ROE	0.721	0.09273699	7.775	0.00002	−0.003
MOM	0.634	1.155373	0.548	0.003	0.001
INVEST	−0.0004	-1.319321	0.0003	0.005	0.002
BETA	4.252	0.693944	6.127	0.001	−0.001
LOGPPE	−6.864	-29.98907 (∗∗∗)	0.229	0.715	0.714
VOLAT	−9.036	-7.71582 (∗∗∗)	1.171	0.125	0.123
SALESGR	1.502	0.3340769	4.496	0.0003	−0.002
EPSGR	−0.032	-0.5946181	0.055	0.001	−0.002

Table B.5: CDP S2: Results from Fama MacBeth regression

Factor	Risk_Premium	T_Stat	Std_Error	R^2	R^2_Adj
CDP_S2	−0.012	-1.21498	0.010	0.005	0.002
LOGSIZE	−9.559	-13.12631 (∗∗∗)	0.728	0.328	0.326
BM	−1.486	-6.736756 (∗∗∗)	0.221	0.114	0.112
LEVERAGE	−2.526	-0.1374254	18.379	0.0001	−0.003
ROE	−98.578	-4.432184 (∗∗∗)	22.241	0.053	0.051
MOM	−0.109	-0.507427	0.214	0.001	−0.002
INVEST	0.001	0.3713813	0.004	0.0004	−0.002
BETA	0.914	0.2534254	3.608	0.0002	−0.003
LOGPPE	−7.368	-11.67841 (∗∗∗)	0.631	0.281	0.279
VOLAT	−1.831	-3.293108 (∗∗)	0.556	0.030	0.027
SALESGR	−1.126	-0.545121	2.065	0.001	−0.002
EPSGR	0.126	1.207225	0.104	0.004	0.001

Table B.6: Tot Energy Cons: Results from Fama MacBeth regression

Factor	Risk_Premium	T_Stat	Std_Error	R^2	R^2_Adj
Energy_Cons	−0.786	-1.007492	0.780	0.002	0.000
LOGSIZE	−9.640	-71.04008 (∗∗∗)	0.136	0.920	0.920
BM	−3.912	-32.69445 (∗∗∗)	0.120	0.714	0.713
LEVERAGE	−5.638	-2.506023 (∗)	2.250	0.015	0.012
ROE	47.409	5.244857 (∗∗∗)	9.039	0.064	0.062
MOM	−2.345	-2.548598 (∗)	0.920	0.015	0.012
INVEST	−0.005	-5.026006 (∗∗∗)	0.001	0.063	0.060
BETA	13.688	1.109364	12.339	0.003	0.001
LOGPPE	−4.968	-15.89396 (∗∗∗)	0.313	0.410	0.408
VOLAT	2.561	0.8570945	2.988	0.002	−0.001
SALESGR	2.210	0.8640049	2.558	0.002	−0.001
EPSGR	0.010	0.3044305	0.033	0.0002	−0.002

B.2 Delta

This table shows the regression results from the multiple linear regression studying the effects of the 12-months change (delta) in environmental metrics on E(SG) scores in time (t+1), with corresponding T-statistics. The R-squared of these regressions on the bottom of the table tells us how much of the E(SG) score is explained by the change in environmental metrics.

Table B.7: MLR: Delta Environmental Metrics on ESG Scores

Metric	Sustainalytics ESG	Bloomberg ESG	Refinitiv E	S&P E	CDP CC
CDP_S1	-0.01	0.00	-0.00	-0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
CDP_S2	0.00	0.00	0.02***	0.00	0.13***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
CDP_Reg_Risk	-0.01*	0.03***	0.07***	0.02***	0.02***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
CDP_Rep_MtCO2e	0.00	0.02***	-0.01**	-0.01**	0.02***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Green_rev	0.06***	-0.00	0.04***	-0.00	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Inv_Op_Sust	-0.01*	-0.01***	0.02***	0.01***	-0.01
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Inv_Sust_Prod	0.02***	0.02***	-0.01*	-0.00	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
RES_Use	0.02***	-0.03***	0.02***	0.01***	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Tot_Energy_Cons	-0.00	-0.00	-0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Raw_Material_Used	0.00	-0.00	0.00	-0.00	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Tot_Waste	-0.01*	-0.01	-0.01***	-0.01***	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
CO2intensity	-0.03***	-0.00	-0.00	-0.00	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Excess_Return	0.30***	0.21***	1.03***	0.21***	0.40***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
R-squared	0.01	0.00	0.02	0.00	0.02
Adjusted R-squared	0.01	0.00	0.02	0.00	0.02
F-statistic	43.73	26.41	146.70	10.28	133.20
DF	13 and 79876	13 and 79876	13 and 79876	13 and 79876	13 and 79876

Bibliography

A. van Dijk, P. d. M. (2020). Assessing climate risk for investment portfolios. https://www.rsm.nl/fileadmin/Faculty-Research/Centres/EPSVC/Assessing_climate_risk_for_investment_portfolios__oct20.pdf

Andersson, M., Bolton, P., & Samama, F. (2016). Hedging Climate Risk [Publisher: Routledge _eprint: https://doi.org/10.2469/faj.v72.n3.4]. *Financial Analysts Journal*, *72*(3), 13–32. https://doi.org/10.2469/faj.v72.n3.4

Berg, F., Kölbel, J. F., & Rigobon, R. (2022). Aggregate Confusion: The Divergence of ESG Ratings*. *Review of Finance*, *26*(6), 1315–1344. https://doi.org/10.1093/rof/rfac033

Bolton, P., & Kacperczyk, M. (2021). Do investors care about carbon risk? *Journal of Financial Economics*, *142*(2), 517–549. https://doi.org/https://doi.org/10.1016/j.jfineco.2021.05.008

Chatterji, A. K., Durand, R., Levine, D. I., & Touboul, S. (2016). Do ratings of firms converge? Implications for managers, investors and strategy researchers [_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/smj.2407]. *Strategic Management Journal*, *37*(8), 1597–1614. https://doi.org/10.1002/smj.2407

Eccles, R. G., Serafeim, G., & Krzus, M. P. (2011). Market Interest in Nonfinancial Information [_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1745-6622.2011.00357.x]. *Journal of Applied Corporate Finance*, *23*(4), 113–127. https://doi.org/10.1111/j.1745-6622.2011.00357.x

Engle, R. F., Giglio, S., Kelly, B., Lee, H., & Stroebel, J. (2020). Hedging Climate Change News (A. Karolyi, Ed.). *The Review of Financial Studies*, *33*(3), 1184–1216. https://doi.org/10.1093/rfs/hhz072

EU Action Plan on Sustainable Finance. (2023). Retrieved June 28, 2023, from https://www.securities-services.societegenerale.com/en/insights/views/news/eu-action-plan-sustainable-finance/

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

Fama, E. F., & MacBeth, J. D. (1973). Risk, Return, and Equilibrium: Empirical Tests [Publisher: The University of Chicago Press]. *Journal of Political Economy*, *81*(3), 607–636. https://doi.org/10.1086/260061

Ferson, W. E., & Harvey, C. R. (1991). The Variation of Economic Risk Premiums [Publisher: University of Chicago Press]. *Journal of Political Economy*, *99*(2), 385–415. Retrieved June 5, 2023, from https://www.jstor.org/stable/2937686

Gibson Brandon, R., Krueger, P., & Schmidt, P. S. (2021). ESG Rating Disagreement and Stock Returns [Publisher: Routledge _eprint: https://doi.org/10.1080/001519 *Financial Analysts Journal*, *77*(4), 104–127. https://doi.org/10.1080/0015198X.2021.1963186

Grundström, G., & Miedel, I. (2021). *Sustainable Investing : On the relation between sustainability rating and greenhouse gas emissions*. Retrieved May 23, 2023, from https://urn.kb.se/resolve?urn=urn:nbn:se:umu:diva-185219

In, S. Y., Park, K. Y., & Monk, A. (2019). Is 'Being Green' Rewarded in the Market?: An Empirical Investigation of Decarbonization and Stock Returns. Retrieved May 28, 2023, from https://papers.ssrn.com/abstract=3020304

LaBella, M. J., Sullivan, L., Russell, J., & Novikov, D. (2019). The Devil is in the Details [_eprint: https://www.firstlinks.com.au/uploads/Whitepapers/lm-qs-the-devil-is-in-the-details-0919.pdf].

Lakonishok, J., Shleifer, A., & Vishny, R. W. (1994). Contrarian Investment, Extrapolation, and Risk [_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1540-6261.1994.tb04772.x]. *The Journal of Finance*, *49*(5), 1541–1578. https://doi.org/10.1111/j.1540-6261.1994.tb04772.x

Lo, A. W., & MacKinlay, A. C. (1990). When are Contrarian Profits Due to Stock Market Overreaction? [Publisher: [Oxford University Press, Society for Financial Studies]]. *The Review of Financial Studies*, *3*(2), 175–205. Retrieved June 5, 2023, from https://www.jstor.org/stable/2962020

NYU. (n.d.). Johannes stroebel. Retrieved June 21, 2023, from https://pages.stern.nyu.edu/~jstroebe/

Pástor, Ľ., Stambaugh, R. F., & Taylor, L. A. (2021). Sustainable investing in equilibrium. *Journal of Financial Economics*, *142*(2), 550–571. https://doi.org/10.1016/j.jfineco.2020.12.011

Pollard, J., & Sherwood, M. (2018). Establishing ESG as Risk Premia. *The Journal Of Investment Management*, *16*.

Soebhag, A., & Vliet, B. V. (2022). Mind Your Sorts.